

Serialization

The process of transforming data from an in-memory representation into a form suitable for storage on disk or transport over the network

Plain text

The closing price for GOOG on Jan 2, 2014 was \$1113.12 and 1,821,400 shares were traded.

Binary encoding

0x10AB04360001B2D0001BCAD8

Encoding (made up):

Day:	5 bits	
Month:	4 bits	
Year:	7 bits	(X-1970)
Symbol:	16 bits	
(1078 th in alphabetic order)		
Closing price:	32 bits	(fixed point)
Shares traded:	32 bits	(unsigned)

Context matters.

CSV

2,Jan,2014,GOOG,1113.12,1821400

XML

```
<stock_info>  
  <symbol>GOOG</symbol>  
  <date>  
    <day>2</day>  
    <month>Jan</month>  
    <year>2014</year>  
  </date>  
  <closing_price>1113.12</closing_price>  
  <shares_traded>18121400</shares_traded>  
</stock_info>
```

JSON

```
{"symbol": "GOOG",  
  "date": {"day": "2",  
            "month": "Jan",  
            "year": "2014"},  
  "closing price": "1113.12",  
  "shares traded": "18121400"}
```

HTML (partial)

```
<tr>  
  <td class="yfnc_td1" align="right">Jan 2,  
2014</td>  
  <td class="yfnc_td1" align="right">1,113.12</  
td>  
  <td class="yfnc_td1" align="right">1,821,400</  
td>  
</tr>
```

URLs

Protocol: //site address/path/filename

Site address: hostname.domain

<http://www.classes.cs.uchicago.edu/archive/2016/winter/12200-2/pa/pa1/index.html>

Hostname: www

Domain name: classes.cs.uchicago

Top-level domain: edu

Path: archive/2016/winter/12200-2/pa/pa1

Filename: index.html


```
urllib.parse.urlparse(url)
ParseResult(
    scheme='http',
    netloc='www.classes.cs.uchicago.edu',
    path='/archive/2016/winter/12200-2/pa/pa1/index.html',
    params="", query="", fragment="")
```

Basic structure of HTML document

```
<!DOCTYPE HTML>
```

```
<html>
```

```
<head>
```

```
<title>...</title>
```

```
</head>
```

```
<body>
```

```
<p> ... </p>
```

```
<p> ... </p>
```

```
<p> ... </p>
```

```
</body>
```

```
</html>
```

Some example tags

headers:

<h1>...</h1> through <h6>...</h6>

paragraph:

<p> ... some text ... </p>

<p class="courseblocktitle">course title...</p>

links:

<a href="<http://college.uchicago.edu>">The College

images:

More example tags

comment:

`<!-- ... -->`

division or section:

`<div> ... </div>`

`<div class="courseblock main">...stuff</div>`

HTML Tables

```
<table>
  <tr>
    <th>...</th>
    ...
    <th>...</th>
  </tr>
  <tr>
    <td>...</td>
    ...
    <td>...</td>
  </tr>
</table>
```