

The WebVTT format

The [W3C Web Media Text Tracks Community Group](#) is working on the "[WebVTT: The Web Video Text Tracks Format](#)" specification, a format used for defining files that will contain text for captions and subtitles, and much more... The WebVTT files are used with the `src` attribute of the `<track>` element, that can be used inside a `<video>...</video>`.

In the example presented in the previous section (located at <http://jsbin.com/luyazi/2/edit>), we used a file called `sintel-captions.vtt`:

```
<video height="272" width="640"
  poster="http://content.bitsontherun.com/thumbs/q1fx20VZ-640.jpg"
  crossorigin="anonymous"
  controls>
  ...
  <track src="http://demo.jwplayer.com/html5-report/sintel-captions.vtt"
7.    kind="captions" label="Closed Captions" default>
</video>
```

And here is an extract of the corresponding `sintel-captions.vtt` file:

```
WEBVTT
00:00:01.000 --> 00:00:02.042
(drumbeat)
00:00:07.167 --> 00:00:12.025
(plaintive violin solo playing)
00:00:15.000 --> 00:00:18.183
10. (wind whistling)
00:00:24.167 --> 00:00:27.025
(orchestra music swells)
00:00:43.033 --> 00:00:43.192
```

(weapons clash)

00:00:44.000 --> 00:00:44.175

(gasps)

20.

00:00:44.183 --> 00:00:45.158

(grunts)

00:00:45.167 --> 00:00:47.058

(groaning)

00:00:54.192 --> 00:00:55.150

(blade rings)

30.

00:00:55.158 --> 00:00:57.008

(bellowing)

00:00:57.017 --> 00:00:58.067

(grunting)

00:00:59.075 --> 00:01:00.133

(panting)

00:01:05.108 --> 00:01:06.125

40.

(cries out in agony)

00:01:08.050 --> 00:01:09.058

(panting)

00:01:12.092 --> 00:01:13.142

(panting)

00:01:14.017 --> 00:01:18.125

(orchestra plays ominous low notes)

50.

00:01:31.058 --> 00:01:35.133

(plaintive violin solo returns)

00:01:46.158 --> 00:01:49.058

This blade has a dark past.

00:01:51.092 --> 00:01:54.108

It has shed much innocent blood.

60.

00:01:57.083 --> 00:02:00.000

You're a fool for traveling alone

so completely unprepared.

00:02:01.100 --> 00:02:03.033

You're lucky your blood's still flowing.

00:02:04.183 --> 00:02:06.075

Thank you.

This format is rather simple, but we still recommend reading [this excellent article from Mozilla Developer Network](#) that explains in details all the different options.

Each "element" in this file has a starting and ending time, plus a value (the text that will be displayed), followed by a blank line (blank lines are separators between elements).

Each element is called "a cue", and may optionally have an ID that will be useful when using the track element JavaScript API (studied in the HTML5 part 2 course), in particular the `getCueById()` method of `TextTrack` objects.

Example of numeric IDs:

```
9
00:00:21.000 --> 00:00:22.000
to hear from <u>you</u>
10
00:00:22.500 --> 00:00:25.000
We want to hear what inspires you as a developer
```

IDs may be also defined as strings, and values can use HTML as well:

```
Opening
00:00:00.000 --> 00:00:30.000
Welcome to our <i>nice film</i>
```

The displayed text can span over multiple lines, but blank lines are not allowed, as they would be interpreted as a separator: