



TTML vs. WebVTT

XML considered harmful for web captions?

Andreas Tai,
Institut für Rundfunktechnik



Institut für Rundfunktechnik



European Broadcasting Union

- EBU Group: Subtitles in XML

Subtitles == Caption

In this presentation!



Context

Context: Subtitle information



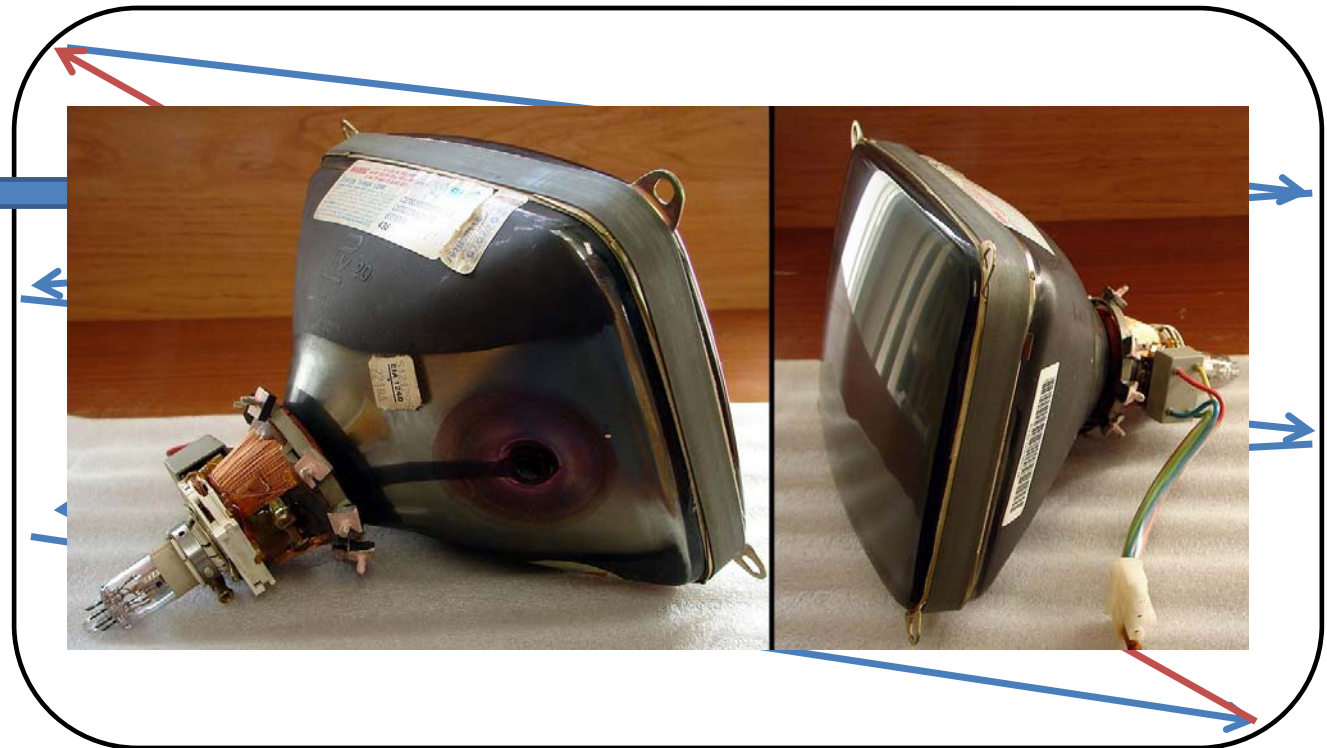
- ½ tweet

+ Formatting

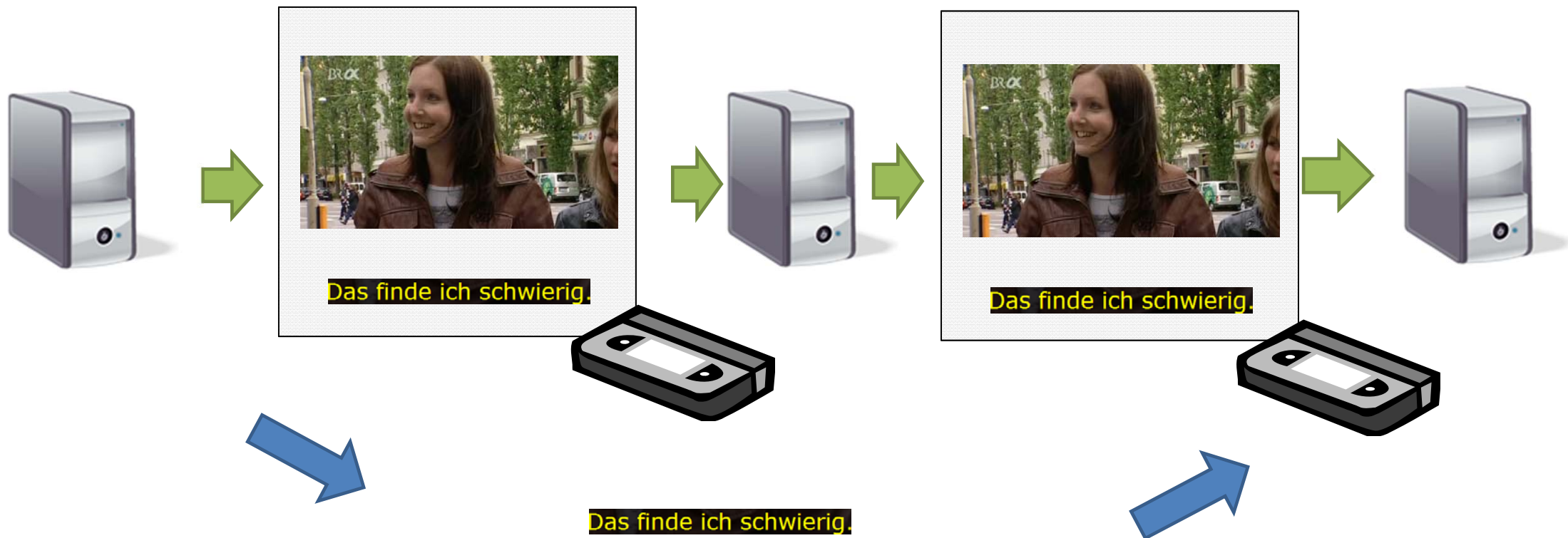
+ Timing

Context: Digital Data in the analog TV signal

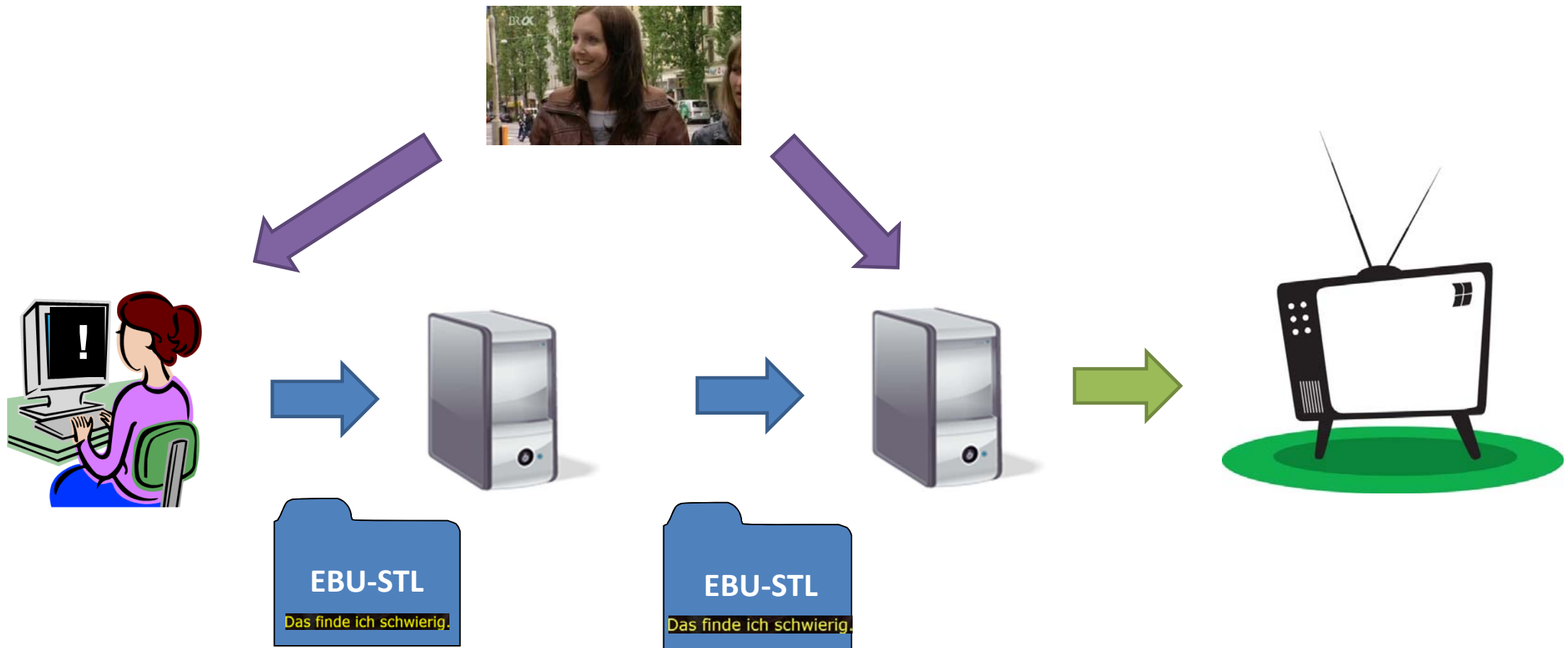
Subtitles in the VBI (Vertical Blank Interval)



Context: Subtitle transmission in TV production



Context: Subtitle transmission in TV production



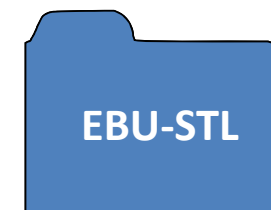
Context: EBU STL – Binary Subtitle Format

Byte	Bytes allocated	Information
0	1	Subtitle Group Number
1..2	2	Subtitle Number
3	1	Extension Block Number
4	1	Cumulative Status
5..8	4	Time Code In
9..12	4	Time Code Out
13	1	Vertical Position
14	1	Justification Code
15	1	Comment Flag
16..127	112	Text Field

EBU-STL

Context: EBU STL – Binary Subtitle Format

		First byte →															
Second byte ↓		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0				SP	0	@	P	`	p			NBSP	°		—	Ω	κ
1				!	1	A	Q	a	q			ı	±	`	ı	Æ	æ
2				"	2	B	R	b	r			¢	²	'	®	Ð	ð
3				#	3	C	S	c	s			£	³	^	©	å	ö
4				¤	4	D	T	d	t			\$	×	~	™	Ë	ë
5				%	5	E	U	e	u			¥	µ	ˉ	♪		ı
6				&	6	F	V	f	v				¶	˘	¬	İ	ij
7				'	7	G	W	g	w			§	•	˙	ı	Ł	ł
8				(8	H	X	h	x				÷	˝		Ł	ł
9)	9	I	Y	i	y			`	'			Ø	ø
A				*	:	J	Z	j	z			"	"	°		Œ	œ
B				+	;	K	[k	{			«	»	ˆ		ø	ß
C				,	<	L	\	l				←	¼	—	½	þ	þ
D				-	=	M]	m	}			↑	½	"	¾	ƒ	ƒ
E				.	>	N	^	n	~			→	¾	˘	¾	Ŋ	ŋ
F				/	?	O	_	o				↓	¿	˘	¾	ŋ	SHY

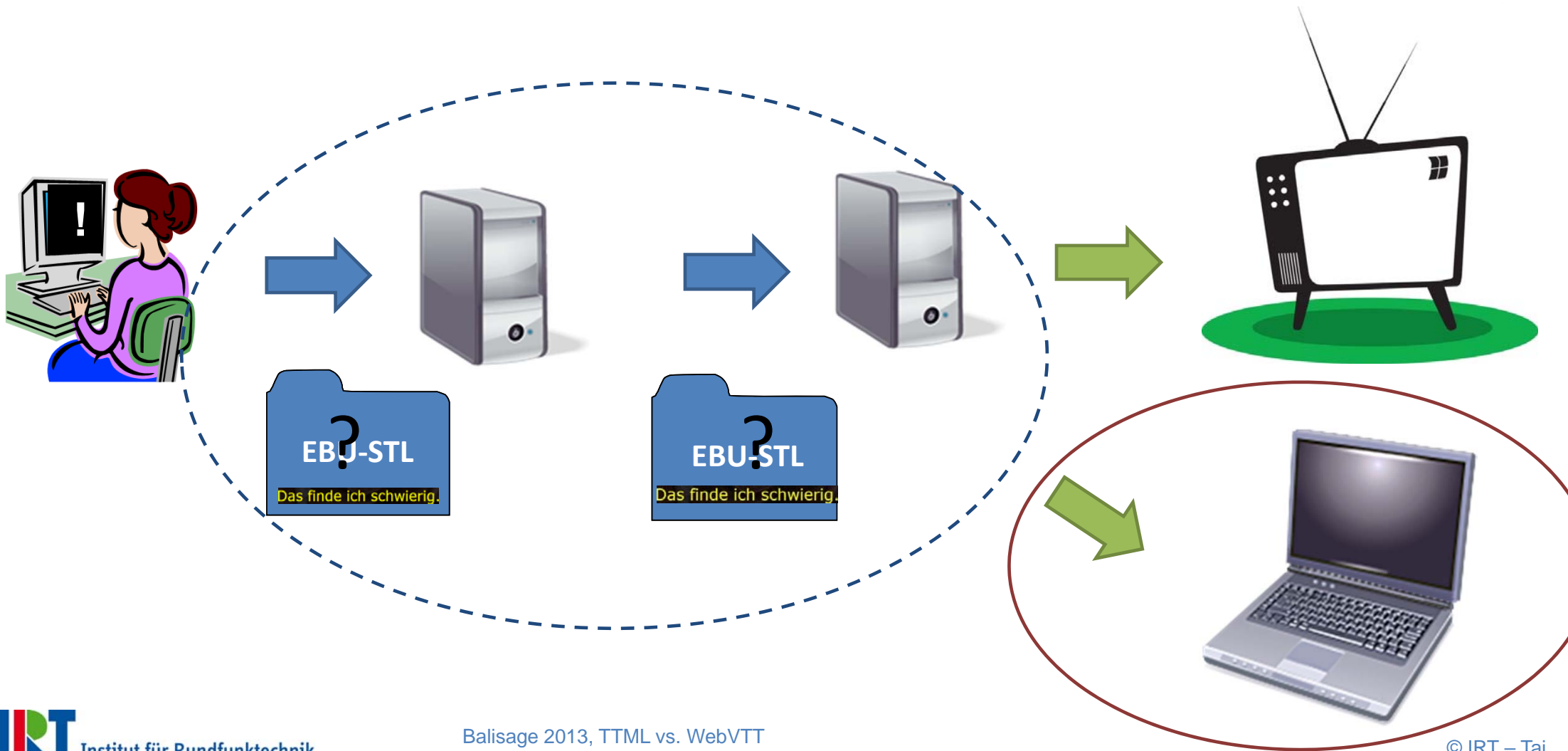


Limitation

- Formatting
- File size
- Character Sets

ISO 6937/2-1983

Context: Requirement for a new format





TTML vs. WebVTT

The match

TTML



2002-2010, 2012 - *

Requirements?

- Standard (!) to synchronise text with media
- XML

Who?





WebVTT

WebVTT: Specification process



WHATWG

2010 -2011



2011 - *

WebVTT: Specification process

Who?



Requirements

- Subtitle Format for media element in HTML5
- Keep it simple

```
<p begin="00:00:00.000" end="00:00:02.000"  
  xmlns="http://www.w3.org/ns/ttml">
```

This is a subtitle

on two lines

```
</p>
```

WebVTT: Example 1

00:00:00.000 --> 00:00:02.000 U+000A

This is a subtitle U+000A

on two lines U+000A

TTML: Example 2

...

```
<tt:style xml:id="s1" tts:color="green"/>
```

...

```
<p xml:id="sub1" begin="00:00:00.000" end="00:00:02.000">
```

This is a green word..

This is an italic word.

```
</p>
```

...

WebVTT: Example 2

sub1

00:00:00.000 --> 00:00:02.000 **size:80%** **position:10%** **line:70%** **align:middle**
<v speaker>This is a green <c.s1>word</c>.
This is an italic <i>word</i>.

U+0020
↓



Formal Model TTML

<p

xml:id = IDREFS

begin = <timeExpression>

end = <timeExpression> >

Content: (#PCDATA | br | span)*

</p>

Reduced info set

- Infoset
- DOM
- XDM



Formal Model WebVTT

Definition in prose

- A WebVTT timestamp representing the start time offset of the cue.
- The string "-->" (U+002D HYPHEN-MINUS, U+002D HYPHEN-MINUS, U+003E GREATER-THAN SIGN).
- A WebVTT timestamp representing the end time offset of the cue.

Model WebVTT: Interface

IDL

```
enum AutoKeyword { "auto" };  
enum DirectionSetting { "" /* horizontal */, "rl", "lr" };  
enum AlignSetting { "start", "middle", "end", "left", "right" };  
[Constructor(double startTime, double endTime, DOMString text)]  
interface VTT Cue : TextTrackCue {  
    attribute DOMString regionId;  
    attribute DirectionSetting vertical;  
    attribute boolean snapToLines;  
    attribute (long or AutoKeyword) line;  
    attribute long position;  
    attribute long size;  
    attribute AlignSetting align;  
    attribute DOMString text;  
    DocumentFragment getCueAsHTML();  
};
```

Model WebVTT: Parsing algorithm

.....

- Let input be the string being parsed.
- Let position be a pointer into input, initially pointing at the start of the string.
- Skip whitespace.
- Collect a WebVTT timestamp. If that algorithm fails, then abort these steps and return failure. Otherwise, let cue's text track cue start time be the collected time.
- Skip whitespace.
- If the character at position is not a U+002D HYPHEN-MINUS character (-) then abort these steps and return failure. Otherwise, move position forwards one character.

....

WebVTT: Parsed Data Structure

```
{  
  "id": "",  
  "startTime": 0,  
  "endTime": 2,  
  "pauseOnExit": false,  
  "direction": "horizontal",  
  "snapToLines": true,  
  "linePosition": "auto",  
  "textPosition": 50,  
  "size": 100,  
  "alignment": "middle",  
  "text": "This is a subtitle\nnon two lines",  
  "tree": {"children": [{  
    "type": "text",  
    "value": "This is a subtitle\nnon two lines"  
  }]}  
}
```

WebVTT: Parsed Data Structure

```
"tree": {"children": [  
  {  
    "type": "text",  
    "value": "A bold "  
  },  
  {  
    "type": "object",  
    "name": "b",  
    "classes": [],  
    "children": [{  
      "type": "text",  
      "value": "word"  
    }]  
  },  
  {  
    "type": "text",  
    "value": " on one line"  
  }  
]}
```



Validation

- TTML W3C XML Schema
- TTML Relax NG
- ttv - TTML Verification Tools

Not normative!

```
<p ... tts:textAlign="left">
```

One line aligned to the left.

```
</p>
```

Validation: TTML XSD sample

```
<xs:attribute name="textAlign">  
  <xs:simpleType>  
    <xs:restriction base="xs:token">  
      <xs:enumeration value="left"/>  
      <xs:enumeration value="center"/>  
      <xs:enumeration value="right"/>  
    </xs:restriction>  
  </xs:simpleType>  
</xs:attribute>
```

00:00:00.000 --> 00:00:02.000 align:left

One line aligned to the left.

- The string "align" as the WebVTT cue setting name.
- A U+003A COLON character (:).
- One of the following strings as the WebVTT cue setting value: "start", "middle", "end", "left", "right"

- If name is a case-sensitive match for "align"
 - If value is a case-sensitive match for the string "left", then let cue's text track cue alignment be left alignment.
 -

WebVTT: Live Validator

Live [WebVTT](#) Validator

WEBVTT

```
00:00:00.000 --> 00:00:02.000 vertical: rl
.enil eno no eltitbus A
```

Kind: subtitles/captions/descriptions ▼

Almost there! (1ms)

1. Line 3, column 31: No value for setting defined.

<http://quuz.org/webvtt/>



Validation: WebVTT File Format & Parsing

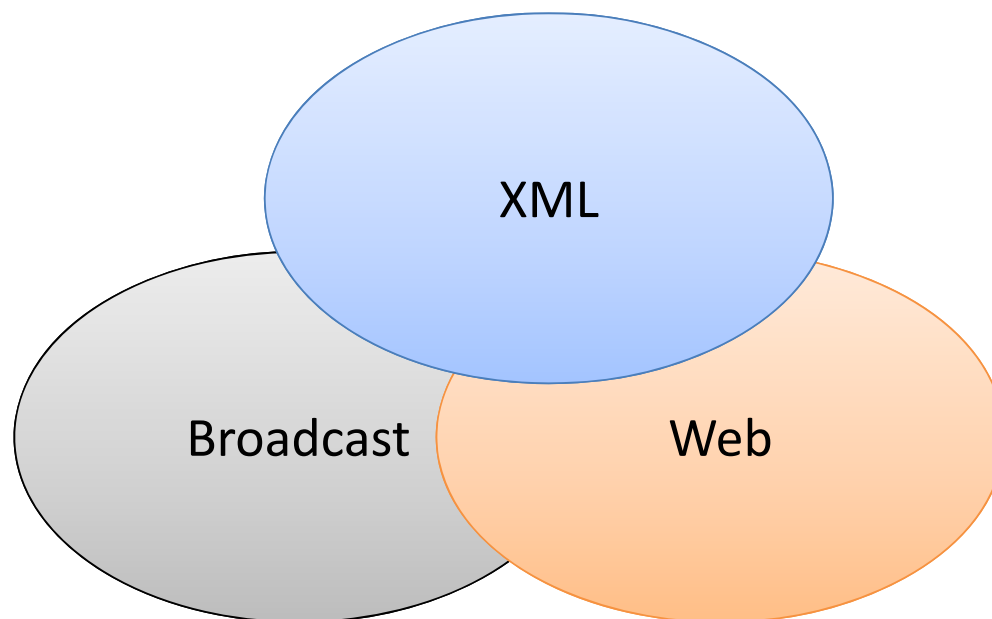
“The file format tells you what's valid--what authors should be writing.

The parser tells you how to deal with every possible input, which includes error handling (inputs that don't follow the file format) and [...] forwards-compatibility.”

Glenn Maynard, 8.3.2013, public-texttracks mailing list



Communities



TTML – Adoption



SMPTE-TT

Broadcast

EBU

EBU-TT



CFF-TT

- Semantic completeness
- Structured Content
- Profiling
- XSD as a contract



Broadcast

Communities: Adoption TTML?



~~TTML~~

- Namespaces
- Error handling
- Reference of XSL:FO

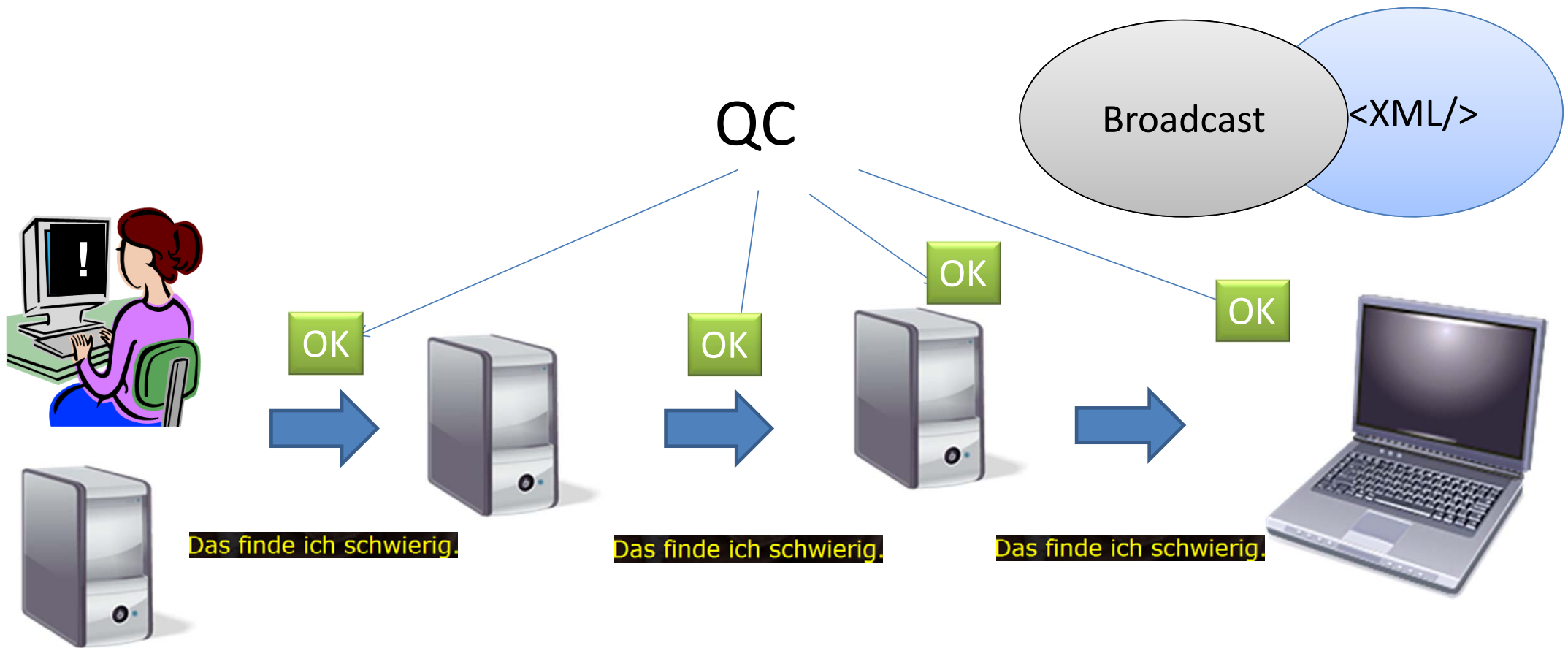
Communities: Adoption WebVTT

Browser
Manufacturer

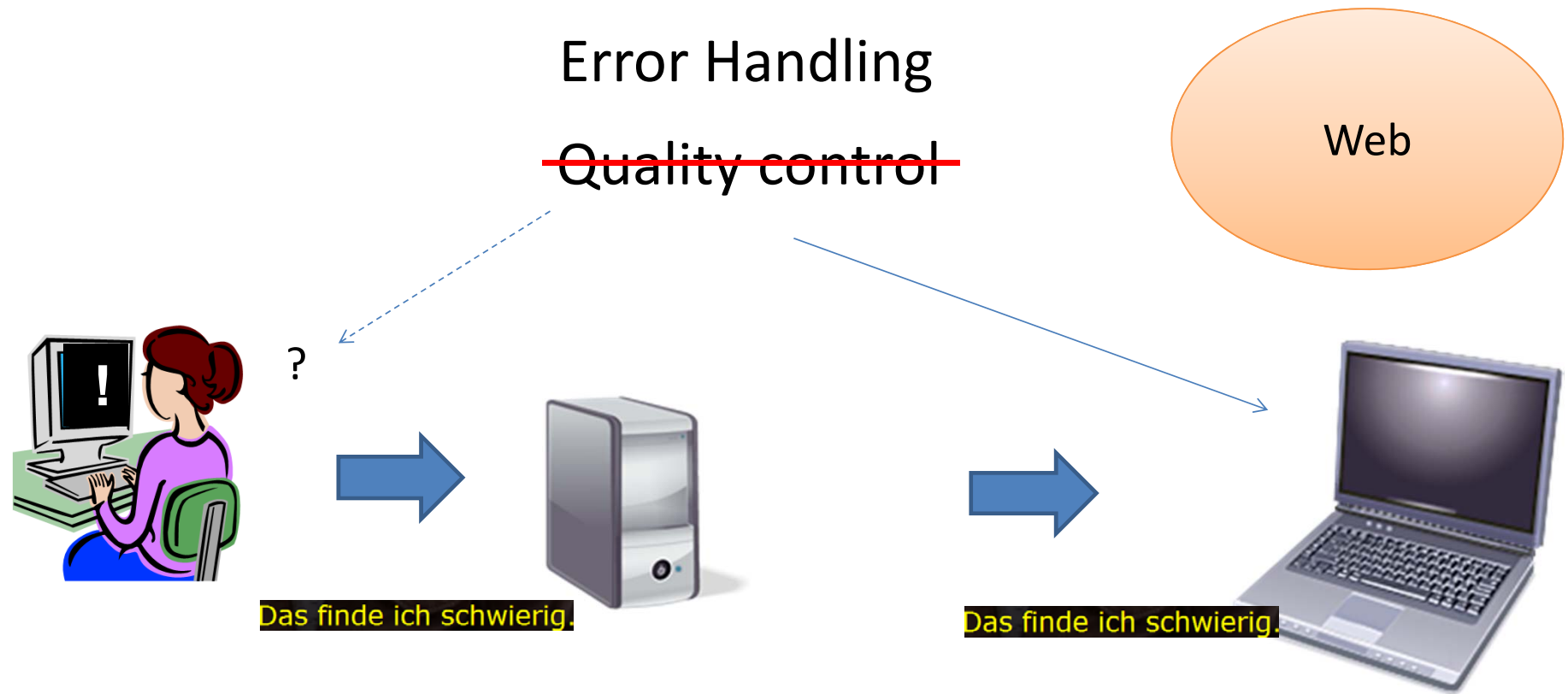


Web

Communities: Quality Control



Communities: Quality Control



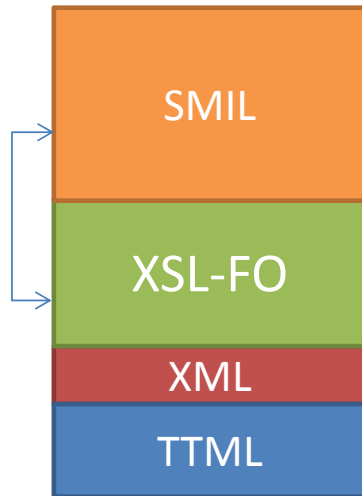
Communities: Innovation



Communities: Innovation

<XML/>

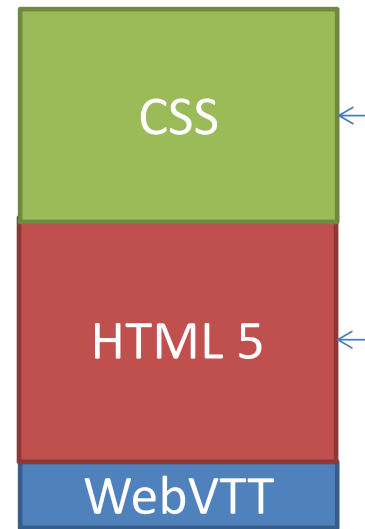
Domain
specific



TTML

Web

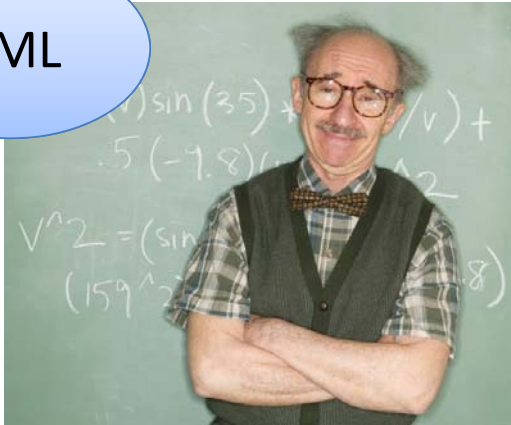
Well known



WebVTT

Communities: Readability

XML



Web



<p **begin**="00:00:00.000" **end**="00:00:02.000">

This is a subtitle

on two lines

</p>

00:00:00.000 --> 00:00:02.000

This is a subtitle

on two lines

Communities: Readability

EBU-STL

001101



Broadcast



Conclusion



Situation now?

W3C TPAC, November 2013



TTWG

&

WebVTT CG



Lessons to be learned?

- Standards as a service
- Accessibility of standards
- Background of the audience