

Introductie EBU Timed Text

Waarom, Waar, Hoe

SUBTITLE



Dutch Guild 13 April 2021 - Marco Slik

Agenda

- Een klein beetje geschiedenis
- De huidige status van de ondertitel markt
- Het waarom, waar en hoe van EBU-TT
- De EBU-TT standard suite
- De noodzaak voor een 'NL joint profile'



Een klein beetje geschiedenis

- 1974 - BBC World Standard Teletext
- 1980 - NPO start Doven en Slechthorenden ondertiteling 888 (Vertalende ondertiteling dan d.m.v. caption keying van ponskaarten)
- 1988 – Publicatie WST Level 2
- 1995 – Level 2.5
- 1991 - STL bestandsformaat (EBU Tech3264-e) als 1^e EU non-proprietary format
- Teletekst technologie na ruim 40 jaar nog steeds gebruikt voor ondertitel transport



Status van de ondertitel markt

- **Verschuiving van lineaire broadcast naar (OTT) multi-device publicatie**
 - Vereist internet format ondertitel standaarden
 - 40 karakter / 23 ondertitel tekstregels teletekst raster → relatieve positioning en schaling, flexibele styling
 - SPMTE tijdcode → media tijdcode
- **Nieuwe video formaten...**
 - Bieden geen Teletekst ondersteuning
 - Pagina 888 voor doven en slechthorenden vervalt
 - Transport route vervalt hiermee ook (VBI/VANC data insertie in SDI)
- **Transitie naar IP infrastructuur**
 - Vereist IP gebaseerde (near real-time) transport protocollen
 - Out of band of als RTP elementary streams
- **Nieuwe use cases zorgen voor nieuw volume**
 - Video snacks zonder audio



Status van de ondertitel markt

- **Verhoogde vraag naar access services volume én nieuwe diensten**
 - Audio Descriptie (vraagt om een gestandaardiseerd authoring format)
 - Aanpassing door de gebruiker (vereist tekst gebaseerde ondertiteling tot in de ontvanger)
 - Player overlay diensten (timed metadata)
 - **Traditionele ondertitel fabrikanten verdwijnen**
 - Ondersteuning voor legacy en proprietary bestandsformaten wordt minder
 - **Nieuwe spelers in de markt**
 - IT based startups
 - Speech-to-text cloud en Text-to-speech (AI) diensten voor metadata verrijking (AWS, Azure, Google, IBM Watson Media, kleinere spelers zoals Spraaklab en Limecraft)
- **Essentieel om naar open internet standaarden te gaan**
= W3C → Timed Text Markup Language (TTML)
= EBU-TT als een broadcast specifieke subset



EBU-TT (TTML) als vervanging voor...

- **Legacy bestandsformaten**
 - Cavena, Pac, ... (proprietary binair)
 - EBU STL (open, binair, teletekst raster legacy)
- **Teletekst transport interfaces**
 - Newfor, Softel Cyclone / XAP
 - WST in VBI door de keten heen
 - OP-47
- **Closed captioning (SMPTE-TT)**
 - CEA-608 (Line 21 Captions), CEA-708
- **Andere internet ondertitelstandaarden**
 - WebVTT
- **DVB bitmapped subtitling**
 - Uitfasering op langere termijn



EBU-TT - toekomstgericht en flexibel

- **Eén gezamenlijke lichte standaard voor broadcast en online**
 - XML gebaseerd
 - Styling, positioning, timing voor traditioneel linear en OTT
 - Teletext/STL legacy support (conversiestrategie en tunneling), indien gewenst
- **Filebased en live workflow vanuit één documentspecificatie**
 - Minimale conversie tussen contributie en distributie
- **Human / machine readable**
 - Eenvoudige rule-based conversie en processing (delay compensatie, spellingscorrectie)
 - Eenvoudige indexatie t.b.v. contextual searches



EBU-TT - toekomstgericht en flexibel

- **Gestandaardiseerd**

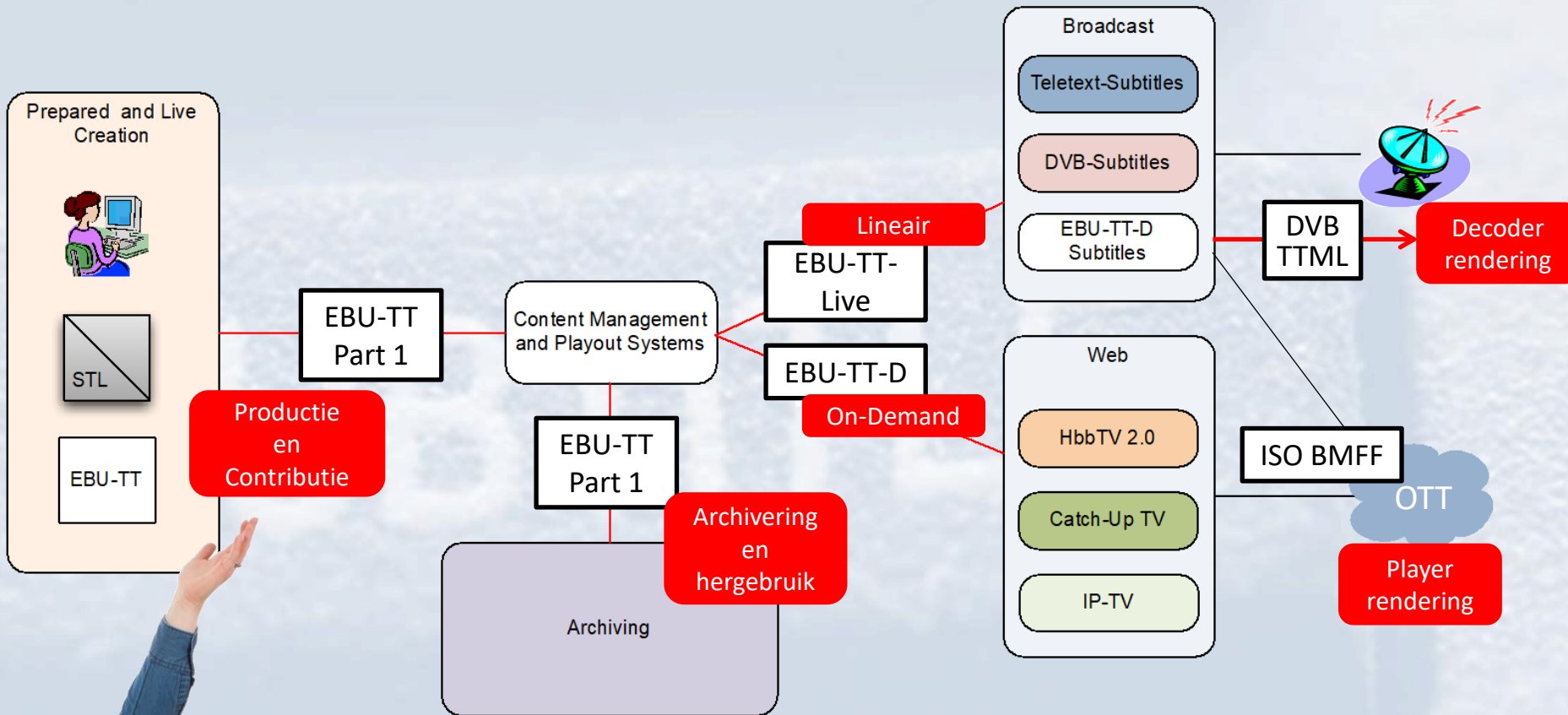
- W3C TTML subset, ETSI EN303560 (DVB TTML), ISO/BMFF (DASH), HbbTV
- IETF RFC 8759 RTP Payload for Timed Text Markup Language (TTML)
- W3C/EBU liaison gericht adoptie op EBU-TT-Live

- **Biedt nieuwe usecases**

- Timed metadata (i.E. geo-tagging, segmentatie, ...)
- Nieuwe access services (i.E. avatar sign language)
- Client side customization



EBU-TT in de publicatie keten



EBU-TT vs IMSC text profile

- **IMSC is het W3C TTML distributie formaat**
 - Parallel ontwikkeld aan EBU-TT-D
 - Beperkt in (custom) metadata. Breder met een image profile
 - In liaison opgetrokken in de IMSC1.1 tekst profile wijzigingen
 - Doorontwikkeling EBU-TT-D → W3C IMSCx.x
- **EBU-TT-D = IMSC1.1 tekst profile**
 - EBU-TT Part 1 gaat veel breder als productieformaat
 - EBU-TT-D en IMSC zijn gestripte exports met Media Tijdcode
- **EBU-TT-Live adoptie hangt in de lucht**
 - Discussiepunt: SMPTE tijdcode ondersteuning in IMSC
- **Hollywood is IMSC als productieformaat gaan gebruiken**
 - IMF bevat IMSC1

Productie en archivering



Open /
Vertalend



Doven &
Slechthorend



Timed
Metadata



Gesproken
Ondertiteling
(Player OS text-to-
speech)

Live transport



EBU
OPERATING EUROVISION AND EURORADIO

TECH 3370

**EBU-TT, PART 3
LIVE SUBTITLING APPLICATIONS**

SYSTEM MODEL AND CONTENT
PROFILE FOR AUTHORING
AND CONTRIBUTION

VERSION: 1.0
SOURCE: SP/MIM - XML SUBTITLES

Geneva
May 2017

System model

+

EBU
OPERATING EUROVISION AND EURORADIO

TECH 3370s1

**EBU-TT, PART 3 SUPPLEMENT 1
LIVE SUBTITLING APPLICATIONS:
CARRIAGE OVER WEBSOCKET**

VERSION: 1.0
SOURCE: SP/MIM - XML SUBTITLES

Geneva
May 2017

Websockets

**IETF
8759RFC**

RTP Payload for
Timed Text Markup Language
(TTML)

RTP

SMPTE STANDARD
Professional Media Over
Managed IP Networks:
SMPTE ST 291-1 Ancillary Data

Page 1 of 8

Table of Contents		
	Page	
1	Scope	3
2	Conformance Notation	3
3	Normative References	3
4	Terms and Definitions	4
5	ANC Data Packet RTP Payload Format	4
5.1	Overview (Informative)	4
5.2	RTP Payload Format	5
5.3	Media Clock and RTP Clock	6
5.4	RTP Timestamps	6
5.5	Keep Alive	6
6	Session Description Protocol (SDP)	7
Bibliography (Informative)		8

Approved April 12, 2018

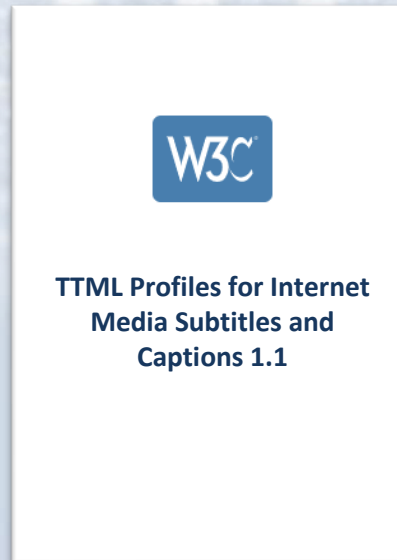
Work in progress



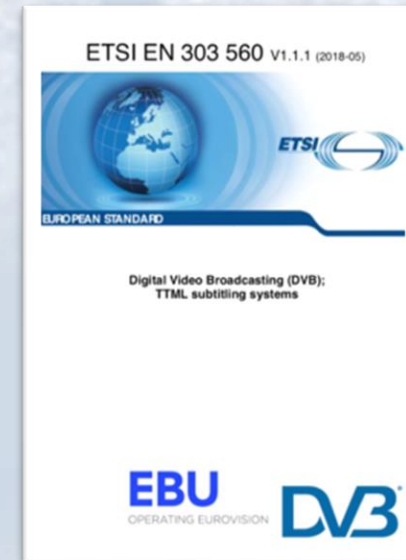
Distributie



Document specificatie



Transport specificatie

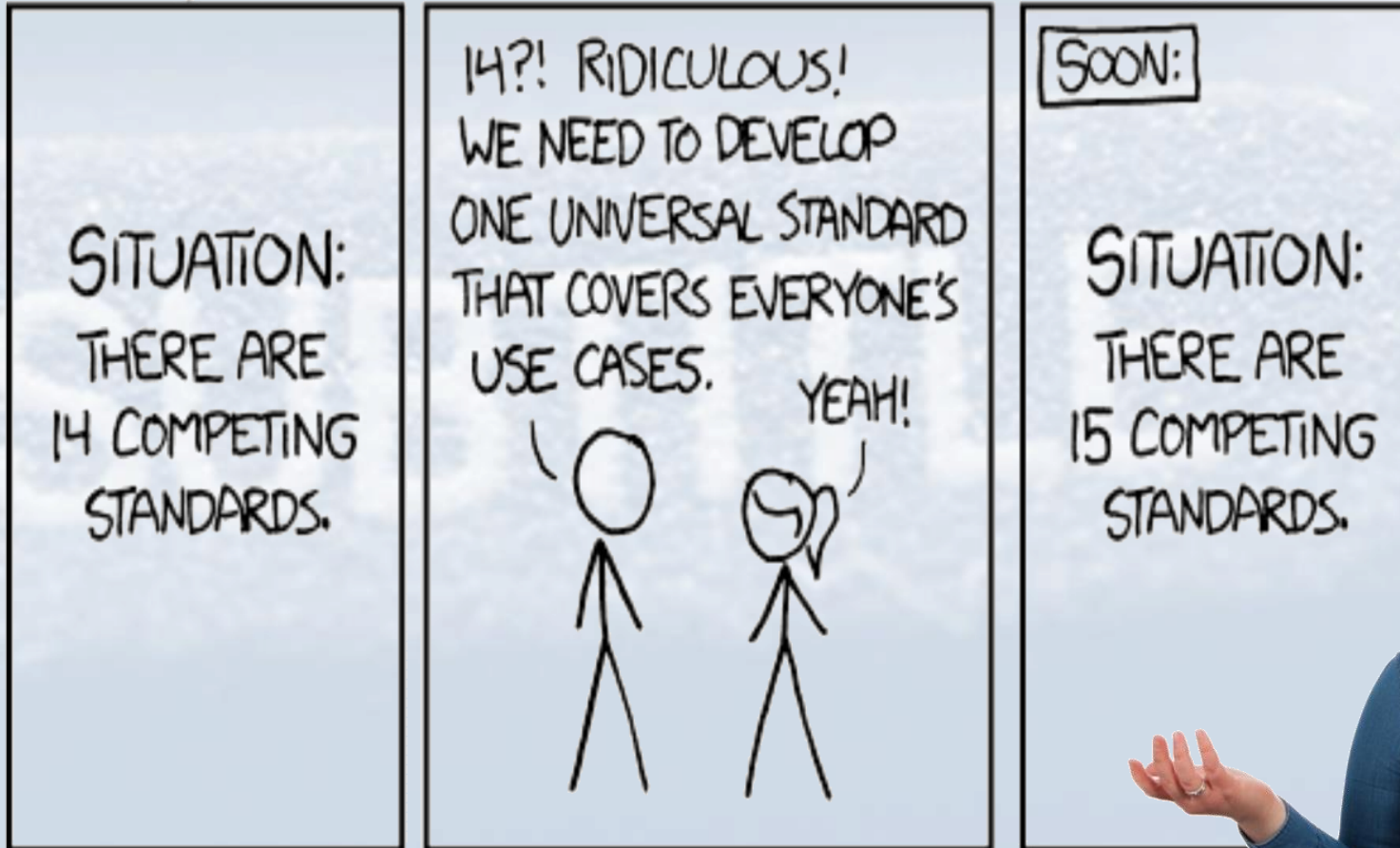


OTT - IPTV

Cable - Satellite - IPTV

Gezamenlijk NL document profiel?

HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)



NIET ONS DOEL!

Gezamenlijk NL document profiel – Waarom?

- Declaratie van styling, positionering etc. is nog veel te vrij
 - Vergemakkelijken van een ondubbelzinnige gefaseerde implementatie door meerdere leveranciers
 - Mitigeren van incompatibiliteitsrisico's en eindeloze aanpassingen
 - Geïsoleerde referentie-tests mogelijk maken
 - Bij voorkeur als ketenbreed Nederlands profiel
- Vingerwijzen is niet toegestaan! ... en niet nodig



Gezamenlijk NL document profiel – Hoe?

- Vernauwen, met behoud van volledige functionaliteit en compatibiliteit
- Styling en document format strategie
 - Logische structuur and declarative t.b.v. onze use cases
 - Leesbaarheid en eenvoud (geen nesting)
 - Right-to-left taalondersteuning (nieuwe Nederlanders)
- Positionering
 - Uniforme on-screen presentatie
 - Gelijke Teletekst raster vertaling
 - Nu gebaseerd op HD resolutie, maar eenvoudig UHDTV aanpasbaar
- Specifieke definitie van:
 - Vereiste en geadviseerde metadata
 - Logische positionering van custom metadata



Nog te doen

- **EBU-TT-Live / IMSC in ST2110-41?**
 - Op basis van IETF RFC 8759 RTP Payload for Timed Text Markup Language (TTML)
- **EBU-TT-Live adoptie door W3C**
- **Test and monitoring tooling**
 - Vendor implementatie is nog erg dun
- **Playout vendor motivatie**
 - Iedereen spreekt over HTML5... TTML rendering is dan toch gemakkelijk?
- **Gezamenlijk NL profiel?**
 - Eenvoud in de authoring markt
 - Efficiënte ontwikkeling en investeringen



Meer informatie

- **EBU specifications and resources**

Open source EBU-TT-Live websocket reference implementation (Python)

Implementation examples

<https://tech.ebu.ch/timedtext>

- **Subtitle profiles & sample files**

NPO <https://spec.npo.nl/ttml>

BBC <https://bbc.github.io/subtitle-guidelines/>

DPP http://dpp-assets.s3.amazonaws.com/wp-content/uploads/2014/03/Subtitle_Ext_Format_Final.pdf

- **Online IMSC/EBU-TT-D tools**

- Sandflow imscJS SDK <https://github.com/sandflow/imscJS>

- IMSC1 Renderer http://sandflow.com/imsc1_1/index.html

- **W3C ‘overview TTML and derivative Captions formats’**

<https://www.w3.org/AudioVideo/TT/docs/TTML-Profiles.html>



Dankjewel voor je aandacht

Vroegtijdig wijzen
naar oplossingen
is beter dan
vingerwijzen

