CONTENTUS
Next Generation Multimedia Libraries

Automatische Erschließung und semantische Verknüpfung von multimedialen Inhalten

Vorstellung der Ergebnisse des Forschungsprojekts CONTENTUS

Christoph Dosch
Christian Fey
Ronald Mies
Birgit Schmidt

THESEUS
Neue Technologien für das Internet der Dienste

Gefördert durch das
Bundesministerium für Wirtschaft und Technologie

© IRT – Dosch, Fey, Mies, Schmidt

Agenda

Problemstellung und Herausforderung ⇒ Christoph Dosch

- Rundfunk-Sicht, Produktion, Archive

Contentus-Prozesskette ⇒ Birgit Schmidt

- Einführung, Partner, Prozesskette
- Was wurde entwickelt (Übersicht)?

IRT-Entwicklungen ⇒ Christian Fey

- Einbindung von bereits existierenden Rundfunkmetadaten und Untertitel

Zusammenfassung und Demo ⇒ Ronald Mies

- Restaurierung, Semantische Multimediasuche

© IRT – Dosch, Fey, Mies, Schmidt

Agenda

Problemstellung und Herausforderung ⇒ Christoph Dosch

- Rundfunk-Sicht, Produktion, Archive

Contentus-Prozesskette ⇒ Birgit Schmidt

- Einführung, Partner, Prozesskette
- Was wurde entwickelt (Übersicht)?

IRT-Entwicklungen ⇒ Christian Fey

- Einbindung von bereits existierenden Rundfunkmetadaten und Untertitel

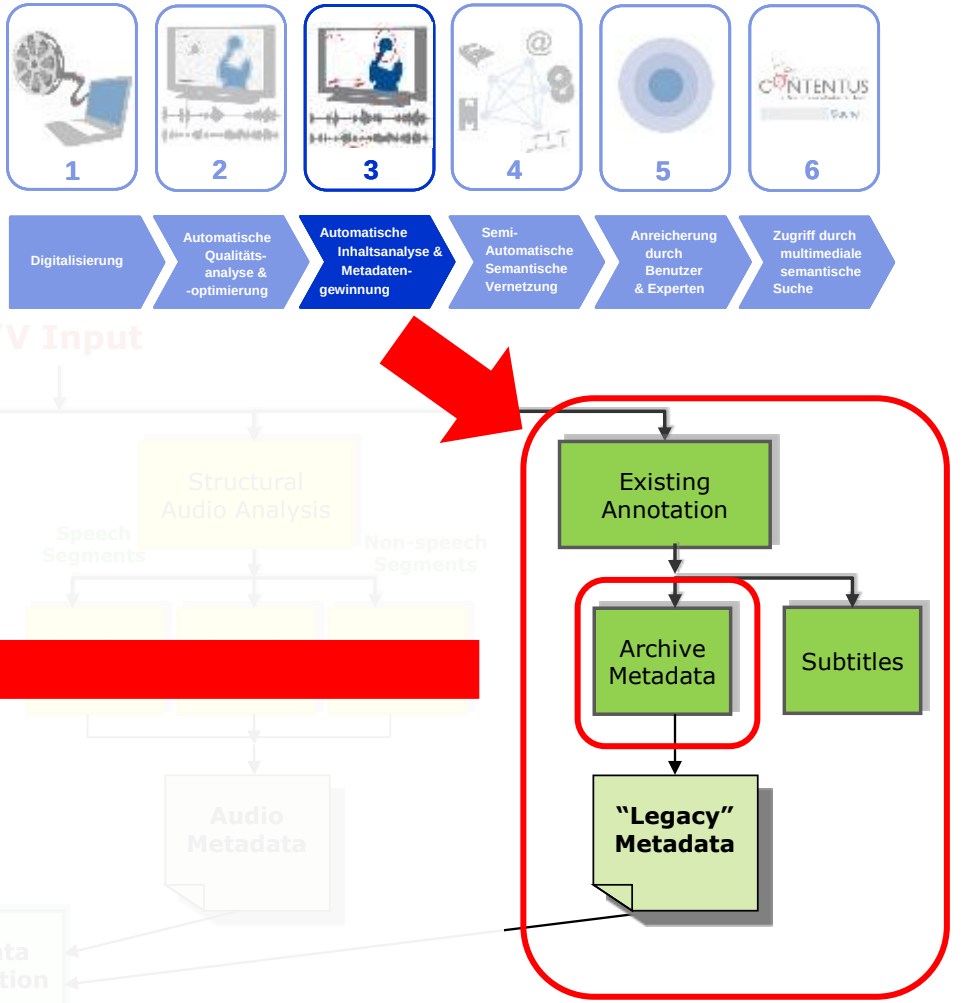
Zusammenfassung und Demo ⇒ Ronald Mies

- Restaurierung, Semantische Multimediasuche

© IRT – Dosch, Fey, Mies, Schmidt

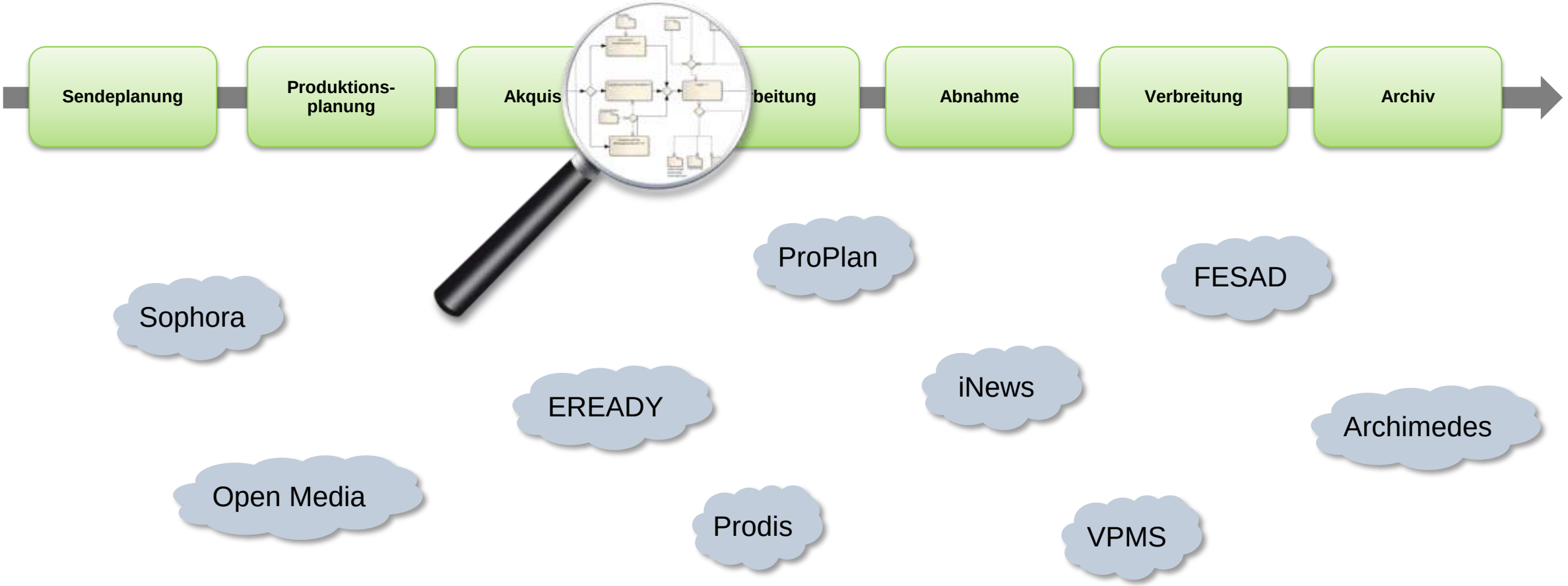
Integration existierender Rundfunk-metadaten in CONTENTUS

- Große Vielfalt der Metadaten-Formate alleine im Archiv
  - ARCHIMEDES, FESAD, Avid Interplay Archive, etc.
- Weitere Metadaten aus dem Produktionsprozess
  - Produktionsmetadaten (z.B. aus Kameras), Teleprompter-Texte, Untertitel etc.
- Zeitlicher Aufwand für die Erschließung sehr hoch
  - Zugriff auf Metadaten aus anderen Produktionsschritten nur eingeschränkt oder gar nicht möglich
  - Hohe Komplexität des gesamten Workflows



© IRT – Dosch, Fey, Mies, Schmidt

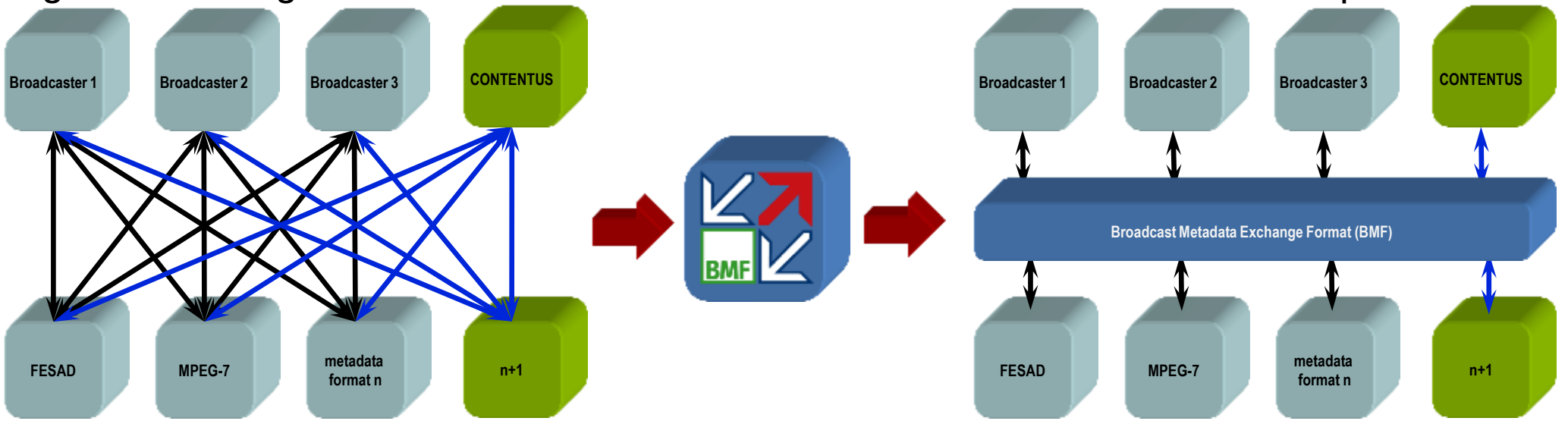
Produktionsprozesskette



© IRT – Dosch, Fey, Mies, Schmidt

BMF als Austauschformat

- **B**roadcast **M**etadata exchange **F**ormat (IRT-Entwicklung)
  - Zentrales Austauschformat zwischen unterschiedlichen (proprietären) Systemen
- Umfangreiches Metadaten-Vokabular für Fernsehen und Hörfunk
  - Abdeckung möglichst vieler Formate aus dem professionellen Rundfunkbereich
- Ermöglicht den IT-gestützten Austausch von Metadaten bzw. erhöht die Interoperabilität



© IRT – Dosch, Fey, Mies, Schmidt

Abbildung von AV-Metadaten in CONTENTUS

Problematik:

- Low-Level Metadaten müssen in der CONTENTUS Prozesskette verarbeitet werden
  - Bsp.: Koordinaten von Objekten wie gefundenen Gesichtern
- Zusätzlich müssen beschreibende Metadaten abgedeckt werden
  - Bsp.: Titel, Bildinhalt, Indexat
- AV-Metadaten müssen <u>eindeutig</u> abgebildet werden
  - Voraussetzung für eine Rücktransformation in andere Formate
  - Bsp.: Arbeitstitel <-> Titel => wichtige Informationen gehen verloren!
- Generisches Format notwendig

Lösung in CONTENTUS:

- Verwendung von MPEG-7 mit zusätzlichen Anpassungen

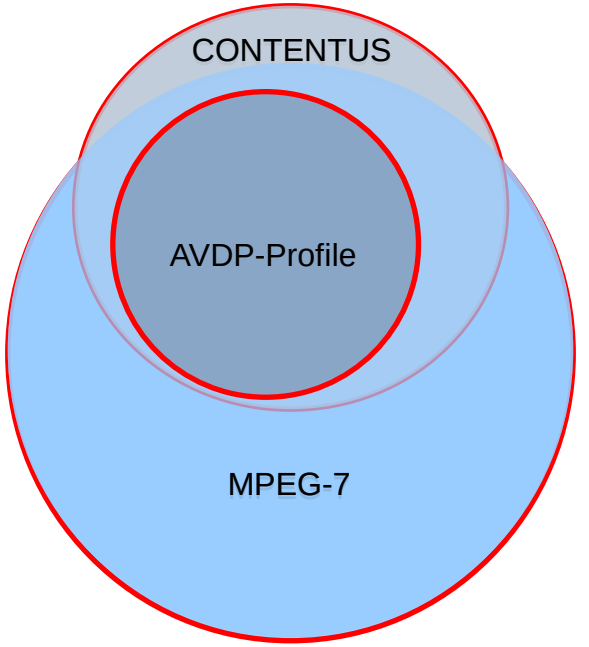
© IRT – Dosch, Fey, Mies, Schmidt

Abbildung von AV-Metadaten in CONTENTUS

- Basis ist das MPEG-7 Datenmodell (ISO/IEC 15938)

- Vorteil:
  - Standardisiertes Format
  - viele Technologien bereits vorhanden

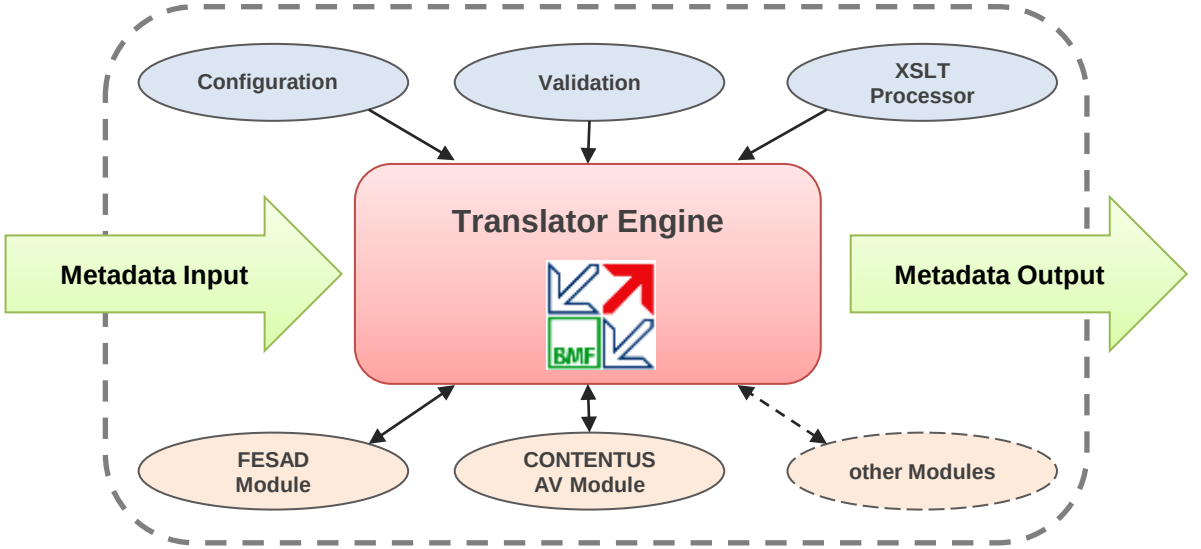
- Anpassung notwendig, da nicht rundfunkspezifisch
  - Einschränkungen in der Struktur
  - Spezifizierung neuer Elemente unter Verwendung der MPEG-7 DDL



- CONTENTUS Metadaten sind Teil des MPEG-7 "Audiovisual description profile" (AVDP)
  - seit Mai 2012 offizieller ISO-Standard ist (ISO/IEC 15938-9:2005/Amd1)

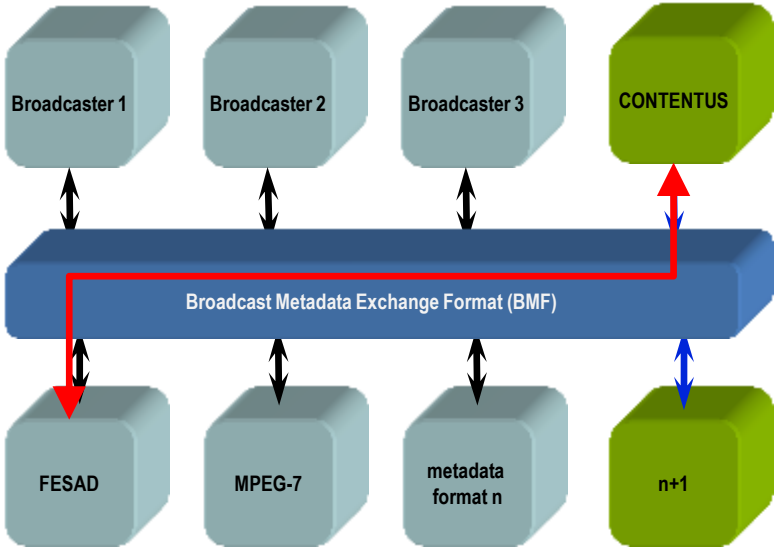
Anbindung der Rundfunkmetadaten in CONTENTUS (1)

- Metadatentranslator
  - Mapping der Metadaten mit Hilfe von XSLT-Transformationen
  - Zentrales Austauschformat BMF zur Abdeckung möglichst vieler Formate aus dem professionellen Rundfunkbereich
  - Vorteil: bei Updates keine Anpassung des Quellcodes notwendig
  - Serviceorientierte Architektur
    - Realisierung als Webservice



Anbindung der Rundfunkmetadaten in CONTENTUS (2)

- Exemplarische Anbindung einer FESAD-Datenbank
  - häufige Nutzung in den Rundfunkanstalten
  - Transformation FESAD <=> BMF <=> CONTENTUS MPEG-7
  - Proof of Concept: Rücktransformation auf Basis der Dokumentarischen Einheit



ib 2011

NAB (2012)

- BMF-Weiterentwicklung (http://bmf.irt.de/):
  - Dokumentation, XML-Schema und Beispiel-Instanzen stehen kostenfrei zur Verfügung

© IRT – Dosch, Fey, Mies, Schmidt

Untertitel als Metadaten nutzen, wieso?

Drei gute Gründe:

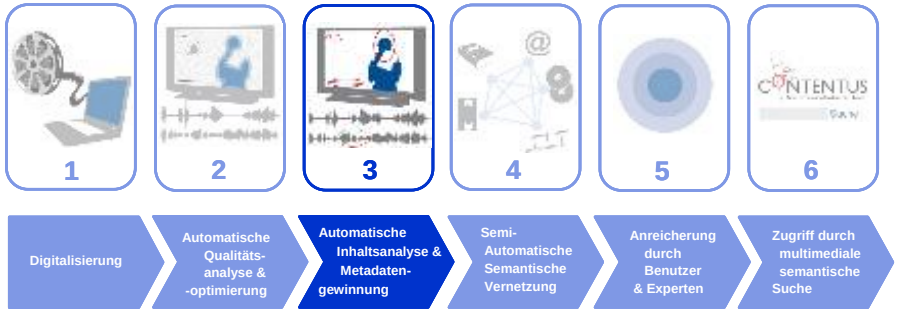
- Umstellung auf einen dateibasierten Workflow, sowie ein einheitliches, modernes UT-Format treiben die technische Erschließung von UT voran.
- Der Anteil der untertitelten FS-Programmen wird auch durch die politischen Entwicklungen ansteigen.
- Untertitel enthalten hochqualitative Informationen, da von Experten erstellt bzw. kontrolliert, und sind deswegen eine zuverlässige Basis für die Essenzbeschreibung.

Technische Entwicklungen im Untertitel-Bereich

- Bisher „einstreifiges“ Verfahren zur Übertragung von Untertiteln: UT und AV-Essenz werden bis zur Ausstrahlung gemeinsam transportiert
  - Analog: VBI (Austastlücke), Digital: SD-SDI
  - Format: EBU-STL (1991 von der EBU definiertes Binärformat)

- Zukünftig „zweistreifiges“ Verfahren: UT und AV-Essenz werden bis zur Ausstrahlung separat transportiert
  - Format: <u>EBU-TT (Timed Text)</u>
    - EBU-AG, Federführung IRT
    - Basiert auf der vom W3C spezifizierten TTML (Timed Text Markup Language)
  - <u>Dateibasiert: strukturierte XML-Daten</u>

Nutzung von Untertiteln als Metadaten in CONTENTUS



Voraussetzungen:

- Erweiterung CONTENTUS MPEG-7-Schema
- Umwandlung UT-Formate in CONTENTUS MPEG-7
- Import der Untertitel in den CONTENTUS-Workflow

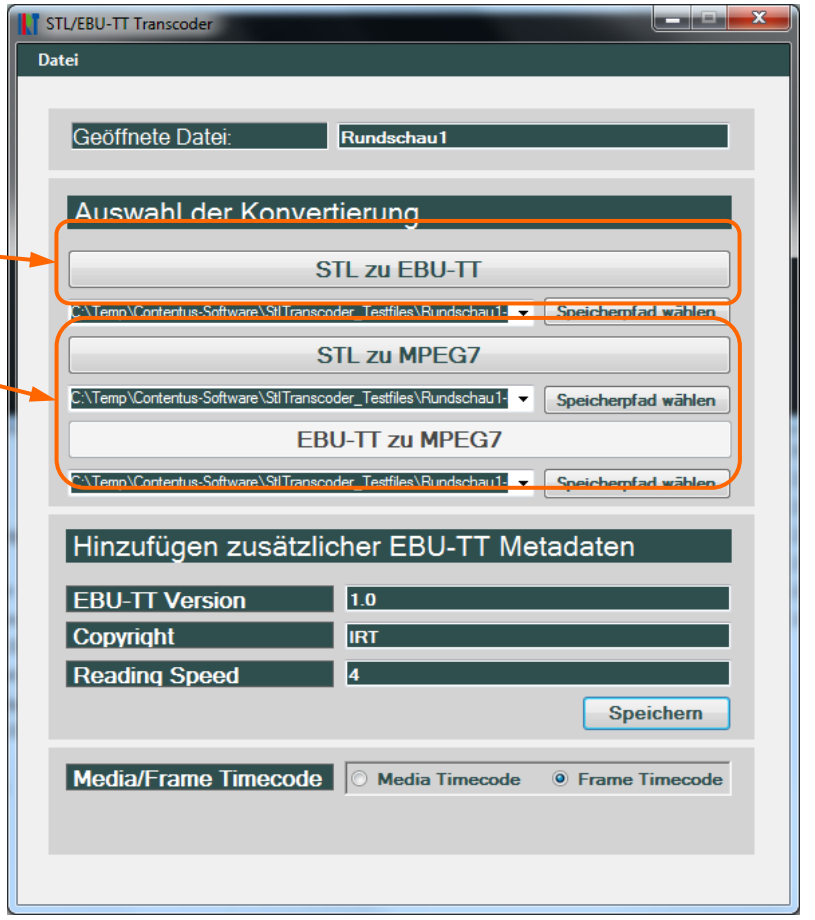
⇒ Anzeige der UT in den jeweiligen <u>CONTENTUS-Endnutzer-Anwendungen</u>

⇒ Nutzung der Untertitel zur Indexierung

© IRT – Dosch, Fey, Mies, Schmidt

Umwandlung von Untertiteln in CONTENTUS-Metadaten

- Prototypische Software-Entwicklung des IRT
- Konvertiert binäre STL-Dateien in EBU-TT
- Wandelt STL und EBU-TT in CONTENTUS-MPEG-7 um
- Erlaubt so die Einbindung von UT in CONTENTUS mit Hilfe des Metadaten-Translators



© IRT – Dosch, Fey, Mies, Schmidt

CONTENTUS
Next Generation Multimedia Libraries

Vielen Dank für Ihre Aufmerksamkeit!

Christoph Dosch (dosch@irt.de), Tel. +49-(0)89-32399-349
Christian Fey (fey@irt.de), Tel. +49-(0)89-32399-287
Ronald Mies (mies@irt.de), Tel. +49-(0)89-32399-503
Birgit Schmidt (schmidt@irt.de), Tel. +49-(0)89-32399-294

Institut für Rundfunktechnik
Floriansmühlstraße 60
80939 München

Die Folien/Dokumente sind durch das Urheberrecht geschützt.
Eine Vervielfältigung ist nur mit Genehmigung des Verfassers gestattet.
Dieser Urheberrechtshinweis darf nicht entfernt werden.