

Metadaten für Audio-/Radiosignale

Übersicht zu Standards und zur Gewinnung

Malik Al-hallak

Bauhaus-Universität Weimar
Fakultät Medien
malik.al-hallak@uni-weimar.de

Julius Bullinger

Bauhaus-Universität Weimar
Fakultät Medien
julius.bullinger@uni-weimar.de

Zusammenfassung

In this paper, we describe the formatting guidelines for the workshop proceedings. Body text of the abstract should be 10 pt. Times New Roman Italic.

Keywords

Keywords are your choice, and are not mandatory.

INTRODUCTION

The proceedings are the records of the workshop. We hope to give these workshop byproducts a single, high-quality appearance. To do this, we ask that authors follow some simple guidelines. In essence, we ask you to make your paper look like this document. The easiest way to do this is simply to use this template and replace the content with your own material.

PAGE SIZE & PAGE LIMIT

All material on each page should fit within a rectangle of 18x23.5 cm, centered on the page, beginning 1.9 cm (.75") from the top of the page and ending with 2.54 cm (1") from the bottom. The right and left margins should be 1.9 cm. The text should be in two 8.45 cm (3.33") columns with a .83 cm (.33") gutter. Camera-ready submissions should be about 8 pages in length or less.

TEXT & HEADINGS

For body text, please use a 10-point Times Roman font, or other Roman font with serifs, as close as possible in appearance to Times Roman in which these guidelines have been set. The goal is to have a 10-point text, as you see here. Please use sans-serif or nonproportional fonts only for special purposes, such as distinguishing source code text. If Times New Roman is not available, try the font named Computer Modern Roman. On a Macintosh, use the font named Times. Right margins should be justified, not ragged.

The following is an example of what a bulleted list should look like:

- All bullets should start at the same point.
- Spacing between the bullets can vary to produce good column and page breaks.

- The use of hanging indent is recommended.
- Numbered lists should follow similar layout.

Title and Authors

The title (Helvetica¹ 16-point bold), authors' names (Helvetica 12-point-Bold) and affiliations (Times New Roman 10-point) run across the full width of the page. We also recommend e-mail addresses for all authors. See the top of this page for three addresses. If only one address is needed, center all address text. For two addresses, use two centered columns, and so on. If more than three authors, you may have to improvise.²

Subsequent Pages

For pages other than the first page, start at the top of the page, and continue in double-column format. The two columns on the last page should be as close to equal length as possible.

Tabelle 1: Table captions should be placed above the table

Graphics	Top	In-between	Bottom
Figures	Good	Body text for the tables should be Times Roman 10 pt.	
Text	Graphs		

References and Citations

Footnotes should be Times New Roman 8-point, and ragged right. Ragged right was decided because of many web addresses being used.

A numbered list at the end of the article, ordered alphabetically by first author, and referenced by numbers in brackets [?]. See the examples of citations at the end of this document. Within this template file, use the style named references for the text of your citation. The body text of the citation/references section should be 10 pt. Times New Roman.

¹ If Helvetica is not available on your machine, Arial may be used as a substitute.

² If necessary, you may place some address information in a footnote, or in a named section at the end of your paper. Please make footnotes Times New Roman 8 pt.

References should be published materials accessible to the public. Internal technical reports may be cited only if they are easily accessible (i.e., give the address to obtain the report within your citation) and may be obtained by any reader. Private communications can be acknowledged, not referenced (e.g., "[Robertson, personal communication]"). Proprietary information may not be cited.

Page Numbering, Headers and Footers

Do not include headers, footers or page numbers in your submission. These will be added when the publications are assembled.

FIGURES/CAPTIONS

Place tables/figures/images in text as close to the reference as possible (see Figure 1). Figures may extend across both columns to a maximum width of 17.78 cm (7").

Captions should be Helvetica bold 9-point. They should be numbered (e.g., "Table 1" or "Figure 2"). Please note that the words "Table" and "Figure" are spelled out. Figure captions should be centered beneath the image or picture, and table captions should be centered above the table body.

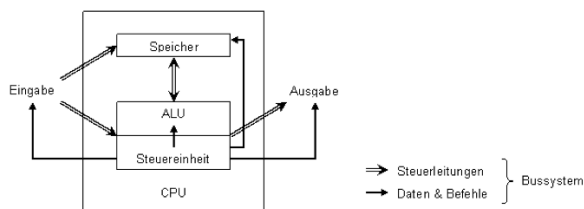


Abbildung 1: Place caption below the figure

EQUATIONS

Equations look very pretty in \LaTeX . Here's a nifty example of a matrix:

$$A_{mn} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad (1)$$

Try that with MS Word! Now, let's have a look at the prettiness of Equations 2 and 3:

$$f(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ -(n+1)/2 & \text{if } n \text{ is odd} \end{cases} \quad (2)$$

$$\int_{-\pi/2}^{\pi/2} \cos(\omega t) dt = \frac{2}{\omega} \sin\left(\frac{1}{2}\omega\pi\right) \quad (3)$$

SECTION HEADS

The heading of a section should be in Helvetica Bold 10-point bold, in all-capitals flush left with an additional 6 points of white space above the section head. Sections and subsequent subsections should flush left.

Subsections

The heading of subsections should be in Helvetica 10-point bold with only the initial letters capitalized. For subsections and subsubsections, a word like *the* or *a* is not capitalized unless it is the first word of the header.

Subsubsections

The heading for subsubsections should be in Helvetica 10-point italic with initial letters capitalized and 6 points of white space above the subsubsection head.

ACKNOWLEDGEMENTS

Your appreciation to employers, co-workers, department heads, and/or institutions that issued you a grant can be acknowledged in this section.

DER METADATEN STANDARD MPEG-7

Einleitung

Der MPEG-7 Standard wurde 2002 von der *Moving Picture Expert Group* als ISO/IEC 15938 verabschiedet. Im Gegensatz zu den früheren MPEG Standards (MPEG-1, MPEG-2 und MPEG-4) kein Kompressionsstandard für multimediale Daten sondern ein Metadaten Standard für selbige.

Er bietet eine auf XML-Basierende umfangreiche Repräsentation von Metainformation für Audio- und Videodaten.

Im folgenden wird der Standard im Hinblick auf Audiosignale detailliert vorgestellt und seine automatisch generierbaren Aspekte vorgestellt.

Inhalt

Der MPEG-7 Standard spezifiziert vier Bestandteile:

- *Descriptors*
- *Descriptor Schemes*
- *Description Definition Language*
- *Binary Format für MPEG-7*

Descriptors (D)

Der erste Teil des Standards sind *Descriptors*. Diese arbeiten auf räumlichen, zeitlichen oder physischen Segmenten der Ressource und halten Daten zur Repräsentation eines Merkmals der Audiodatei.

Dabei werden entweder gleiche Zeitintervalle der Ressource beschrieben oder sie wird in Segmente unterteilt und Ähnlichkeiten und Differenzen werden aufgezeichnet. Beispiele für solche Descriptors sind der *AudioWaveform Descriptor* welcher einem Intervall des Signals z.B. maximaler und minimaler Tonfrequenz zuweist.

Diese Strukturen sind Low-Level und beinhalten lediglich die technischen Eigenschaften der Audio-Ressource. Es werden derzeit 17 *Feature Descriptors* und ein spezieller *Silence Descriptor* standardisiert. Diese 18 Strukturen sind in Untergruppen Organisiert (siehe Abb. 2).

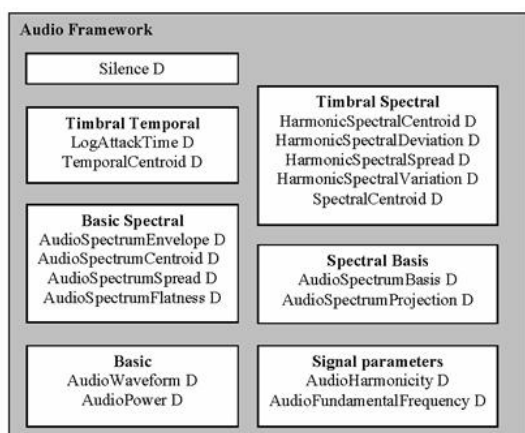


Abbildung 2: Low-Level Descriptors in MPEG-7

Der *Silence Descriptor* ist einfach die Darstellung dass an dieser Stelle kein signifikanter Ton vorhanden ist und kann z.B. dazu eingesetzt werden ein Segment mit dieser Eigenschaft nicht weiter zu bearbeiten was die Erstellung der Metadaten effizienter macht.

Description Schemes (DS)

Der MPEG-7 Standard liefert neben den Low-Level Descriptors auch High-Level Description Tools in Form von *Description Schemes*. Description Schemes beschreiben die Audiodatei auf einer höheren Abstraktionsebene und setzen dafür Descriptors und andere Description Schemes in Beziehung zueinander.

Description Schemes nutzen die Rohdaten aus Descriptors und anderen Description Schemes um Audiodaten auf einem hohen Abstraktionsgrad zu beschreiben. Sodass Description Schemes die Struktur und die Semantik liefern in welcher die Descriptors und andere Description Schemes in Verbindung gesetzt wird. Damit kann man neben der rein technischen Low-Level Beschreibung der Ressource auch Charakteristiken wie Ton-, Instrumenten- und Melodie-Erkennung erzeugen und vermerken.

Description Definition Language (DDL)

Neben den bereits genannten Möglichkeiten bietet der MPEG-7 Standard ein weiteres mächtiges Werkzeug. Mit der Description Definition Language ist es möglich weitere, auf die Anforderungen des jeweiligen Nutzers zugeschnittene, Descriptors und Description Schemes zu entwickeln.

Die Description Definition Language ist eine Erweiterung von XML-Schema.

Binary Format für MPEG-7 BiM

Das BiM ist entwickelt worden um die Metadaten schnell übertragen und streamen zu können. Das BiM bietet eine hohe Kompressionsrate der XML-Metadaten (bis zu 98%) und einen schnellen Zugriff auf Daten Entitäten im komprimierten Stream

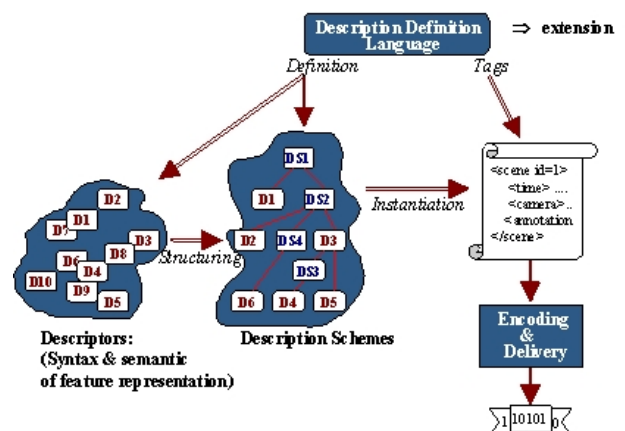


Abbildung 3: MPEG-7 Aufbau

DER METADATEN STANDARD EBUCORE

Einleitung

Der EBUCore ist ein Metadaten set welches von der *European Broadcasting Union (EBU)* entwickelt wurde. Die EBU ist ein Zusammenschluss von 74 Rundfunkanstalten in 56 Ländern Europas, Nordafrikas und Vorderasiens. Dieser Zusammenschluss produziert z.B. auch den Eurovision Song Contest.

Der EBUCore ist eine Erweiterung des *DublinCore*, welcher 1995 in Dublin/Ohio festgesetzt wurde. Der DublinCore spezifiziert eine Reihe von Metadaten zur Beschreibung von Dokumenten im Internet.

Im Jahr 2000 war es das anfängliche Ziel der EBU den DublinCore auf Audioarchive zu übertragen. Dieses Ziel wurde 2008 erweitert. Das neue Ziel war es den EBUCore so zu erweitern, das man mit den Metadaten nicht nur Daten im Archiv beschreiben kann, sondern mit ihrer Hilfe auch Archive durch z.B. Suchmaschinen im Browser durchsuchen kann.

Die EBU sagt, dass heutige Ziele nicht mehr auf Audioressourcen und Archive beschränkt sind.

Der bekanntesten Einsatzort des EBUCore ist derzeit die *Europeana*. Die Europeana ist ein Projekt welches darauf abzielt, das kulturelle Erbe Europas zu digitalisieren und für jedermann zugänglich zu machen. Mithilfe des Webportals hat man Zugriff auf viele Millionen multimediale Ressourcen aus Europa und der Welt.

Damit man diese Ressourcen auch nach der gewünschten durchsuchen kann, braucht es ein einheitliches Metadaten Format für Europa.

Der Standard

Der EBUCore ist eine minimale Liste von Attributen die man benötigt um eine multimediale Ressource zu beschreiben. Getreu dem Motto :

Wenn du's nicht findest, hast du's auch nicht.

versucht die EBU durch den EBUCore einen allumfassenden Standard bereit zu stellen welcher eine Verbindung zwischen kulturellen Datenbanken, Produzenten multimedialer Ressourcen, Rundfunkanstalten, Archiven und Netzentologien herstellt.

Auch der EBUCore nutzt das XML Format zur Repräsentation der Metainformation. Der Standard ist unter anderem in der EBU tech3293 spezifiziert.

Anders als der MPEG-7 Standard muss der EBUCore komplett manuell erstellt werden. Dazu wird einfach ein XML-Instanz Dokument erstellt welches dem XML-Schema entsprechen muss.

Das XML-Schema definiert über 30 Tags welche zur Beschreibung der Ressource dienen. Dabei werden neben den üblichen Beschreibungsformen, wie Titel oder Sprache, auch Tags angeboten die Informationen über das physikalische oder digitale Format, Version oder Herausgabe Historie. Diese Tags spiegeln hier beispielhaft das bestreben produzierenden und konsumierende Gruppen unter einen Standard zusammen zu fassen.

Es folgen nun zwei Beispiele wie ein XML-Instanz Dokument für ein Bild und eine Audioressource aussehen kann.

Dabei kann man gut erkennen das der Standard komplett unabhängig von der zu beschreibenden Ressource fungiert.

Listing 1: Beispiel-XML EBUCore für Albrecht Dürers Bild: Adam und Eva

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <schema xmlns:xsi=
3  "http://www.w3.org/2001/XMLSchema-instance" ...>
4
5
6  <ebuCoreMain>
7    <coreMetadataType>
8      <title>Adam and Eve</title>
9      <creator>Duerer, Albrecht (german painter) </creator>
10     <description>...</description>
11     <source>VADS</source>
12     <rights>The Bowes Museum Castle, Co. Durham</rights>
13     <provider>CultureGrid; UK</provider>
14     <subject>religion (Adam and Eve)</subject>
15     <type>Physical Object</type>
16   </coreMetadataType>
17 </ebuCoreMain>
18 </schema>
```

Listing 2: Beispiel-XML EBUCore für die Europahymne

```
19 <?xml version="1.0" encoding="UTF-8"?>
20 <schema xmlns:xsi=
21 "http://www.w3.org/2001/XMLSchema-instance" ...>
22
23
24 <ebuCoreMain>
25   <coreMetadataType>
26     <title>Europahymne</title>
27     <creator>van Beethoven, Ludwig</creator>
28     <description>...</description>
29     <source>
30       Dernier mouvement de la Neuvieme symphonie
31       (Ode a la joie) ...
32     </source>
33     <rights>
34       Centre Virtuel de la Connaissance sur l Europe (CVCE)
35     </rights>
36     <provider>ASSETS; Luxembourg</provider>
37     <language>fr; en; de</language>
38     <identifier>
39       doi_cvce_audiovisual_4b6e5671-a0fc-45fa-
40       8368-f84e4817dc14
41     </identifier>
42     <type>Physical Object</type>
43   </coreMetadataType>
44 </ebuCoreMain>
45 </schema>
```

Die beiden Beispiele sind sehr minimalistisch aber sie dienen dem Zweck einen großen Vorteil des EBUCore zu zeigen, die einfache Bedienung bei gleichzeitig relativ großer Beschreibungsmöglichkeit. Durch die weiteren Tags und zugehörige Attribute kann man von der Erstellung der Ressource (zum Beispiel mit dem Version-Tag) über die Auslieferung (zum Beispiel Title oder Creator Tag) bis zur Archivierung (zum Beispiel das Subject Tag).

EBUCore und Semantic Web

EBUCore liefert auch Bestandteile die für das Semantic Web genutzt werden können. Mit dem *<realtion>* – Tag ist es möglich einzelne Ressourcen mit anderen in Verbindung zu setzen. Man kann aus verschiedenen Typen die gewünschte Relation ausdrücken. Beispiele hierfür sind : *isPartOf*, *isRequiredBy* oder *isFormatOf*.

Ein weiterer Mechanismus der für das Semantic Web genutzt werden könnte, ist der *<identifier>* – Tag. Mit diesem kann eine eindeutige Referenz auf die Ressource vermerkt werden, wie zum Beispiel ein DOI.

VERGLEICH

Nutzer

Die verschiedenen Standards haben auch verschiedene Nutzer. Der **MPEG-7** Standard wird vorrangig von Algorithmen zur Suche verwendet um zum Beispiel "Query By Humming" Anfragen zu bearbeiten. Das liegt auch daran dass fertige MPEG-7 XML Dateien durch die vielen Descriptors welche beispielsweise im 10 ms Intervall eine Audioressource abtasten.

Der **EBUCore** kann durch seine einfache Bedienung und seine dennoch große Vielfalt von jedem eingesetzt werden. Derzeit ist der größte Nutzer die Europeana und die angeschlossenen Projekte wie: ECLAP, Digital Contemporary Art, EUScreen und andere.

NewsML-G2 ist ein Standard der nur von Nachrichtenagenturen eingesetzt wird. Er ist weltweit im Einsatz und dient zur Kommunikation zwischen den Agenturen.

TV-Anytime wird vom Endverbraucher genutzt. Er kann über TV-Anytime Informationen zur Ressource erhalten und auch selber Informationen wie zum Beispiel Meinungen eintragen.

Abstraktionsgrad

Der Abstraktionsgrad der Standards unterscheidet sich stark. **MPEG-7** hat einen sehr hohen Abstraktionsgrad. Er bietet die Möglichkeit die Audioressource über die physikalischen Eigenschaften hinaus auf ein hohes Niveau bis

zur Notenerkennung und Automatischer Spracherkennung (Automatic Speech Recognition, ASR). **TV-Anytime** ist ebenfalls auf einem hohen Level der Abstraktion. Es bietet neben der Beschreibung von Inhalten, Regie und Mitwirkende auch die Möglichkeit Bewertungen von Benutzern über den Inhalt einzusehen und so auch eine Kommunikation statt eine reine Distribution zu gewährleisten.

Dagegen sind die beiden anderen Spezifikationen eher auf einer geringen Stufe der Abstraktion. **EBUCore** und **NewsML-G2** bieten lediglich eine Beschreibung der Ressource auf allgemeiner Ebene wie Inhalt und beteiligte Personen. Dies ist jedoch keinerlei Nachteil, da sie genau dafür entwickelt wurden.

Generierung

Man kann die Standards hinsichtlich der Generierung in zwei Gruppen teilen. Die eine Gruppe bildet der **MPEG-7** Standard welcher mithilfe von Applikationen und Werkzeugen nahezu automatisch generiert werden kann.

Die andere Gruppe sind **EBUCore**, **TV-Anytime** und **NewsML-G2**. Diese lassen sich nicht automatisch generieren sondern werden manuell erzeugt.

Stelle in der Verwertungskette

Zweck

VERHÄLTNIS ZWISCHEN XML UND RDF/OWL (SEMANTIC WEB) UND SKOS FÜR CLASSIFICATION SCHEMES

Absatz

RDF, RDFS und RDF/XML

RDF

Das Resource Description Framework (RDF) ist eine universelle formale Sprache zur Beschreibung von strukturierten Informationen (ursprünglich: Metadaten im Web). Mittels RDF werden Informationen über so genannte Ressourcen gespeichert. Ressourcen sind Uniform Resource Identifiers (URIs), die eine Quelle referenzieren. Diese Ressourcen können Webseiten, ISBNs für Bücher und Zeitschriften oder DOIs (zum Beispiel doi:10.1000/182) sein.

Metadaten werden als Tripel – Subjekt, Prädikat, Objekt – gespeichert. Dabei wird die Beziehung zwischen Subjekt und Objekt angegeben. Die Beziehung »Die Glocke« ist von Friedrich Schiller geschrieben« ergibt das Tripel »Die Glocke« (Subjekt), »ist geschrieben von« (Prädikat), »Friedrich Schiller« (Objekt). Zur Notation nutzt man häufig die (inoffizielle) *Turtle*-Syntax. Auf diese soll hier jedoch nicht eingegangen werden, da in der Praxis überwiegend die Serialisierung mittels Extensible Markup Language (XML) genutzt wird.

RDF/XML

Dieses vom W3C standardisierte Format wird RDF/XML genannt. Listing 3 zeigt ein Beispiel für obige Beziehung in RDF/XML. Dabei wurden zusätzlich einige Elemente der Dublin-Core-Spezifikation genutzt.

RDFS

RDF benötigen zur *Interpretation* der dargestellten Beziehungen ein gemeinsames Vokabular, zum Beispiel Dublin Core. RDF Schema (RDFS) stellen dafür ein Vokabular zur Verfügung, mit dessen Hilfe innerhalb von RDFS-

Dokumenten Aussagen über semantische Beziehungen der Termini gemacht werden können. Es handelt sich dabei nicht um ein thematisches Vokabular; vielmehr

OWL

Web Ontology Language (OWL) ist eine Spezifikation des W3C, die entwickelt wurde, um...

OWL basiert auf RDF/XML, fügt jedoch weitere Ausdrücke zur Beschreibung von Eigenschaften und Klassen hinzu. Damit werden Aussagen ähnlich der Prädikatenlogik ermöglicht (Disjunktheit, Aussage "genau einer", Äquivalenz...)

OWL bietet drei Untersprachen mit steigender Ausdruckstärke (*OWL Lite*, *DL* (description logic) und *Full*) und unterschiedlichen Zielgruppen. OWL Lite ist eine (echte) Teilsprache von OWL DL, welches wiederum eine (echte) Teilsprache von OWL Full ist.

OWL 2

SKOS

Simple Knowledge Organization System (SKOS) ist eine vom W3C spezifizierte formale Sprache zum Verteilen und Verlinken von Informationssystemen wie Thesauri, Taxonomien und Normdateien im Web. SKOS basiert auf RDF und RDFS.

...

SKOS soll die Nutzung von concept schemes im Web im Vergleich zu OWL erleichtern. OWL verfolgt das Ziel, komplexe Konzeptstrukturen auszudrücken, um so ausführliche Metadaten zu generieren und Schlussfolgerungswerkzeuge zu unterstützen. Die Erstellung sinnvoller Web-Ontologien ist jedoch sehr aufwändig, und in vielen Fällen werden die zahlreichen Möglichkeiten der OWL nicht benötigt. Deshalb kann hier der Einsatz von SKOS vorteilhaft sein.

Listing 3: Beispiel für XML/RDF

```
1 <?xml version="1.0" encoding="UTF-8" ?>
2 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3   xmlns:dc="http://purl.org/dc/elements/1.1/"
4   xmlns:ex="http://example.org/">
5
6   <rdf:Description rdf:about="http://example.org/DieGlocke">
7     <dc:title>Das Lied von der Glocke</dc:title>
8     <ex:geschriebenVon>
9       <rdf:Description rdf:about="http://example.org/FriedrichSchiller">
10         <dc:author>Friedrich Schiller</dc:author>
11       </rdf:Description>
12     </ex:geschriebenVon>
13   </rdf:Description>
14 </rdf:RDF>
```