

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.

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

²If 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.

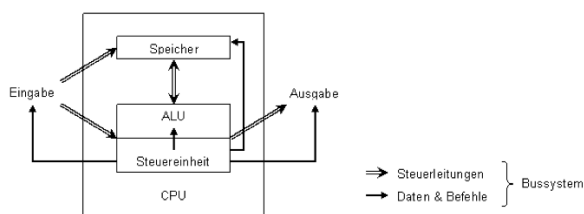


Abbildung 1: Place caption below the figure

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:

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

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

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.

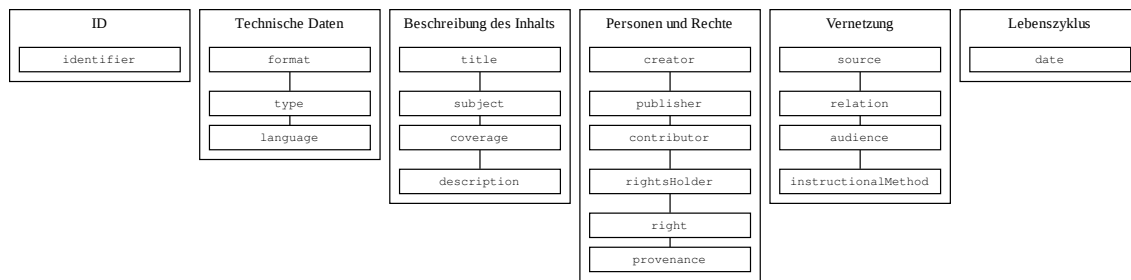


Abbildung 2: Kernelemente des Dublin-Core-Metadaten-Sets

EBUCORE METADATA SET [B]

Im Februar 1950 schlossen sich öffentliche Rundfunkanstalten aus 23 europäischen Ländern zu einer Union zusammen, um »eine Interessensgruppe für den Rundfunk in technischen und rechtlichen Belangen« zu gründen. Heute gehören der Europäischen Rundfunkunion (*European Broadcasting Union, EBU*) 74 Mitglieder aus 56 Ländern Europas sowie 36 assoziierte Mitglieder der ganzen Welt an. (Stand: Mai 2011).

Im Jahr 2000 kam der Wunsch auf, die Spezifikation des *Dublin-Core*-Metadaten-Sets auf Audio-Archive zu übertragen.

Dublin Core ist ein 1994 entworfener Standard für Metadaten mit dem Ziel, Dokumente im Internet leichter auffindbar zu machen. Dafür werden 15 Kernfelder zur Verfügung gestellt, welche technische Daten, eine Beschreibung des Inhalts, Personen und Rechte sowie die Vernetzung und den Lebenszyklus eines Dokuments beschreiben.

Im Jahr 2008 wurde dieser Wunsch erweitert auf die Spezifikation eines gemeinsamen Kerns an Informationen zur Durchsuchbarkeit von Archiven aller Art. Heute ist EBU-Core spezifiziert als »eine minimale Liste an Attributen zur Beschreibung von Audio- und Videoressourcen für eine große Auswahl an Rundfunkanwendungen, inklusive Archivierung, Austausch und Veröffentlichung«. Dafür werden die 15 Kernelemente von Dublin Core erweitert:

Outline:

EBU

Aufbau auf Dublin-Core

Dublin-Core

Scope

Aufbau

Grafik der Verwertungskette (Archive)
und noch irgendeine Grafik?!

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

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 Identifier (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 von W3C standardisierte Format wird RDF/XML genannt. Listing 1 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 Ausdrucksstärke (*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.

```
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>
```

Listing 1: Beispiel für XML/RDF