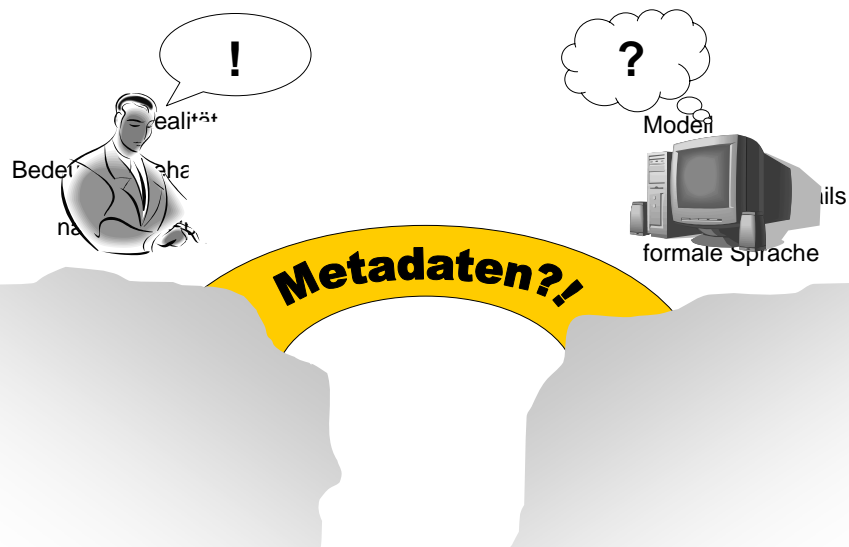


Metadaten für Multimedia im Überblick

Präsentation von
Florian Hoffmann
betreut von
Prof. Dr. H. Stuckenschmidt

Seminar Bridging the Semantic Gap SS 2006
unter der Leitung von
Prod. Dr. W. Effelsberg

Bridging the Semantic Gap



1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Inhalt

1. Bridging the semantic gap
2. Metadaten allgemein
 - Definition, Systematik, Kritik, Ziele, ...
3. Metadaten für Multimedia
 - aus der Multimediatechnik
 - ID3
 - JPEG 2000
 - MPEG-7
 - MPEG-21
 - TV-Anytime
 - aus dem Bereich Semantisches Netz
 - DCMI
 - LOM
 - RDF
 - RDF Schema
 - OWL
 - SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

2. Metadaten allgemein

Definition

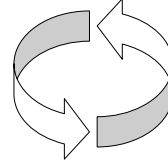
Metadaten sind Daten über Daten.

Rohdaten oder Metadaten?

The Lady of Shalott
Alfred Lord Tennyson
On either side of the river lie
Long fields of barley and of rye
That clothe the wold and meet the sky,
And through the field the road runs by
To many-tower'd Camelot.
...

→ Rollenwechsel

Metadaten über Metadaten
über Metadaten über ...



→ Rekursivität

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Unterscheidung

- ... nach dem **Inhalt**
 - Ressourcen beschreibend
 - Inhalt beschreibend
- ... nach der **Veränderlichkeit**
 - unveränderlich
 - veränderlich
- ... nach der **logischen Funktion**
 - subsymbolisch (Rohdaten)
 - symbolisch (zur Beschreibung)
 - logisch (zum Schlussfolgern)

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Lebenszyklus

1. Erzeugung

- Wann, wo und von wem?

2. Veränderung

- bei Veränderung der Daten
- bei Weiterverarbeitung der Daten (Metaproduktion)

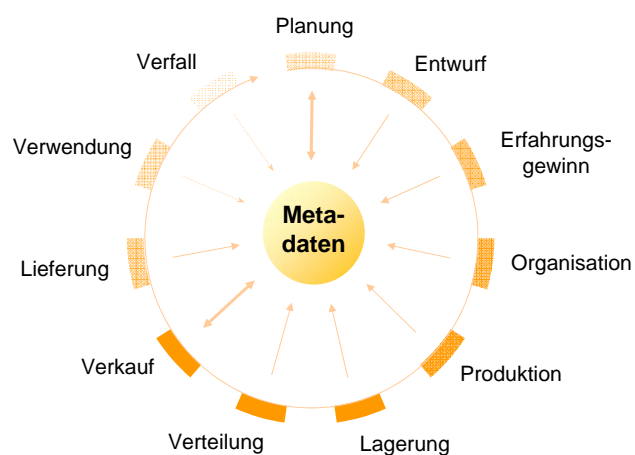
3. Zerstörung

- Sollten Metadaten ihre Daten überleben?

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Erzeugung



→ Metadaten sollten gleich bei Anfall erfasst werden

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Weitere Fragen

- **Welche Metadaten werden benötigt?**

- Anwendung sollte vorher bekannt sein
Negativbeispiel: EXIF, tote Metadaten für Digitalkameras

- **Wo bewahrt man sie auf?**

- intern bei den Daten?
- extern in eigener Datei?

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Ontologie

→ Vokabular eines Wissensbereichs und Zusammenhang der Fachwörter

- **Komplexität**

- einfach (disjunkte Kategorien)
- mit Vererbung (Untermengenbildung)
- komplex (komplette Mengenlehre)

- **Erstellung**

- vorgefertigt?
- von den Benutzern angelegt?

- **Darstellung**

- implizit = nur für Menschen verständlich?
- explizit = auch für Computer verständlich?

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Kritik

- Zu teuer / zeitaufwändig
- Zu kompliziert
- Zu subjektiv / kontextabhängig
- Überflüssig

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Ziele

- Genaueres Suchen
- Wissensrepräsentation
- Komprimierung
- Variable Darstellung
- Workflow-Automatisierung
- ...

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

3. Metadaten für Multimedia

Einfluss aus zwei Richtungen

Multimedia

ID3
JPEG 2000
MPEG-7
MPEG-21
TV-Anytime

WWW

DCMI
LOM
RDF
RDF Schema
OWL
SMIL

Identifier on MP3

Multimedia

ID3

JPEG 2000

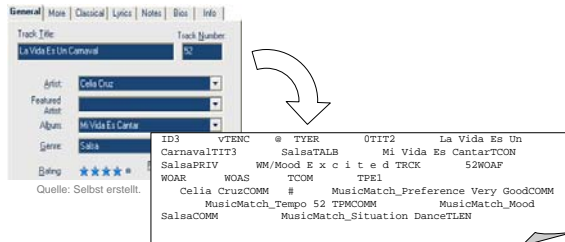
MPEG-7

MPEG-21

TV-Anytime

Ziel: Genaueres Suchen

Anwendung: Playlists, P2P, Online-Shops



- Text in vordefinierten Feldern
- einfaches Kategoriensystem
- (Klar)Text, in Ressource

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

JPEG 2000

Multimedia

ID3

JPEG 2000

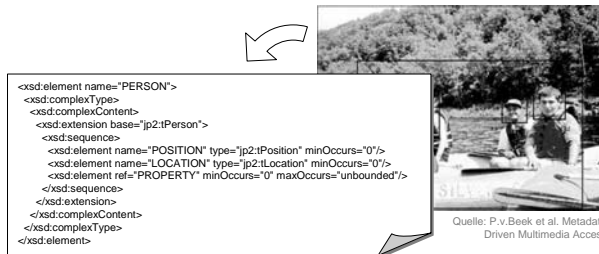
MPEG-7

MPEG-21

TV-Anytime

Ziel: Variable Darstellung

Anwendung: Bilder im Handyformat / als Text



Quelle: M. Bollek (Editor) et al. JPEG 2000 Part II Final Committee Draft.

- Regions of Interest
- ähnlich wie MPEG-7
- XML, in Ressource

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

MPEG-7

Multimedia

ID3

JPEG 2000

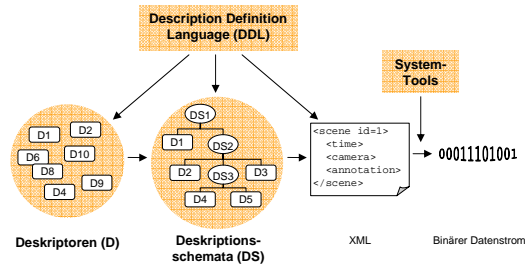
MPEG-7

MPEG-21

TV-Anytime

Ziel: Alle?!

Anwendung: Alle?!



- DDL / Deskriptoren / Schemata
- Tools für Binärisierung / Streaming / ...
- XML, extern

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

MPEG-21

Multimedia

ID3

JPEG 2000

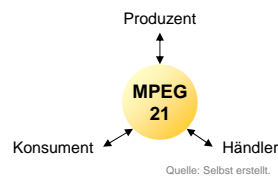
MPEG-7

MPEG-21

TV-Anytime

Ziel: Infrastruktur für Multimediadaten

Anwendung: Alle



- Teil 3: Identifikation digitaler Inhalte
- Teil 6: Rights Data Dictionary
- Teil 7: Metadaten transport zw. versch. Systemen

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

TV-Anytime

Multimedia

ID3

JPEG 2000

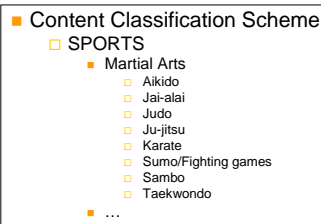
MPEG-7

MPEG-21

TV-Anytime

Ziel: Genaueres Suchen

Anwendung: Multimedia-Server



Quelle: The TV-Anytime Forum, Specification Series: S-3.
On: Metadata (Normative), Part A: Metadata Schemas.

- Einfaches Kategoriensystem nach MPEG-7
- XML, extern

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Dublin Core Metadata Initiative

Multimedia

ID3

JPEG 2000

MPEG-7

MPEG-21

TV-Anytime

Ziel: Genaueres Suchen

Anwendung: WWW-Dokumentensuche

```
<meta name="DC.format"      scheme="DCTERMS.IMT"      content="text/html">
<meta name="DC.type"       scheme="DCTERMS.DCMIType"  content="Text">
<meta name="DC.publisher"  content="Jimmy Whales">
<meta name="DC.subject"    content="Dublin Core Metadaten-Elemente">
<meta name="DC.creator"    content="Björn G. Kulms">
```

Quelle: http://de.wikipedia.org/wiki/Dublin_Core

- Text in 15 vordefinierten Tags
- XML, in Ressource

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Learning Object Metadata

Multimedia

ID3

JPEG 2000

MPEG-7

MPEG-21

TV-Anytime

Ziel: Genaueres Suchen

Anwendung: Lernobjekt-Suche im e-Learning

■ Metadata Categories

1. General
2. Lifecycle
3. Meta-Metadata
4. Technical
5. Educational
6. Rights
7. Relation
8. Annotations
9. Classification

- Text in 60 vordefinierten Tags
- XML, in Ressource

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Resource Description Framework

Multimedia

ID3

JPEG 2000

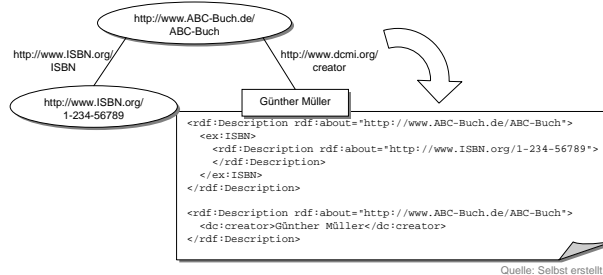
MPEG-7

MPEG-21

TV-Anytime

Ziel: Suchen / Wissensrepräsentation

Anwendung: Semantisches Netz



- Tripel Subjekt-Prädikat-Objekt
- frei erweiterbar
- notationsneutral, extern

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

RDF Schema

Multimedia

ID3

JPEG 2000

MPEG-7

MPEG-21

TV-Anytime

Ziel: Ontologie-Erstellung für RDF

Anwendung: Semantisches Netz

- erweitert RDF um die Funktion, Ontologien mit Vererbung zu erstellen
- basiert auf RDF

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Web Ontology Language

Multimedia

ID3

JPEG 2000

MPEG-7

MPEG-21

TV-Anytime

Ziel: Ontologie-Erstellung für RDF

Anwendung: Semantisches Netz

```
<owl:Class rdf:ID="Porina">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Trichotheliaceae"/>
  </rdfs:subClassOf>
</owl:Class>
```

Quelle: FungalTaxonomy.owl

- erweitert RDF um die Funktion, beliebig komplexe Ontologien zu erstellen
- basiert auf RDF

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Synchronized Multimedia Integration Language

Multimedia

ID3

JPEG 2000

MPEG-7

MPEG-21

TV-Anytime

Ziel: Inhaltsbeschreibung

Anwendung: Scriptsprache im WWW

```
<!-- Metadata about a scene of the video -->
<rdf:Description rdf:about="#scene1"
  dc:title="RDF intro"
  dc:description="Introduction to RDF functionalities"
  dc:language="en"
  smilmetadata:duration="30 secs"
  <smilmetadata:containsShots>
    <rdf:Seq rdf:id="ChronologicalShots">
      <rdf:li>Panorama-shot</rdf:li>
      <rdf:li>Closeup-shot</rdf:li>
    </rdf:Seq>
  </smilmetadata:containsShots>
</rdf:Description>
```

Quelle: Selbst erstellt.

Dublin
Core
SMIL-
Schema

WWW

DCMI

LOM

RDF

RDF Schema

OWL

SMIL

- Aufbau ähnlich wie Javascript
- Metadaten per RDF + RDF Schema
- XML, extern

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Fazit

- **Pro**
 - Brückenansatz über Semantischen Graben
 - vielseitig anwendbar
 - erfolgreiche Beispiele
- **Contra**
 - teuer und aufwändig
 - einige Ansätze kompliziert

Mein Fazit:

Metadaten sind viel versprechend, es werden aber dringend Tools benötigt, die sie auch verwenden.

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Danke für Ihre Aufmerksamkeit!

Noch Fragen?

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia

Quellen

- J. R. Smith, P. Schirling. **Metadata Standards Roundup**. *IEEE MultiMedia*, 13(2):84-88, IEEE Computer Society Press, Los Alamitos, Ca, USA, April 2006
- T. Bray. **RDF and Metadata**. 9. Juni 1998. Artikel auf XML.com, <http://www.xml.com/pub/a/98/06/rdf.html> WWW, ausgewertet am 18. Mai 2006
- G. Stamou, J. v. Ossenbruggen, J. Pan, G. Schreiber. **Multimedia Annotations on the Semantic Web**. *IEEE MultiMedia*, 13(1):86-90, IEEE Computer Society Press, Los Alamitos, Ca, USA, Januar-März 2006
- D. C. A. Bultermann. **Is it Time for a Moratorium on Metadata?** *IEEE MultiMedia*, 11(4):10-17, IEEE Computer Society Press, Los Alamitos, Ca, USA, Oktober-Dezember 2004
- H. Kosch, L. Bözörményi, M. Döller, M. Libsle, P. Schojer, A. Kofler. **The Life Cycle of Multimedia Metadata**. *IEEE MultiMedia*, 12(1), IEEE Computer Society Press, Los Alamitos, Ca, USA, Januar 2005
- C. Wroe, C. Goble, M. Greenwood, P. Lord, S. Miles, J. Papay, T. Payne, L. Morsiau. **Automating Experiments Using Semantic Data on a Bioinformatics Grid**. *IEEE Intelligent Systems*, 19(1):48-55, Januar/Februar 2004
- M. Horstmann, M. Lorenz, A. Warkowski, et al. **Automated interpretation and accessible presentation of technical diagrams for blind people**. *The New Review of Hypermedia and Multimedia*, 10(29:141-163, Taylor & Francis Inc., Pa, USA, 2004
- Aktuelle ID3v2.4.0 Spezifikationsdokumente auf ID3.org, <http://www.id3.org/develop.html> WWW, ausgewertet am 18. Mai 2006
- M. Boliek (Editor) et al. **JPEG 2000 Part II Final Committee Draft**. ISO / IEC JTC 1 / SC 29 / WG 1, JPEG 2000, 7. Dezember 2000. Verfügbar als PDF-Dokument über JPEG.org, <http://www.jpeg.org/public/tcd15444-2.pdf> WWW, ausgewertet am 18. Mai 2006
- J. M. Martínez (Editor). **MPEG-7 Overview (version 10)**. ISO / IEC JTC 1 / SC 29 / WG 11, Oktober 2004. Verfügbar über MPEG-Website, <http://www.chiariglione.org/MPEG/standards/mpeg-7/mpeg-7.htm> WWW, ausgewertet am 18. Mai 2006
- J. van Ossenbruggen, F. Nack, L. Hardman. **That Obscure Object of Desire: Multimedia Metadata on the Web (part I)**. *IEEE MultiMedia*, 11(4):38-48, IEEE Computer Society Press, Los Alamitos, Ca, USA, Oktober 2004
- J. van Ossenbruggen, F. Nack, L. Hardman. **That Obscure Object of Desire: Multimedia Metadata on the Web (part II)**. *IEEE MultiMedia*, 12(1):54-63, IEEE Computer Society Press, Los Alamitos, Ca, USA, 2005
- J. Bormans, K. Hill (Editoren). **MPEG-21 Overview v.5**. ISO / IEC JTC 1 / SC 29 / WG 11, Shanghai, Oktober 2002. Verfügbar über MPEG-Website, <http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm> WWW, ausgewertet am 18. Mai 2006
- The TV-Anytime Forum. **Specification Series: S-3. On: Metadata (Normative), Part A: Metadata Schemas**. 15. Dezember 2002. Verfügbar über TV-Anytime.org, <ftp://ftp.bbc.co.uk/Specifications/SP003v13-withWORD.zip> FTP-Server, ausgewertet am 18. Mai 2006
- Dokumentation NewsML 1.2 Guidelines V 1.0 auf NewsML.org, http://www.newsml.org/IPTCNewsML1.2/documentation/NewsML_1.2-doc-Guidelines_1.00.pdf PDF-Dokument, ausgewertet am 18. Mai 2006
- Standard-Dokumentation auf SportsML.com, <http://www.sportsml.com/IPTC/SportsML1.6/documentation/sportsml-comments.html> WWW, ausgewertet am 18. Mai 2006
- Verschiedene Dokumente auf DublinCore.org, <http://dublincore.org/> http, ausgewertet am 18. Mai 2006
- W. Hodgins (Vorsitz), E. Duval (techn. Editor) et al. **Draft Standard for Learning Object Metadata**. IEEE Standards Department, 19. Juli 2002. Verfügbar über http://tsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf PDF-Dokument, ausgewertet am 18. Mai 2006
- M. McClelland. **Metadata Standards for Educational Resources**. *Computer*, 36(11):107-109, IEEE Computer Society Press, Los Alamitos, Ca, USA, November 2003
- D. Becket (Editor). **RDF/XML Syntax Specification (Revised)**. W3C Recommendation, 10. Februar 2004. Verfügbar über W3.org, <http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/> WWW, ausgewertet am 18. Mai 2006
- D. Brickley, R. V. Guha (Editoren). **RDF Vocabulary Description Language 1.0: RDF Schema**. W3C Recommendation, 10. Februar 2004. Verfügbar über W3.org, <http://www.w3.org/TR/2004/REC-rdf-schema-20040210/> WWW, ausgewertet am 18. Mai 2006
- S. Bechhofer, F. v. Harmelen, J. Hendler, I. Horrocks, D. L. McGuinness, P. F. Patel-Schneider, L. A. Stein. **OWL Web Ontology Language Reference**. W3C Recommendation, 10. Februar 2004. Verfügbar über W3.org, <http://www.w3.org/TR/2004/REC-owl-ref-20040210/> WWW, ausgewertet am 18. Mai 2006
- D. Bultermann, G. Grassel (Editoren) et al. **Synchronized Multimedia Integration Language (SMIL 2.1)**. W3C Recommendation, 13. Dezember 2005. Verfügbar über W3.org, <http://www.w3.org/TR/2005/REC-SMIL2-20051213/> WWW, ausgewertet am 18. Mai 2006
- A. James, M. Christel, S. Gilles, R. Sarukkai, W.-Y. Ma. **Multimedia Information Retrieval: What is it, and why isn't anyone using it?** The MIR 2005 Panel.

1. Juni 2006

Florian Hoffmann – Metadaten für Multimedia