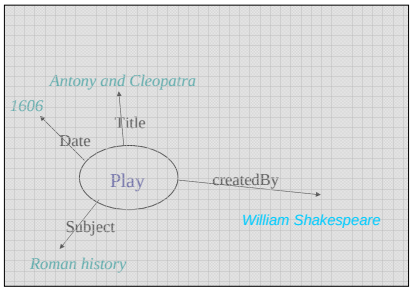
Ressource mit Eigenschaften

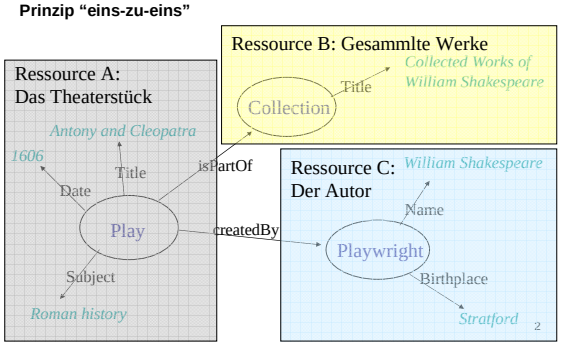
Ein Theaterstück mit Titel “Antony and Cleopatra”. Verfasser William Shakespeare am Datum 1606



1

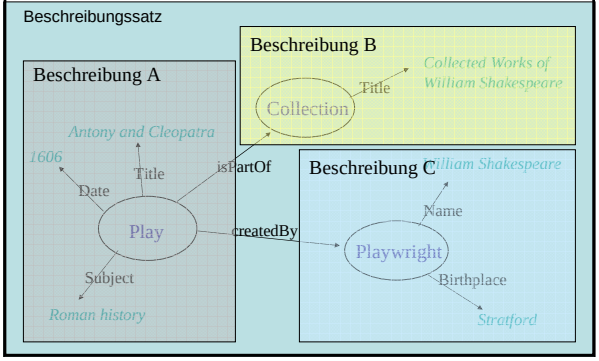
…auf andere Ressourcen bezogen

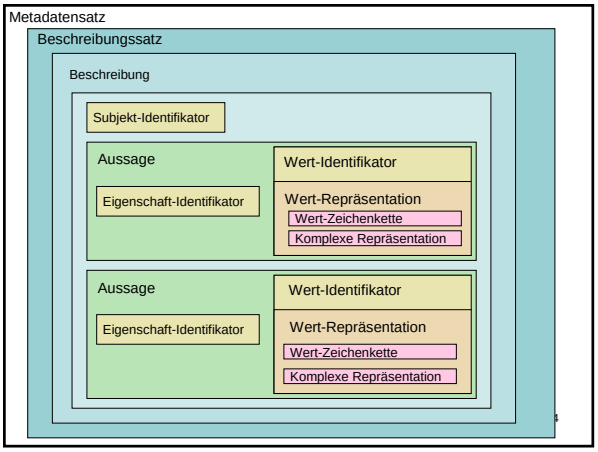
Prinzip “eins-zu-eins”



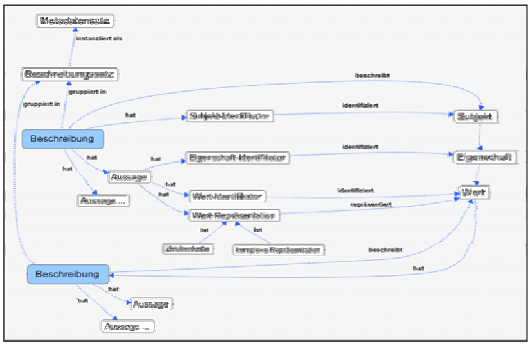
2

Drei Beschreibungen in einem Beschreibungssatz

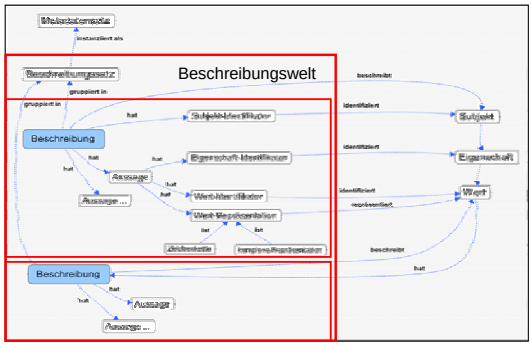




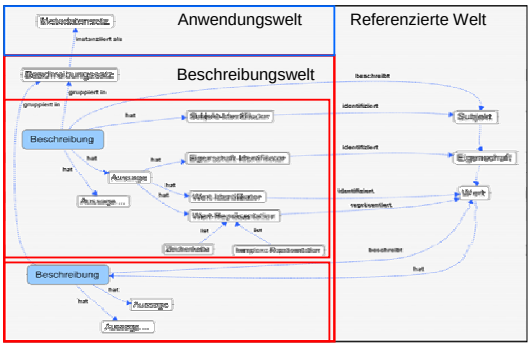
DCMI-Abstraktmodell



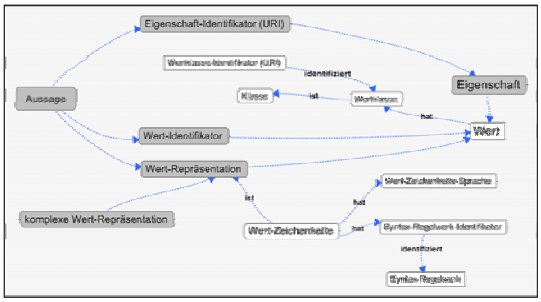
DCMI-Abstraktmodell



DCMI-Abstraktmodell

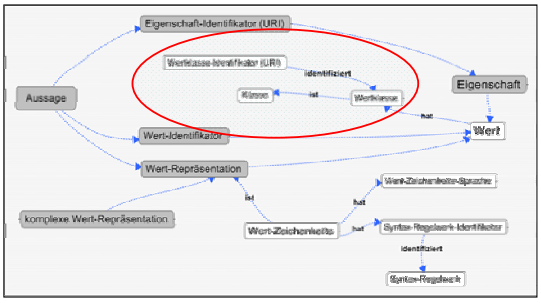


Encoding Schemes



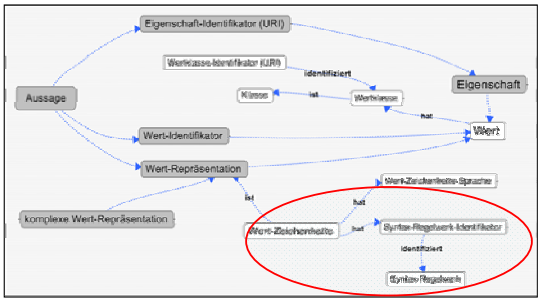
8

Vocabulary Encoding Scheme heißt Wertklasse



9

Syntax Encoding Scheme heißt Syntax-Regelwerk



10

W3C Semantic Web Activity

Semantic Web Best Practices and Deployment

Report from the
W3C Semantic Web
Best Practices Working Group

Thomas Baker, Göttingen State and University Library
Vienna Meeting: W3C and Semantic Web
Vienna, 20 June 2005

11

http://www.w3.org/2001/sw/BestPractices

- Begun 2004 “to provide hands-on support for developers of Semantic Web apps”
  - Working Group Notes (and drafts)
  - Applications and Demos Weblog
  - Semantic Web Tutorials
- Manages “SKOS Core” – a vocabulary for describing simple knowledge structures

12

Task forces

- RDF-in-HTML
- Applications and demos
- Ontology-engineering patterns
- Porting thesauri
- XML schema datatypes
- RDF/Topic Maps interoperability
- Vocabulary management
- WordNet
- Software engineering
- Link with Ontology Definition Metamodel activity

13

This talk

- RDF-in-HTML
- Applications and demos
- Ontology-engineering patterns
- Porting thesauri
- XML schema datatypes
- RDF/Topic Maps interoperability
- Vocabulary management
- WordNet
- Software engineering
- Link with Ontology Definition Metamodel activity

14

RDF-in-HTML Task Force

- http://www.cs.vu.nl/~guus/public/carroll-rdf-html.pdf
- Problem: embedding self-descriptions in Web documents
  - Who wrote the document, its title, etc
  - License restrictions (Creative Commons)
- Existing solutions are problematic
  - <meta> tags since HTML4 (1999)
  - Links to external RDF files
  - RDF (redundantly) in <HEAD> or comments

15

Example: embedded IP rights

16

RDF/XML in comment

- **<!-- **
- <rdf:RDF xmlns="http://web.resource.org/cc/">
- xmlns:dc="http://purl.org/dc/elements/1.1/">
- xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
- <Work rdf:about="">
- <dc:title>Caffiend</dc:title>
- <dc:date>2004</dc:date>
- <dc:description>Theme song for Ducksnorts baseball blog.</dc:description>
- <dc:type rdf:resource="http://purl.org/dc/dcmitype/Sound" />
- <license rdf:resource="http://creativecommons.org/licenses/by-nd/2.0/" />
- </Work>
-
- <**License rdf:about="http://creativecommons.org/licenses/by-nd/2.0/">
**
- **<permits rdf:resource="http://web.resource.org/cc/Reproduction" />
**
- **<permits rdf:resource="http://web.resource.org/cc/Distribution" />
**
- **<requires rdf:resource="http://web.resource.org/cc/Notice" />
**
- **<requires rdf:resource="http://web.resource.org/cc/Attribution" />
**
- </License>

17

RDF-in-HTML solutions under development

- GRDDL
  - Using stylesheets to generate RDF descriptions from XHTML markup
- New W3C specification for XHTML 2.0
  - Aims at compatibility with RDF
    - http://www.w3.org/TR/2005/WD-xhtml2-20050527/
    - Methods for layering RDF on XML/XHTML documents

18

Semantic Web assumptions

- Good descriptions and annotations (metadata) improve the reusability of information
  - Precise and meaningful statements about conceptual aspects of their content
  - Overcome limitations of content-based search engines
- Implies shared vocabularies of concepts
- Need to express 'vocabularies' of concepts machine-understandably
- Need to share a model (grammar) for statements using those vocabularies

19

RDF is the foundation

- Great for simple statements about Web resources and things in the world
- The very simplicity of the base RDF vocabulary limits its expressivity
- RDF Vocabularies layered on top of RDF provide expressivity

20

OWL – powerful ontology engineering

- Can define complex conceptual structures
- Class-oriented, logically precise modeling
  - Demanding in terms of expertise, effort, cost
- SW Best Practices Ontology Engineering and Patterns Task Force
  - "Representing Classes as Property Values"
    - http://www.w3.org/TR/swbp-classes-as-values/
    - Options for using "classes" (e.g., subject headings) when describing resources using RDF Schema or OWL Full as opposed to OWL DL or OWL Lite

21

SKOS Core – "Semantic Web Lite"

- **S**imple **K**nowledge **O**rganisation **S**ystem
  - An RDF-based model for simple knowledge structures such as thesauri
  - Porting ("Webifying") thesauri: representing and sharing classifications, glossaries, thesauri developed in the Print World
  - Examples of existing knowledge structures:
    - In the Print World: Dewey Decimal Classification, Art and Architecture Thesaurus
    - In the Web World: DMOZ categories

22

SKOS Origins and Maintenance

- Product of SWAD-Europe
  - Project of EU IST Fifth Framework Program
- Community ("Open Source") Development
  - Participation of thesaurus standards developers
  - Mailing list public-esw-thes@w3.org
  - Wiki http://esw.w3.org/topic/SkosDev
- Maintained by SW Best Practices WG
  - Public Working Drafts (10 May 2005)
    - SKOS Core Guide
    - SKOS Core Vocabulary Specification
    - Quick Guide to Publishing a Thesaurus on the SW

23

Example: Entries in a Glossary - 1

From RDF Semantics Glossary:

Illustrates description of "concepts" with "labels" and "definitions"

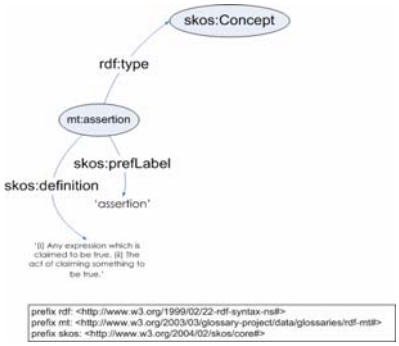
"Assertion"
Definition: "(i) Any expression which is claimed to be true. (ii) The act of claiming something to be true."

"Class"
Definition: "A general concept, category or classification. Something used primarily to classify or categorize other things."

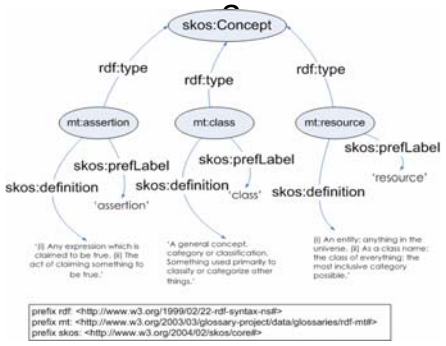
"Resource"
Definition: "(i) An entity; anything in the universe. (ii) As a class name: the class of everything; the most inclusive category possible."

24

Example: Entries in a Glossary -



Example: Entries in a Glossary -



Example: Taxonomies - 1

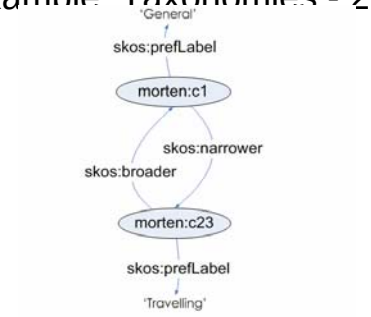
From MortenF's weblog categories

Illustrates "broader" and "narrower" properties.

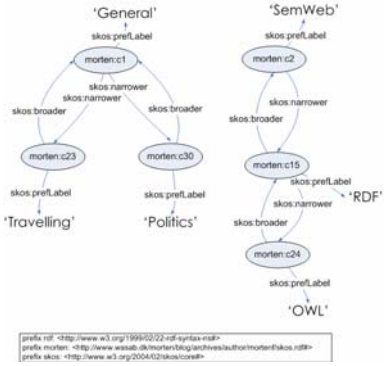
General
- Travelling
- Politics

SemWeb
- RDF
  - OWL

Example: Taxonomies - 2



Example: Taxonomies - 3



Example: Thesaurus - 1

From UK Archival Thesaurus ...

Term: **Economic cooperation**

Used For: Economic co-operation

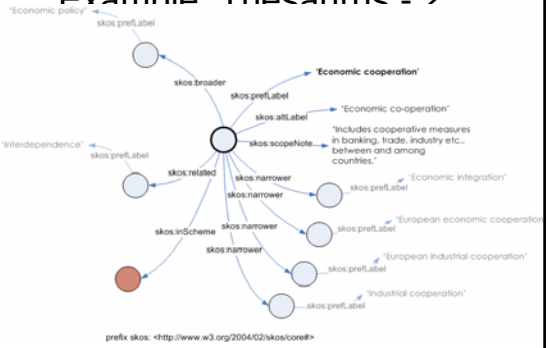
Broader terms: Economic policy

Narrower terms: Economic integration, European economic cooperation, European industrial cooperation, Industrial cooperation

Related terms: Interdependence

Scope Note: Includes cooperative measures in banking, trade, industry etc., between and among countries.

Example: Thesaurus - 2



SKOS Core Overview

- Basic description
  - Concept, ConceptScheme, inScheme, hasTopConcept
- Labelling
  - prefLabel, altLabel, hiddenLabel, prefSymbol, altSymbol … (audio labels?)
- Documentation
  - definition, scopeNote, changeNote, historyNote, editorialNote, publicNote, privateNote
- Semantic relations
  - broader, narrower, related
- Subject indexing
  - subject, isSubjectOf, primarySubject, isPrimarySubjectOf
- Grouping
  - Collection, OrderedCollection, CollectableProperty, member, memberList

Complementarity of OWL and SKOS

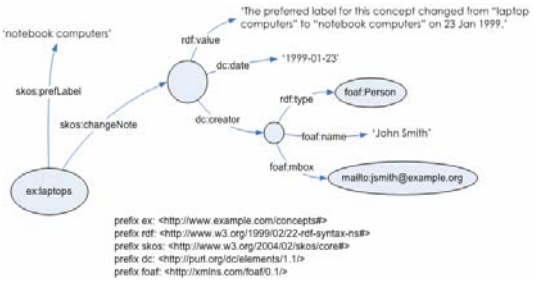
- OWL precision not always necessary or appropriate
  - OWL a sledge hammer / SKOS a nutcracker
  - OWL a Harley / SKOS a mountain bike
  - Complement each other, can be used in combination to optimize cost/benefit
- SKOS
  - Bridging the worlds of library classification and Web technology
  - Simple enough to be undemanding in terms of cost and required expertise

SKOS World View

Extensibility of SKOS

- Not “take it or leave it”, but “take what you want, create what you need”
- Extension “by refinement”
  - Define sub-properties or sub-classes of SKOS Core properties or classes
  - Extensions are backwards-compatible via RDFS inference
- Extension “by combination”

Combining SKOS with other vocabularies



Core and extension vocabularies

- Core-like RDF vocabularies
  - SKOS Core: about knowledge systems
  - Dublin Core: about information resources
  - FOAF: about people and organizations
  - Shared underlying (RDF) model
  - Shared mechanisms for extensibility
- Extension-like vocabularies
  - Web as context for community creation of vocabularies

37

Vocabulary Management Task Force

- Lessons from FOAF, DC, SKOS for other vocabulary communities
- How to declare an RDF vocabulary
  - Identify your terms with URI references
  - Articulate your maintenance policies and expectations for stability and persistence
  - Identify historical versions
  - Document your terms on Web pages
  - Share an RDF schema of your terms for processing by machines

38

Diverse vocabularies in a common model: Dublin Core at 10 years

Thomas Baker
DC-2005: Vocabularies in Practice
Madrid, 12 September 2005

39

1994: The Need

- **Web had arrived**: WWW conference, Chicago
- Clear: librarians would not "scale" to what was coming…
- "We need a **simple template** for describing Web pages"
- Librarians and IT people met in Dublin, Ohio…

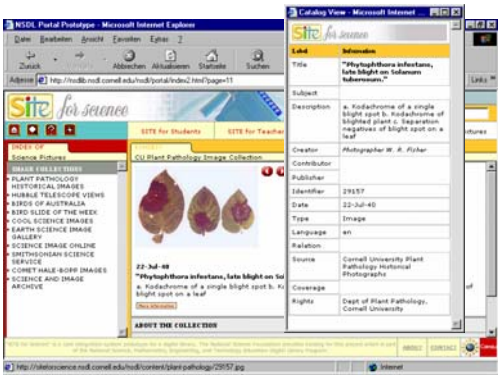
40

1995: "The Dublin Core"

- Title, Date, Creator, Subject…
- Simple enough for non-experts to understand
- Seen as "library catalog card" for Web objects

41

42



43

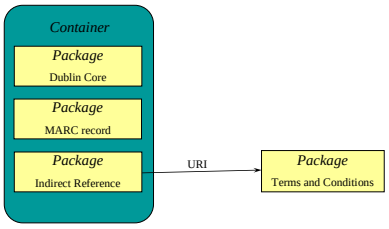
1996: Modular metadata

- Not: “One size fits all”.
- Rather: many ways to describe one object:
  - MARC: complete, for library items
  - DC: simple, for helping people find things
  - Terms and conditions: who allowed to use it?
- Recognized need: General framework for different types of metadata

44

1996: Putting DC in a context

- Diverse “packages” (Warwick Framework idea)



Work starts on general framework for metadata: Resource Description Framework (RDF), Semantic Web

45

1997: Simple, but also precise

- Not just any Date, but Date *Created*
- Not just any Subject, but *Library of Congress* heading
- Dumb-down: ignore extra details to see just a “core” description

46

1998: DC elements in Web space

- http://purl.org/dc/elements/1.1/title: identifiers rooted in Web infrastructure
- New assumptions
  - How can machines and software process this automatically?
- Work on the underlying data model…

47

1999-2000: Application profiles

- **Customized implementations**
  - DC alone is too small – **more terms**
  - **More precision** for specialized purposes
  - **Local rules** and guidelines
- Application profile provides documentation so that others can follow
- Not “take it or leave it”, but **“take what you want, create what you need”**

48

1999-2000: Growing the vocabulary

Elements	Refinements		Encodings	Types
1. Identifier	Abstract	Is referenced by	Box	Collection
2. Title	Access rights	Is replaced by	DCMIType	Dataset
3. Creator	Alternative	Is required by	DDC	Event
4. Contributor	Audience	Issued	IMT	Image
5. Publisher	Available	Is version of	ISO3166	Interactive Resource
6. Subject	Bibliographic citation	License	ISO639-2	Moving Image
7. Description	Conforms to	Mediator	LCC	Physical Object
8. Coverage	Created	Medium	LCSH	Service
9. Format	Date accepted	Modified	MESH	Software
10. Type	Date copyrighted	Provenance	Period	Sound
11. Date	Date submitted	References	Point	Still Image
12. Relation	Education level	Replaces	RFC1766	Text
13. Source	Extent	Requires	RFC3066	
14. Rights	Has format	Rights holder	TGN	
15. Language	Has part	Spatial	UDC	
	Has version	Table of contents	URI	
	Is format of	Temporal	W3CTDF	
	Is part of	Valid		

49

2001: Usage Board

- Editorial reviewers to grow vocabulary
- Balance **bottom-up** with **top-down**
- New terms from working groups and projects approved as DCMI terms

50

2003-2005: Abstract Model

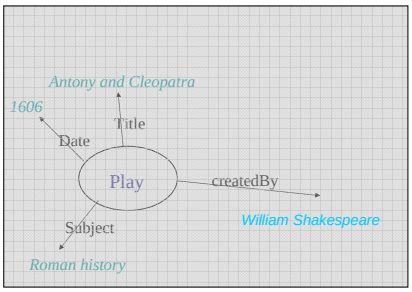
- **Formal modeling basis** for Dublin Core metadata
- Like a “grammar” for Dublin Core
- Strong link with parallel development of Resource Description Framework

51

Basic model

Resource with properties

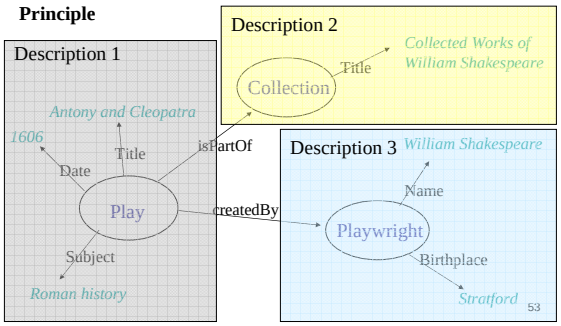
The Play has the title “Antony and Cleopatra”, was written In 1606 by William Shakespeare, and is about “Roman history”.



52

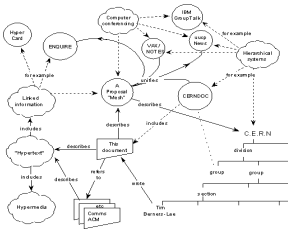
…related to other resources

One-to-one Principle

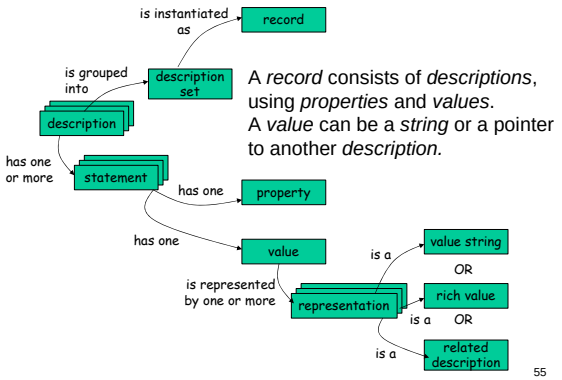


53

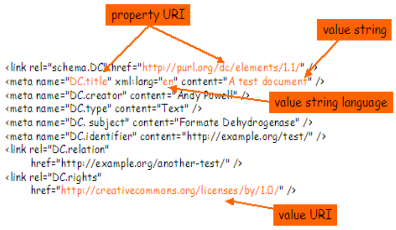
Same as Semantic Web model



54



A basis for comparing syntax options



Example of Simple Dublin Core in XHTML

One model, many ways to implement

- **Choose what's best for you**
  - HTML, XML, XHTML, Databases…: appropriate for different contexts
- Interoperability depends on **shared model**
  - Evolving guidelines for Dublin Core in HTML, XML, XHTML…

DCMI Model today

- **Core semantics**
  - The Dublin Core and related vocabularies
- **Data model**
  - Abstract Model a DCMI Recommendation since March 2005
- **Application Profiles**
  - Use core semantics
  - Based on the Abstract Model
  - Customize for specific purposes

Three legs of Interoperability

- **Shared semantics**
  - Everyone understands a Core
- **Shared model** (independent of syntax)
  - DCMI Abstract Model
- **Content-level agreement**
  - Same way to write a date ("2005-09-12"), a name ("Baker, Thomas")…

Like a language

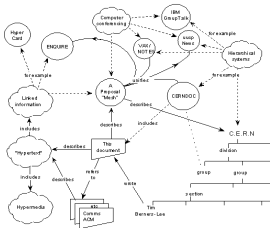
- **Small vocabulary** (like a pidgin)
  - Without shared words, no understanding
- **Basic grammar**
  - Without Subject – Verb – Object, no sentences
- **Jargons and styles**
  - Emerging agreement within communities

DCMI Today

- **Core elements** for describing information
  - Find, share, manage information
- **Formal framework** for interoperability
  - The "grammar" of metadata language
- **Support for communities**
  - Community usage in application profiles
  - Documentation for usage in different areas
  - Platforms for people to come together to share experiences.

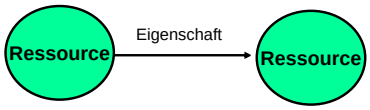
61

Semantic Web,1990



62

Grammatik für Metadatenaussagen, 1998

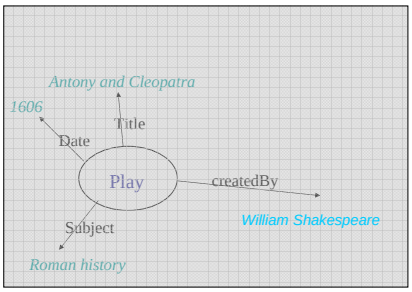


http://dublincore.org/workshops/dc6/pp/miller-datamodel.ppt, 1998

63

Ressource mit Eigenschaften

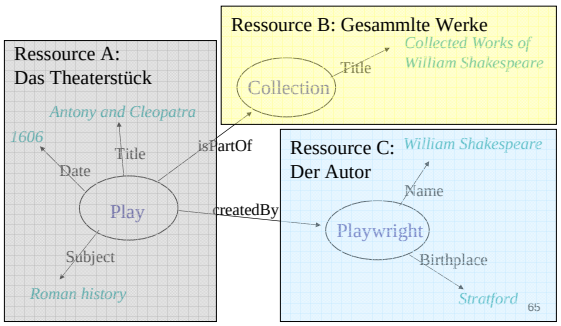
Ein Theaterstück mit Titel "Antony and Cleopatra". Verfasser William Shakespeare am Datum 1606



64

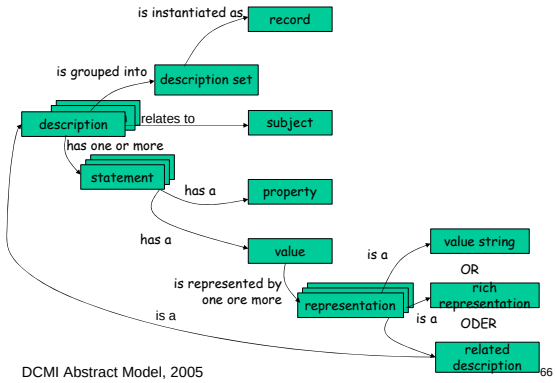
…auf andere Ressourcen bezogen

Prinzip "eins-zu-eins"

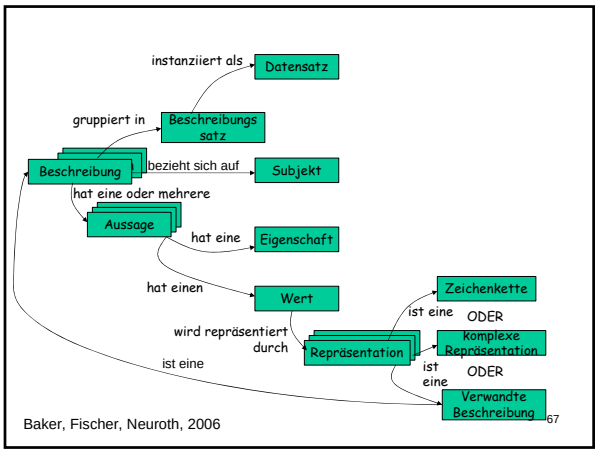


65

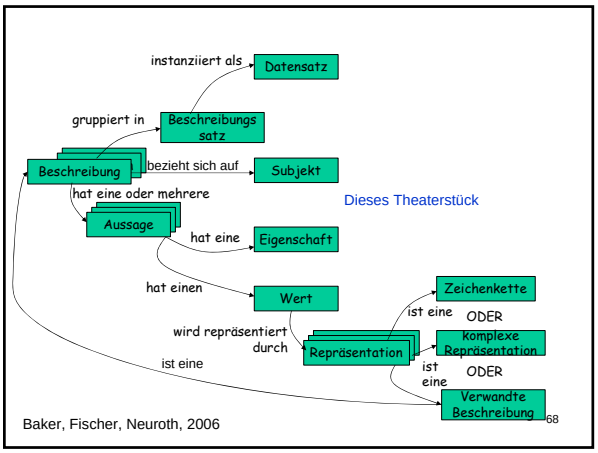
DCMI Abstract Model, 2005



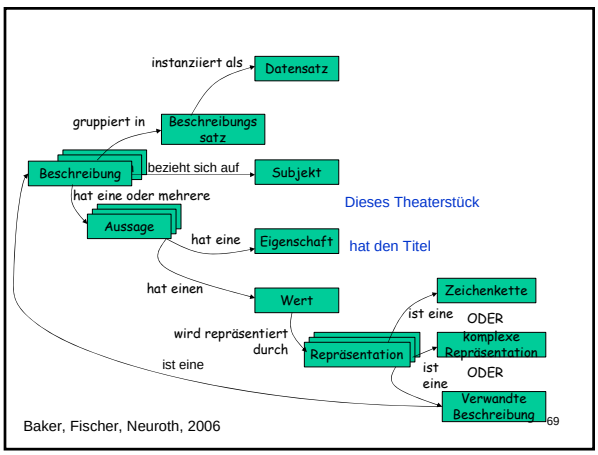
66



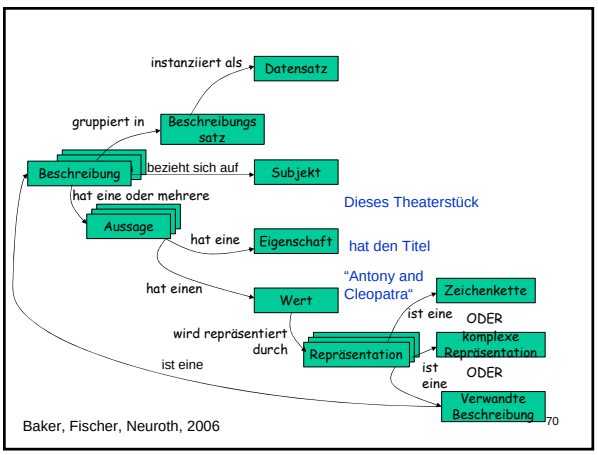
Baker, Fischer, Neuroth, 2006 67



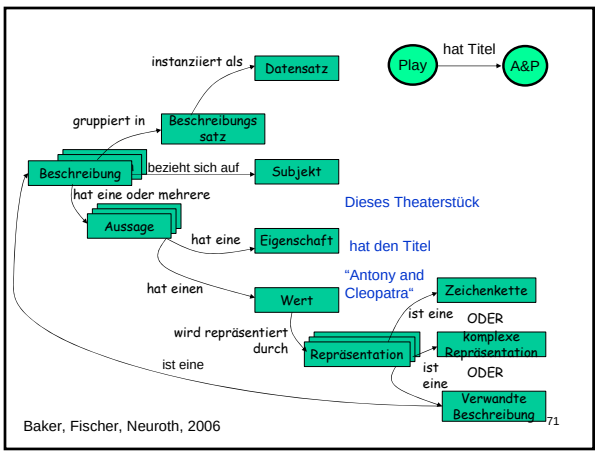
Baker, Fischer, Neuroth, 2006 68



Baker, Fischer, Neuroth, 2006 69

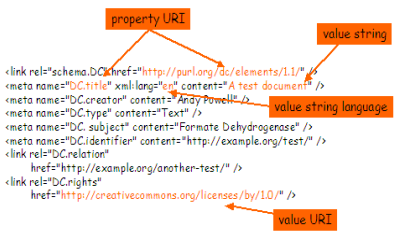


Baker, Fischer, Neuroth, 2006 70



Baker, Fischer, Neuroth, 2006 71

Gemeinsames Grundmodell – Vielfältige Implementierungstechniken

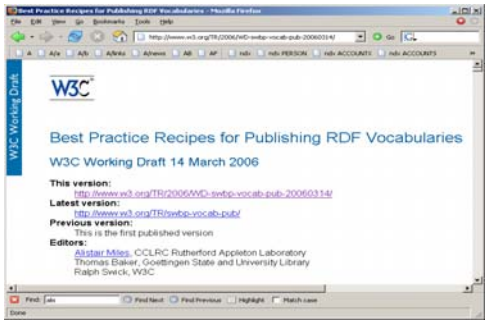


Was kann sinnvoll gemappt werden?

Beispiel: "Simple Dublin Core" in XHTML

72

73



74

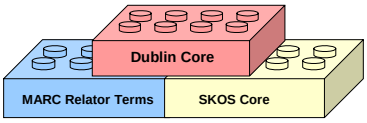
Praxis-Richtlinien für Metadatenvokabulare

- Maschinenenverarbeitbare RDF-Vokabulare
  - Dublin Core, SKOS Core, FOAF, MARC Relators…
- Parallel als Webseiten (für Menschen) und RDF (für die Maschinen)
  - "Kochbuch" mit "Rezepten" zur Konfiguration eines Apache-Servers
- Auf allgemeingültige Prinzipien harmonisieren
  - URI-Vergabe
  - Versionierung…

75

Baukastenprinzip

- Modularität, Flexibilität, Wiederverwendbarkeit, Interoperabilität der Metadaten
- Modellbasierte Rekombinierbarkeit



76

Mit DCMI ist es uns gelungen:

- Eine Wissensgemeinschaft zu mobilisieren
- Eine bekannte Marke zu etablieren
- Ein Gremium für die Weiterentwicklung eines ISO-Standards zu etablieren
- Eine internationale Konferenzserie aufzubauen
- Eine internationale Organisation zu entwickeln:
  - Stakeholder sind im Aufsichtsrat vertreten
  - DCMI-Affiliates bilden Kompetenz auf regionaler Ebene
- Bibliotheken, Regierungen, Behörden, Firmen zu interessieren

77

baker@sub.uni-goettingen.de

78