



Planung des Forschungsdaten-Managements: Metadaten und Identifikatoren

Timo Gnad
SUB Göttingen
gnadt@sub.uni-goettingen.de

6. Dezember 2011, Göttingen

GEFÖRDERT VOM



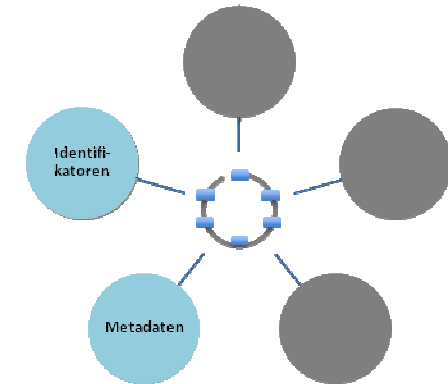
Bundesministerium
für Bildung
und Forschung



Überblick

Metadaten

- Motivation
- Planungsaspekte
- „Dimensionen“ von Metadaten

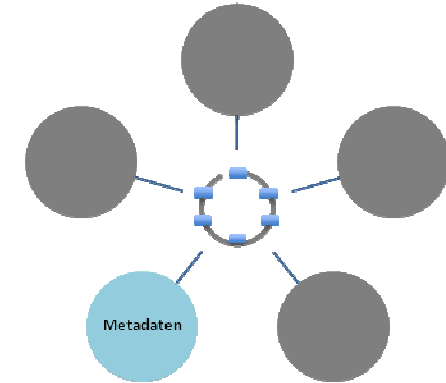
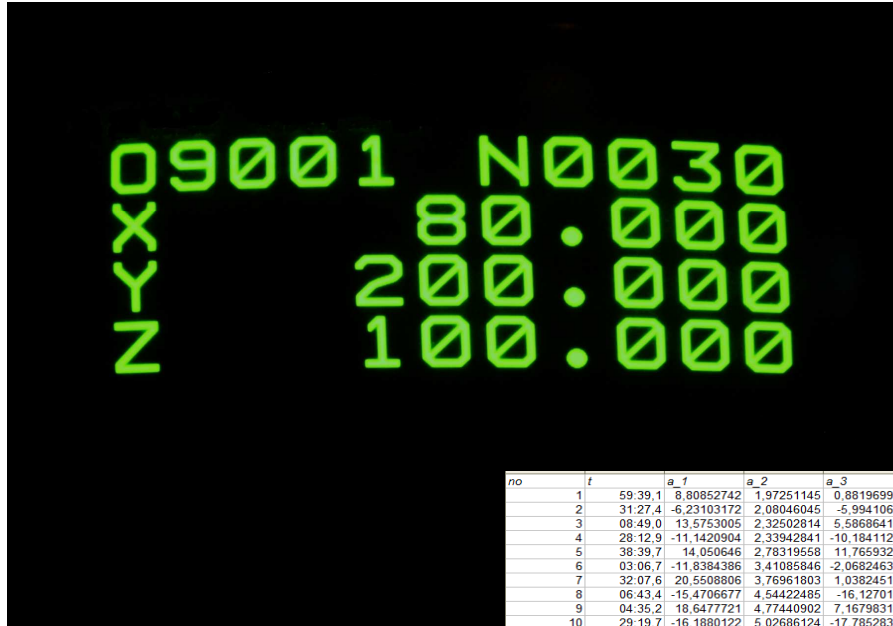


Identifikatoren und Informationsobjekte

- Was sind Identifikatoren?
- Planungsaspekte
- Verwendung und Auflösung



Motivation



no	t	a_1	a_2	a_3	low_s	high_s	front_rad	rear_rad	l_type	r_type
1	59.39.1	8.80852742	1.97251145	0.88196992	-18.5308868	17.3206605	1.95	99.37	f1	f2
2	31.27.4	-6.23103172	2.08046045	-5.9941069	-29.8743318	20.9521546	8.39	92.48	f1	f3
3	08.49.0	-13.5753005	2.32502814	5.58686418	24.0203049	29.9271356	12.57	89.00	f3	f1
4	28.12.9	-11.1420904	2.33942841	-10.1841123	-17.4042415	12.7567398	32.82	68.29	f1	f4
5	58.39.7	14.050646	2.78319558	11.7659322	-16.3919712	8.31209784	2.08	98.48	f2	f1
6	03.06.7	-11.8384386	3.41085846	-2.06824631	-23.8232119	23.5258713	0.57	100.55	f5	f2
7	32.07.6	20.5508806	3.76961803	1.03824515	-6.33254098	0.40728418	3.71	97.84	f2	f5
8	06.43.4	-15.4706677	4.54422485	-16.127016	-25.9316205	24.2721816	7.80	92.41	f1	f2
9	04.35.2	18.6477721	4.77440902	7.16798318	-2.32248349	4.18519581	83.02	17.83	f1	f3
10	29.19.7	-16.1880122	5.02686124	-17.7852831	7.12383087	15.679935	36.59	63.54	f1	f3
11	19.49.0	20.8052089	5.17090161	13.8597993	-28.4439866	18.5700769	1.90	98.98	f2	f1
12	53.25.6	-20.2882746	5.3802059	-8.53850749	-14.6679151	14.0664523	14.84	86.58	f3	f1
13	20.59.5	20.8933519	5.65145486	14.4020537	-24.1408717	17.6632725	4.53	96.11	f3	f5
14	17.17.5	-18.0255035	6.42283132	-8.00031855	-7.93354641	2.69418156	33.36	67.26	f3	f2
15	48.27.0	21.7028364	6.92651694	0.15318583	-15.0571695	9.31603059	89.12	11.05	f1	f4
16	55.53.6	-16.0448301	7.60662952	-6.11826178	-24.1205071	21.6158062	1.14	99.55	f4	f2
17	16.04.4	16.642596	7.63810602	-1.86416347	14.0820632	16.367242	14.72	86.23	f4	f4
18	44.57.5	-9.62408154	7.66674083	-5.15996053	7.96596918	10.61292	2.80	99.14	f1	f3
19	42.41.3	11.4530123	7.76589475	-0.88397324	-29.8759361	29.8575606	39.35	62.30	f1	f3
20	57.26.2	-7.30557556	8.3302276	-6.46507912	-18.2116402	13.3868636	63.87	37.99	f1	f5
21	38.08.6	11.9797574	8.38606586	-1.5607582	-15.6098092	6.69845149	7.16	93.87	f1	f3
22	43.12.6	-5.39447348	8.50476598	-5.88021996	-25.814806	18.8275613	3.87	97.34	f5	f1
23	53.44.5	14.1597736	9.00285124	0.54320278	-3.6112011	6.99701566	18.57	81.69	f2	f2
24	58.37.4	-4.63630885	9.19825619	-7.93512353	-10.7602103	6.81880681	9.44	91.87	f4	f1
25	26.47.2	7.63131249	9.49245926	0.31253235	24.198029	31.4417769	4.36	96.65	f5	f1
26	34.56.8	-7.18144542	10.0353152	-8.24741386	0.67726651	8.2876792	16.99	83.28	f1	f2
27	35.34.1	8.64014373	10.1235639	-3.43462166	-27.7645485	19.063625	0.79	101.17	f2	f4
28	23.58.4	-7.19911634	10.2704193	-10.9245425	-14.1503908	13.978891	9.52	90.89	f1	f3
29	58.46.1	8.99864657	11.1135519	2.46145235	-25.5829086	21.7947088	6.70	94.60	f1	f3
30	25.02.5	-1.4161269	11.3422121	-6.72800182	-29.9176139	21.4684751	2.73	97.98	f2	f5
31	23.09.5	7.0345392	11.5573695	-5.07144114	-0.32997323	3.12146098	30.42	71.10	f4	f1
32	34.51.2	-6.76082573	11.7289912	-11.473453	-29.2578838	22.3209727	12.88	88.73	f3	f2
33	36.33.3	14.2624841	12.2128325	4.26406545	-17.5222232	12.9352606	0.64	100.36	f3	f2
34	55.41.5	-10.0818754	12.5404322	-13.6728988	41.1801304	48.2819284	0.12	100.81	f5	f1
35	50.18.6	20.0737279	12.91326	6.73255231	-4.55377006	2.40780698	16.83	85.05	f1	f4
36	53.49.1	-17.2754713	12.9889775	-8.21428886	-23.9651531	17.1623398	63.09	37.21	f2	f4
37	59.09.5	21.6119922	13.0488678	9.5343767	55.857366	62.6479784	13.32	87.83	f2	f1
38	19.46.8	-18.7174264	13.0946298	-11.6009636	14.3046948	20.3270007	24.63	76.81	f2	f1
39	40.42.3	20.5122994	13.1487306	7.67273973	45.3022996	52.0068715	8.54	92.35	f2	f1
40	44.33.8	-12.2639788	13.6234446	-11.7811988	-7.9894689	6.89170842	15.46	85.53	f2	f1
41	05.07.1	12.8310451	13.9831314	-1.40297731	-0.90292354	8.43367843	45.11	55.63	f2	f2
42	11.45.4	-9.3797382	14.2127593	-11.7233707	-14.7074046	-14.076661	70.63	31.32	f3	f1
43	19.13.2	12.4174732	14.2338871	-2.51146614	10.1226859	16.5031371	36.15	63.96	f1	f3
44	24.43.3	-4.86528635	14.4412488	-8.54234572	16.9426005	18.5069259	29.14	71.40	f4	f1



Was sind Metadaten?

Metadaten dienen der strukturierten *Dokumentation* von Daten hinsichtlich deren

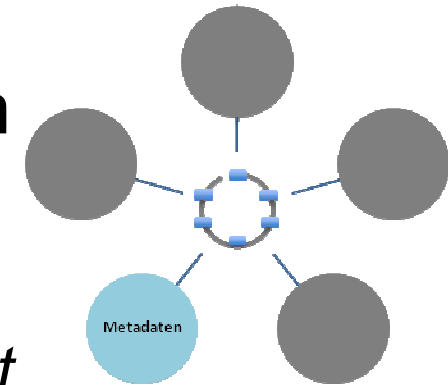
- *Erstellung, Verarbeitung, Kontext, Inhalt, Bedeutung, ...*

mit dem *Ziel* der

- *Verwaltung, Identifizierung/Überprüfung/Beglaubigung, Nachnutzung, ...*

und enthalten *Informationen* über die

- *Datensätze selbst, beteiligte Personen, Vorgänge, Zeiten, Systeme (Hardware / Software), Rechte, ...*





Was sind Metadaten?

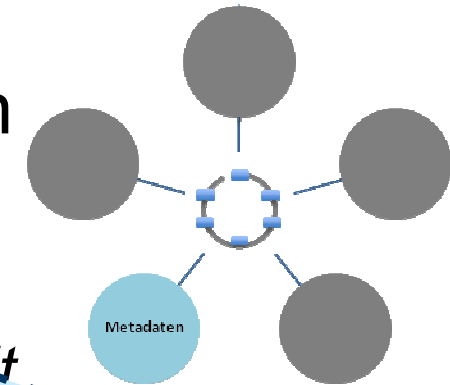
Metadaten dienen der strukturierten *Dokumentation von Daten*

hinsichtlich deren

- *Erstellung, Verarbeitung, Bedeutung, ...*

mit dem *Ziel* der

- *Verwaltung, Identifikation, Nachnutzung*



„Daten über Daten.“

„Informationen über Informationen.“

„...the who, what, when, where, why and how of the resource.“

Informationen über die

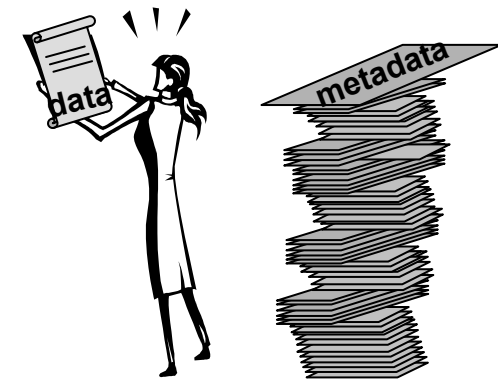
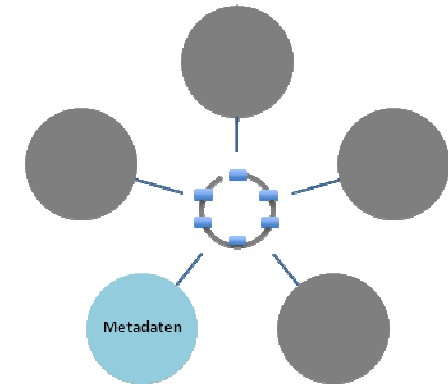
Autoren, Orte selbst, beteiligte Personen, Vorgänge, Zeiten, Systeme (Hardware / Software), Rechte, ...



Planungsaspekte

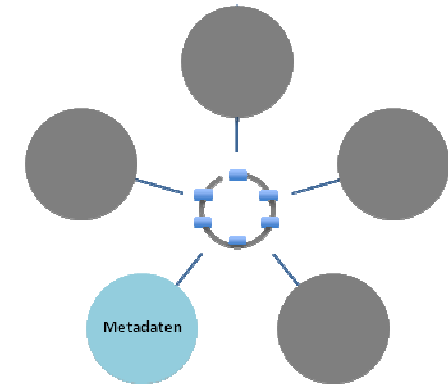
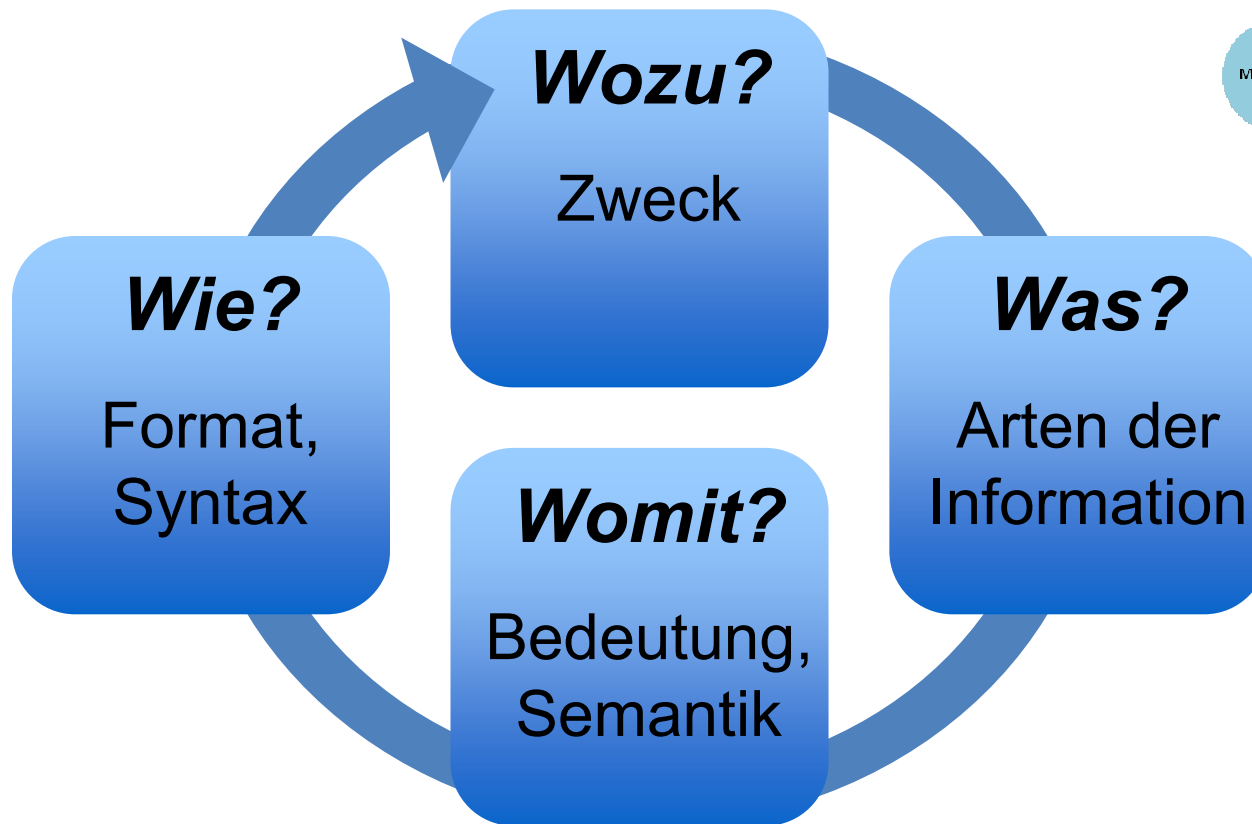
Wieviele ~~×~~ Metadaten brauchen wir?

- Übergang zwischen Domänen
- Zeitpunkt und Automatisierung der Erstellung
- Voraussetzungen
 - Hard-/Software
 - Vor-/Fachkenntnisse
- Standards





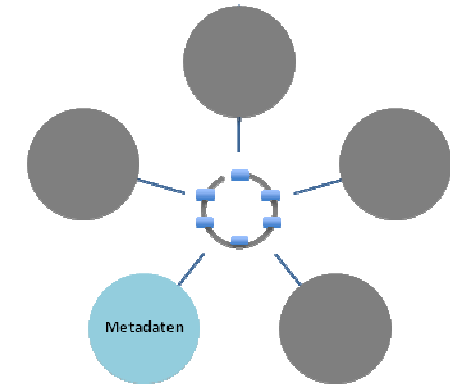
„Dimensionen“ von Metadaten





Zweck von Metadaten

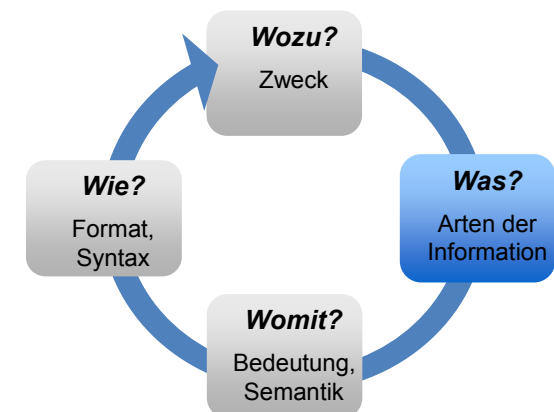
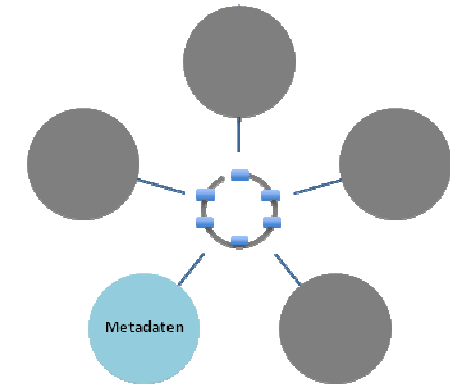
- Datenbestand *sichtbar* machen
- Daten *interpretierbar* machen
- Daten *austauschen*
- *Verwaltung* und *Pflege*
- *Präsentation*





Arten der Information

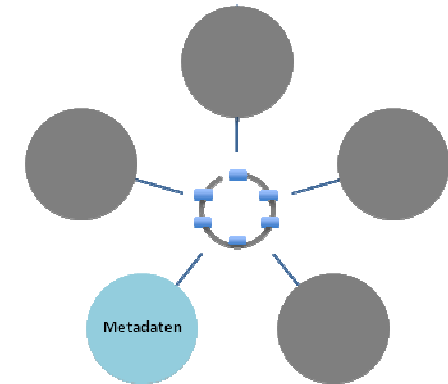
- Art/Typ der *Objekte*
- Beteiligte *Akteure*
- *Quellen* der Daten
- *Vorgänge* bei Verarbeitung
- *Ergebnisse*, z.B. für Präsentation





Semantik von Metadaten

- Rahmensystem
 - z.B. RDF, MARC, MODS, INSPIRE
- Universelle und fachspezifische Standards
 - erleichtern Austausch, Verwaltung



Typische Elemente:

- Identifikator / Kennzeichen
- Bezeichnung / Titel
- Datenformat
- Beteiligte Akteure
- Thematische Informationen

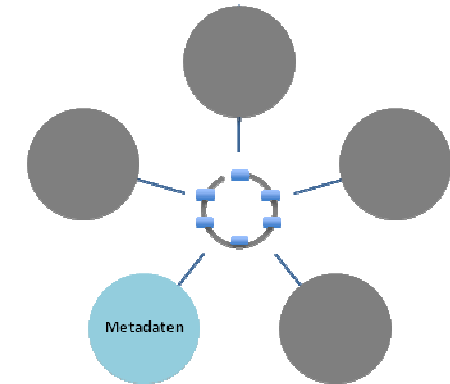




Dublin Core Metadata Initiative

Dublin Core Metadata Element Set (DCMES):

15 optionale, wiederholbare Elemente



ID: identifier

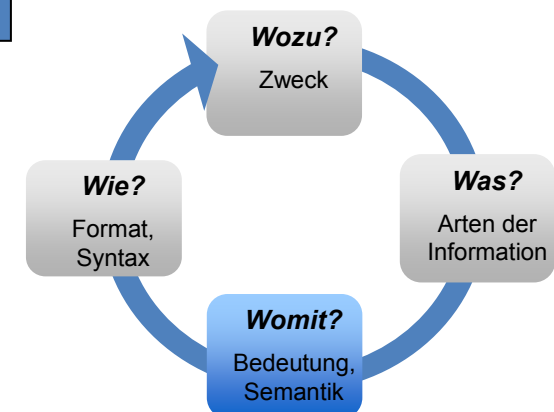
Technische Daten: format, type, language

Inhalt: title, subject, coverage, description

Personen und Rechte: creator, publisher, contributor, rights

Vernetzung: source, relation

Lebenszyklus: date

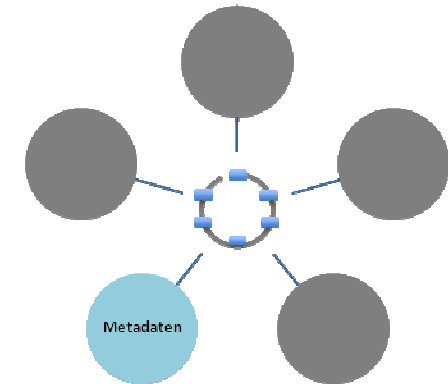




Dublin Core Metadata Initiative

Dublin Core Metadata Element Set (DCMES):

15 optionale, wiederholbare Elemente



ID: identifier

Technische Daten: format, type, language

Inhalt: title, subject, coverage, description

Personen und Rechte: creator, publisher, contributor, rights

Vernetzung: source, relation

Lebenszyklus: date

Erweiterbar auf 55 Elemente (DCMI Metadata Terms):

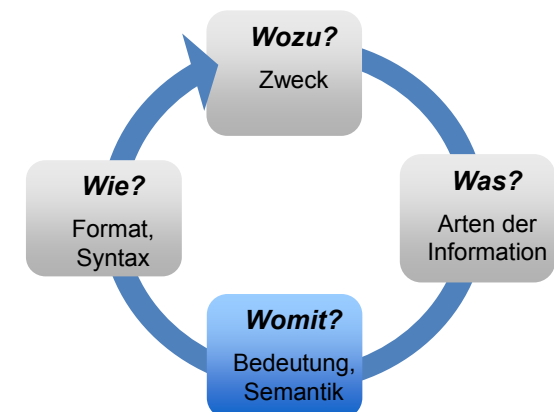
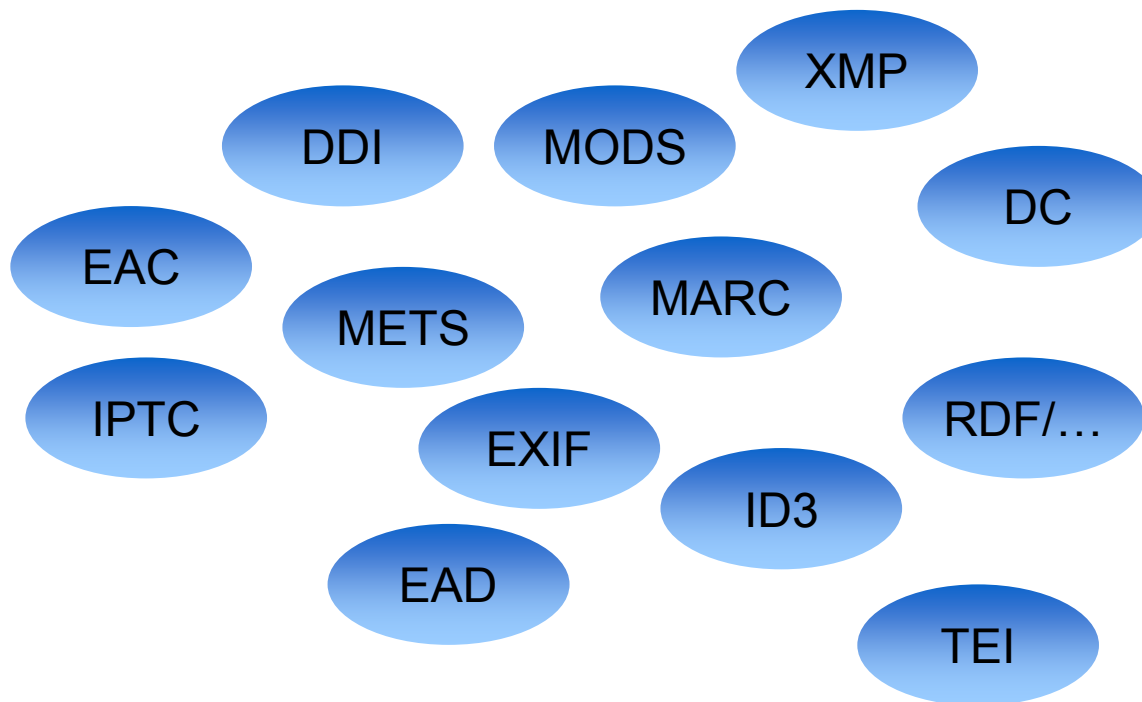
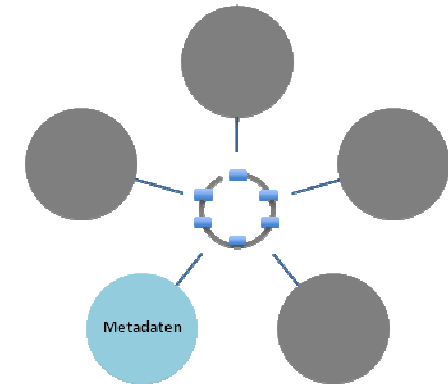
abstract, accessRights, accrualMethod, accrualPeriodicity, accrualPolicy, alternative, audience, available, bibliographicCitation, conformsTo, created, dateAccepted, dateCopyrighted, dateSubmitted, educationLevel, extent, hasFormat, hasPart, hasVersion, instructionalMethod, isFormatOf, isPartOf, isReferencedBy, isReplacedBy, isRequiredBy, issued, isVersionOf, license, mediator, medium, modified, provenance, references, replaces, requires, rightsHolder, spatial, tableOfContents, temporal, valid



Existierende Standards

Vorhandene Formate / Standards verwenden oder eigene erstellen?

→ Existierende Standards anpassen bzw. erweitern

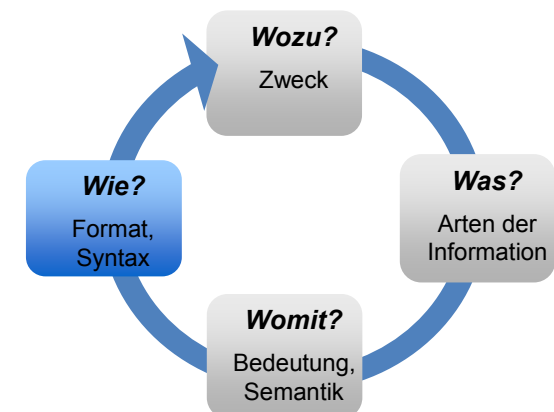
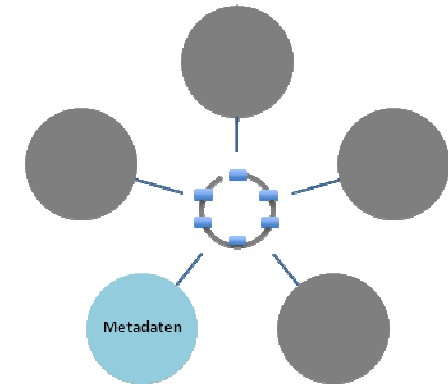




Syntax von Metadaten

Syntax für Erfassung, Speicherung, Austausch

- Transportabilität
- Interpretierbarkeit
- Datei- und Datenformat
 - Interne (Datenbank-) Formate
 - CSV
 - XML
 - RDF/XML
 - Graphen





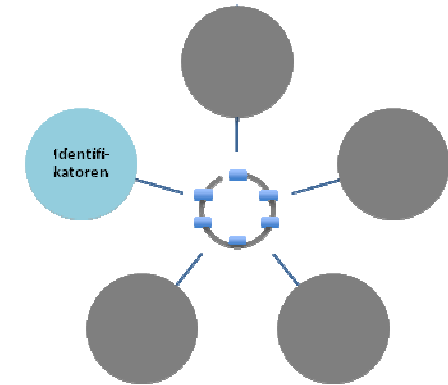
Identifikatoren und Informationsobjekte





Was ist ein Identifikator?

- Künstlich zugewiesenes Merkmal
- Dient der Identifizierung eines Objekts



Beispiele:

- DOI: 10.1594/WGCC/CLM A1B ZS
- URN: urn:nbn:de:gbv:089-3321752945
- PLZ: D-37073
- ISBN-13: 978-3-836460-10-1
- Telefonnummer (?): +49-551-39-3883
- Landesflagge (?):



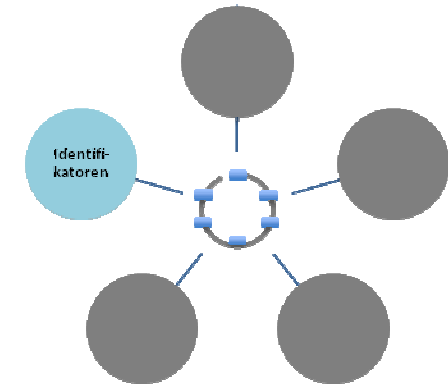
112





Planungsaspekte

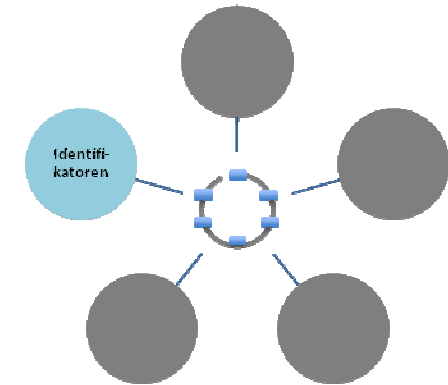
- *Welche Dinge existieren* im Forschungsbereich?
 - Daten, Instrumente, Akteure, Prozesse, ...
- *In welchen Verhältnissen* stehen die Objekte zueinander?
 - Hierarchie? Abhängigkeiten?
 - Ontologie / Informationsmodell?
- *Welche Dinge sind wichtig?*
 - Dauerhafte Identifikatoren (*Persistent Identifiers*) benötigt ?





Persistent Identifiers

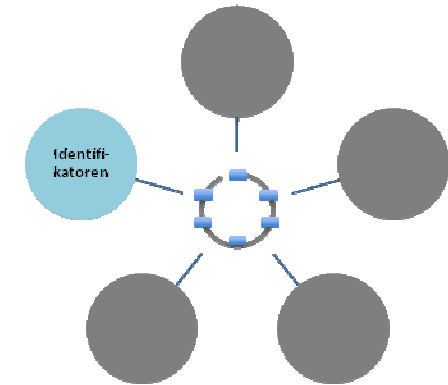
- *Weltweit gültige, dauerhafte Referenz auf bestimmtes Objekt*
 - auch wenn sich dessen Speicherort ändert
 - auch wenn eine neue Version erstellt wird
 - Beispiele: URN, DOI, ...
- *Setzt Verwaltung und Pflege voraus*
 - durch Resolver
 - durch Repository





Verwendung und Auflösung

- *Welche Identifikatoren werden benutzt (Standards/Syntax)?*
- *Wie und wo werden die Identifikatoren aufgelöst?*
 - Mögliche Ausfälle von Resolvern
- *Wann werden Dinge identifiziert?*
 - bei Erstellung, Bearbeitung, Archivierung?





Fragen und Kommentare?

Metadaten und Identifikatoren

Leitfaden, Abschnitt 2, Kap. 10/11

Fragenkatalog, Kap. 10/11