

101. Bibliothekartag 2012

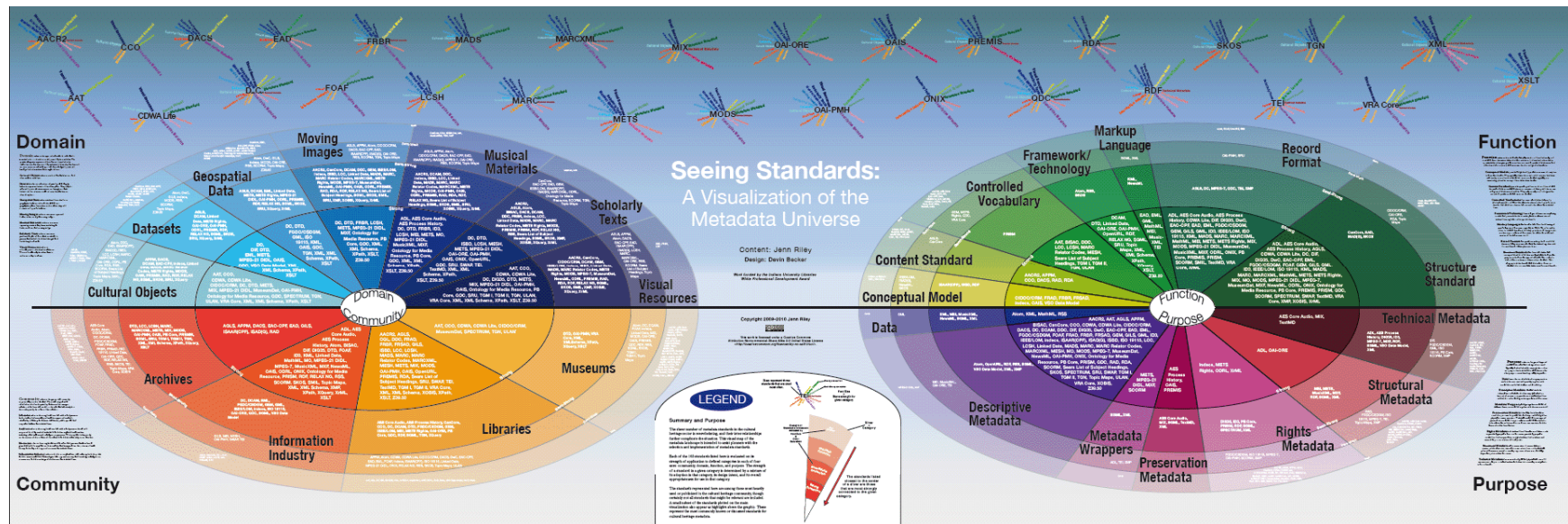
Metadaten als Basis für die semantische Vernetzung von Forschungsdaten

Günther Neher (FH Potsdam), Bernd Ritschel (GFZ Potsdam)

24.05.2012

Vorspann: Problematisierung Status-Quo

„Metadaten-Universum“

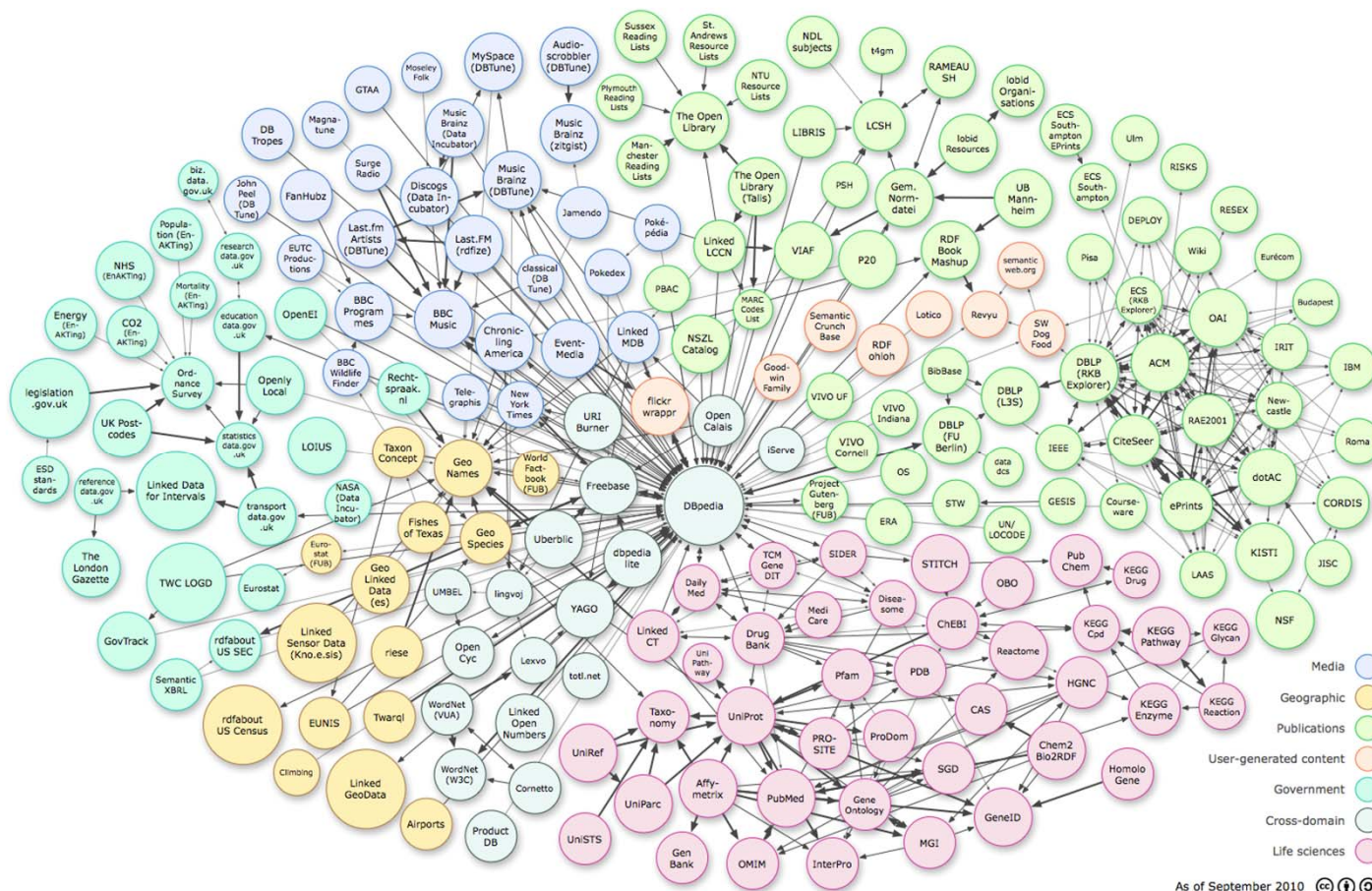


<http://www.dlib.indiana.edu/~jenrile/metadatamap/>

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Vorspann: Problematisierung Status-Quo

„Linked Data Cloud“

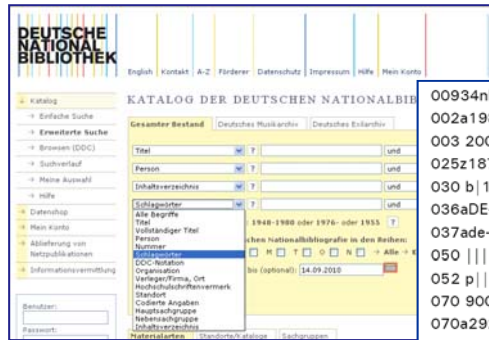


*„ In the Semantic Web,
it is not the Semantic
which is new,
it is the Web which is new“*

Chris Welty, IBM

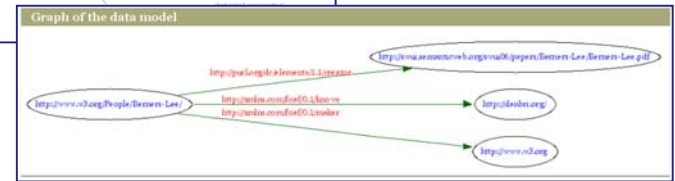
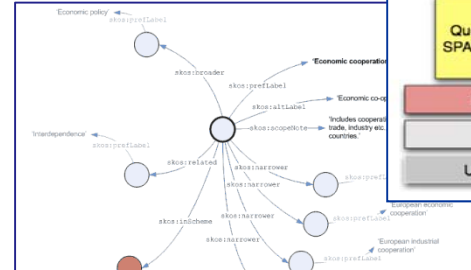
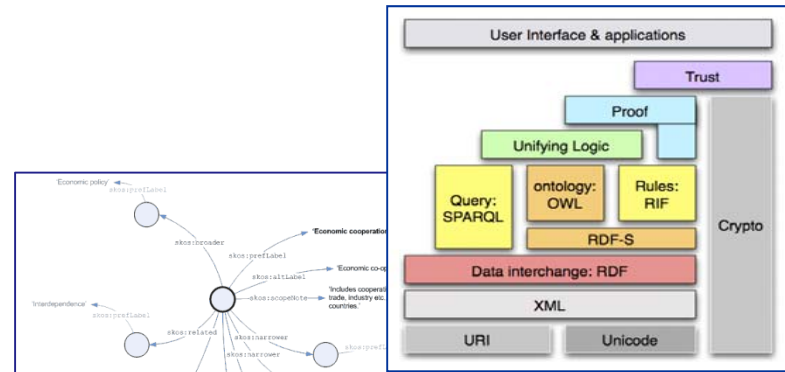
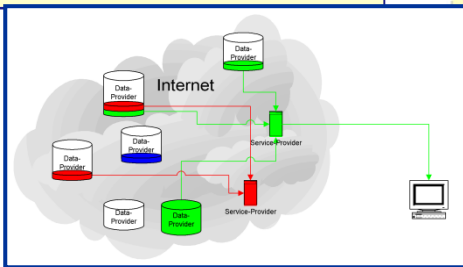
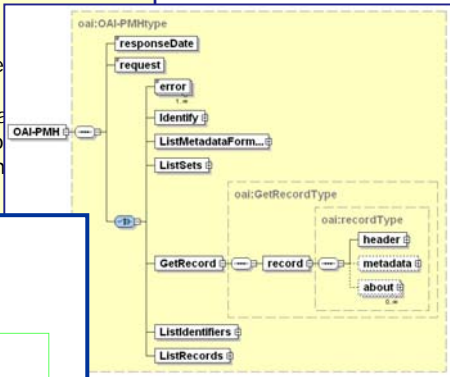
Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro



00934nM2.01200024 h001 187056-
 002a19890420-
 003 20021120131909-
 025z 187056-
 030 b |1dcz|z|||37-
 036aDE-
 037ade-
 050 ||||||g-
 052 p||||||z||||||
 070 9000-
 070a292-
 070b1200-
 331 Journal of neurology für Testzwecke und Validation
 335 official journal of the European Neurological Society-
 341 Zeitschrift für Neurologie-

- RVK** = Regensburger Verbund Klassifikation
- DDC** = Dewey Decimal Classification
- LCC** = Library of Congress Classification
- ...
- PND** = Personen-Normdatei
- SWD** = Schlagwort-Normdatei
- LCSH** = Library of Congress Subject Headings
- ...
- RAK** = Regeln Alphabetischer Katalogisierung
- RSWK** = Regeln für den Schlagwortkatalog
- AACR** = Anglo American Cataloguing Rules
- ...



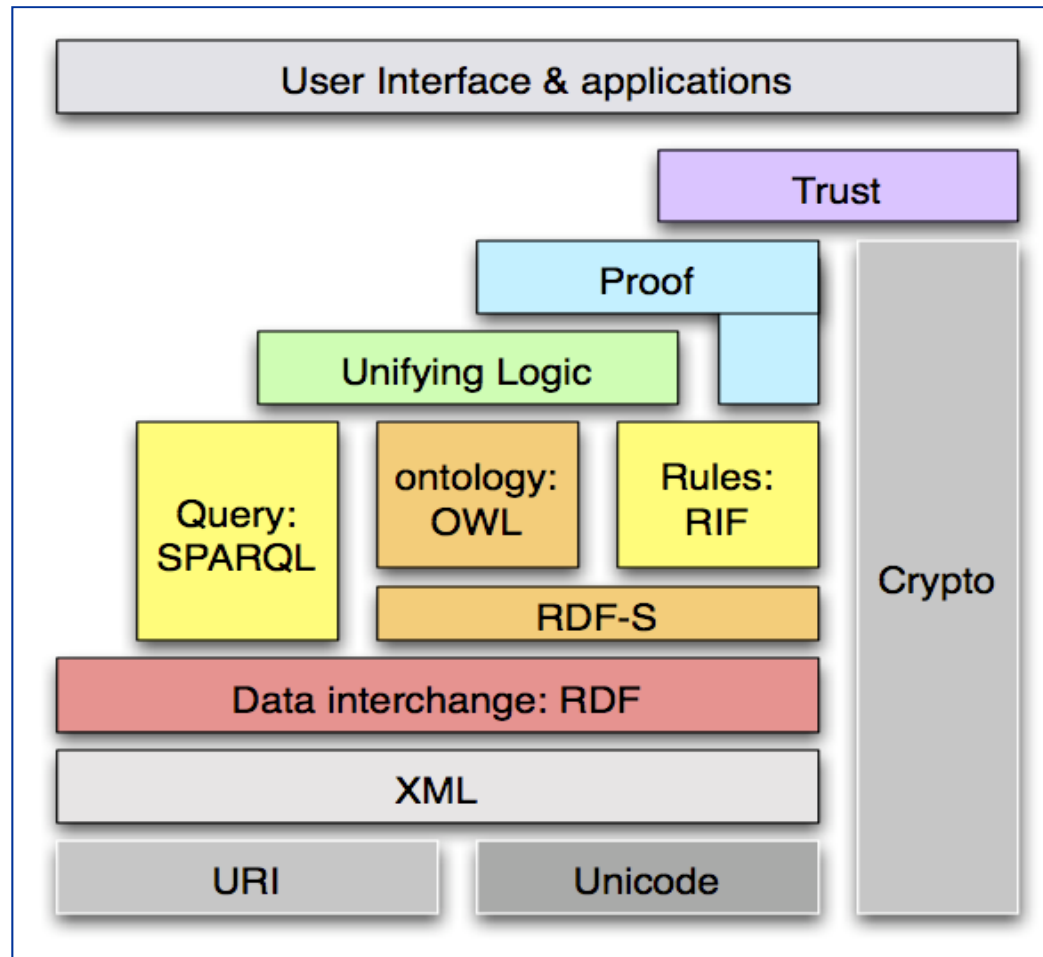
These namespaces are used to describe entities in the number of documents

Namespaces (759 known namespaces)	Number of documents
http://xmlns.com/foaf/0.1/	1,261,218
http://purl.org/oodrelations/v1#	643,606
http://blogs.vandev.nl/schema/foaf/	594,761
http://www.w3.org/2002/07/owl#	449,540
http://files.wissenschaften.uni-berlin.de/siglib/	252,364
http://www.w3.org/2003/01/geo/wgs84_pos#	157,647
http://rdfs.org/sioc/ns#	157,513
http://rdfs.org/sioc/types#	117,529
http://semantic-med3.owl.org/sword1_0/	64,280
http://augusta.org/infex/other/SpecialURI...	25,867
http://xmlns.com/foaf/spec/	24,817
http://usefulinc.com/ns/foap#	19,117
http://purl.org/dc/dcmitype/	17,766
http://www.radameetworks.com/shazam#	8,061
http://www.w3.org/2001/XMLSchema#URIProperty	6,967
http://www.radameetworks.com/2007/09/12/	6,113
http://www.geonames.org/ontology#	5,471
http://purl.org/rss/1.0/	4,566
http://smw.ontoware.org/2005/smw#	4,459
http://www.w3.org/1999/02/22-rdf-syntax-ns#	4,301



- Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro



Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

... aus Nutzersicht

Demo: Informationsverknüpfung auf Basis von HTML-Hyperlinks

Tim Berners-Lee ist der Gründer des W3C

Tim Berners-Lee gründete das W3C im Jahr 1994

Tim Berners-Lee ist Autor des Artikels "Tabulator: Exploring and Analyzing Linked Data on the Semantic Web"

Der Artikel "Tabulator: Exploring and Analyzing Linked Data on the Semantic Web" stammt von Tim Berners-Lee

Tim Berners-Lee kennt Dan Brickley

... aus Rechnersicht:

`Tim Berners-Lee`

???

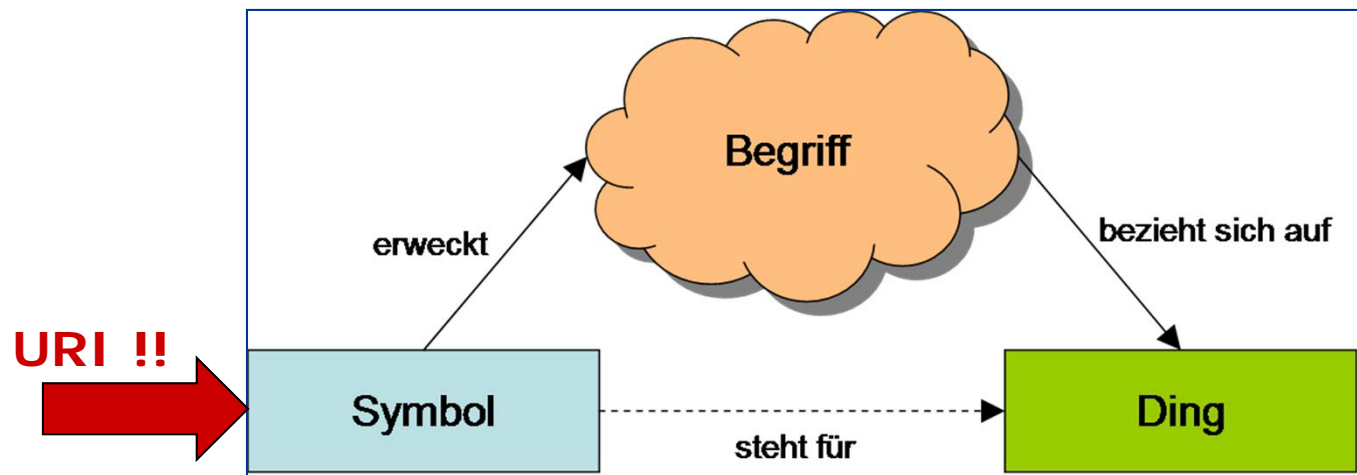
`Dan Brickley`

Idee: aus Zeichenketten „eindeutig identifizierbares Vokabular“ machen

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

Herausforderung: „Semiotisches Dreieck“



Lösung: Eindeutige Festlegung von **Vokabular** auf Basis von URIs

SW: Everything has a URI

Don't say "colour" say `<http://example.com/2002/std6#col>`

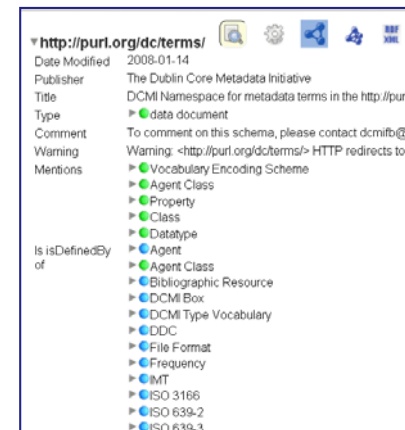
Exkurs: „Semantic Web/Linked Data“ - Intro

aus:

`<TimBL> ist Autor von <Tabulator:...>`

wird:

`<TimBL> <http://purl.org/dc/terms/creator> <Tabulator...>`



<http://purl.org/dc/terms/>

aus:

`<TimBL> kennt <DanBrickley>`

wird:

`<TimBL> <http://xmlns.com/foaf/0.1/knows> <DanBrickley>`



<http://xmlns.com/foaf/0.1/>

■ Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

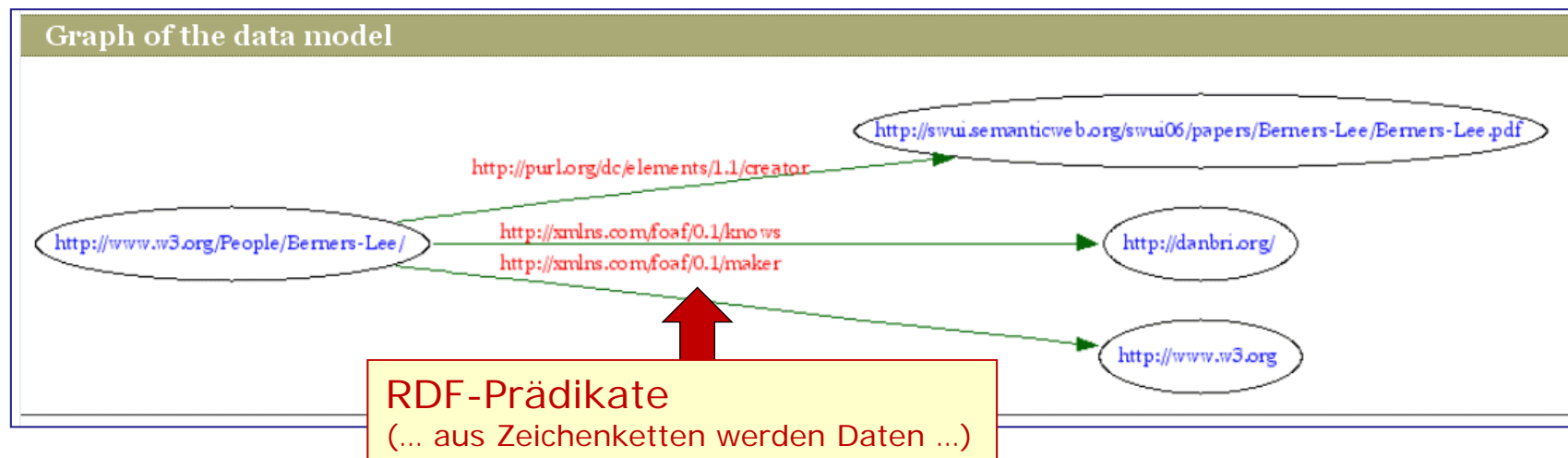
RDF - Resource Description Framework: Datenmodell

Elementare Einheit des RDF-Datenmodells ist ein „**Statement**“ in Form eines Tripels
„**Subjekt-Prädikat-Objekt**“
Subjekt und Prädikat MÜSSEN, Objekt KANN durch einen **URI** repräsentiert werden.
Jedes Objekt eines Statements KANN Subjekt eines neuen Statements sein.

RDF-Statement in „Notation 3“

<TimBL>	< http://xmlns.com/foaf/0.1/knows >	< http://danbri.org >
(„Subjekt“)	(„Prädikat“/„Property“)	(„Objekt“)

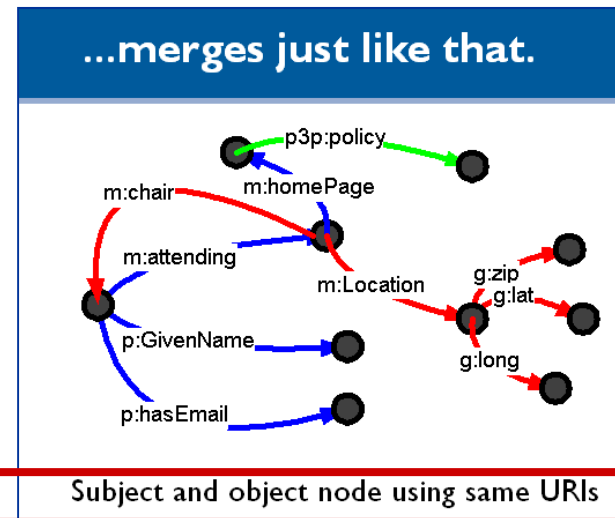
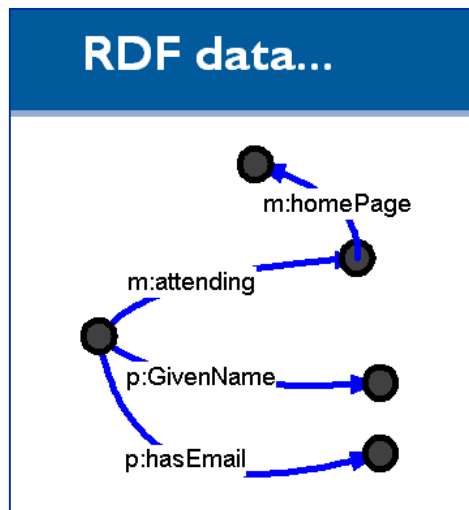
RDF-Statement als gerichteter Graph



Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

Das „Objekt“ jedes Statements kann „Subjekt“ eines neuen Statements sein



(Quelle: Tim Berners-Lee: „[Semantic Web: Building On What Exists](#)“, 2006)

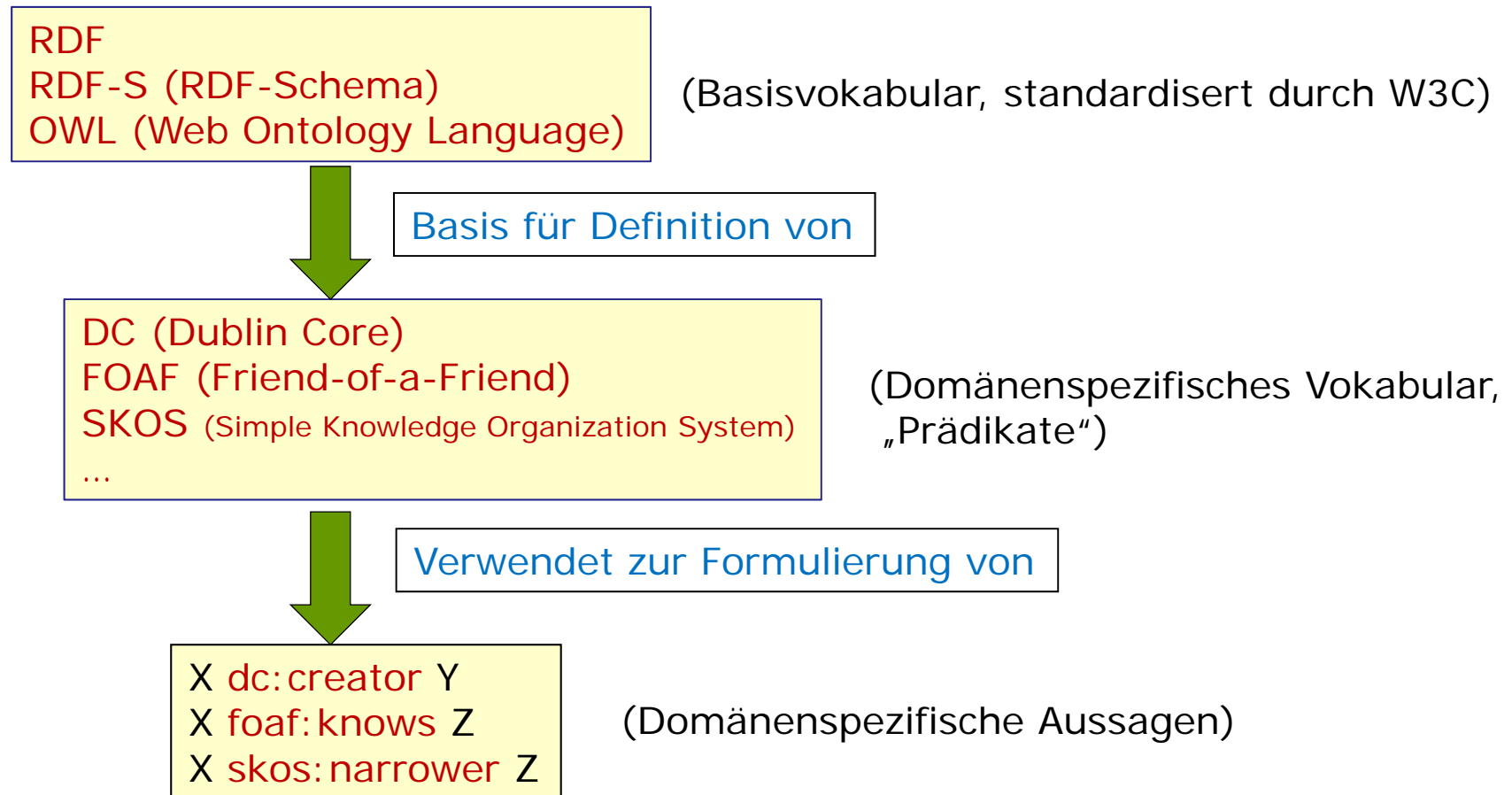


„Linked Data“

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

Formulierung domänenspezifischer Aussagen



Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

RDF (25)	RDFS (16)	OWL (Auswahl)
rdf: RDF	rdfs: Class	owl: AllDifferent
rdf: about	rdfs: Resource	owl: Class
rdf: Description	rdfs: Datatype	owl: FunctionalProperty
rdf: Property	rdfs: domain	owl: InverseFunctionalProperty
rdf: type	rdfs: range	owl: Restriction
rdf: resource	rdfs: Container	owl: SymmetricProperty
rdf: Statement	rdfs: isDefinedBy	owl: Thing
rdf: subject	rdfs: equivalentClass	owl: TransitiveProperty
rdf: predicate	rdfs: subClassOf	owl: allValuesFrom
rdf: object	rdfs: subPropertyOf	owl: cardinality
rdf: value	rdfs: seeAlso	owl: complementOf
rdf: Bag	rdfs: ContainerMembershipProperty	owl: disjointWith
rdf: Seq	rdfs: member	owl: equivalentClass
rdf: List	rdfs: comment	owl: equivalentProperty
rdf: Alt	rdfs: label	owl: hasValue
rdf: XMLLiteral	rdfs: Literal	owl: intersectionOf
rdf: ID , rdf: nodeID		owl: inverseOf
rdf: parseType		owl: sameAs
rdf: datatype		owl: unionOf

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

SKOS: Transfer klassischer, kontrollierter Vokabulare ins SW

Namensraum-URI: " <http://www.w3.org/2004/02/skos/core#> "

Classes:

[Collection](#)
[Concept](#)
[ConceptScheme](#)
[OrderedCollection](#)

Properties:

[altLabel](#)
[broadMatch](#)
[broader](#)
[broaderTransitive](#)
[changeNote](#)
[closeMatch](#)
[definition](#)
[editorialNote](#)
[exactMatch](#)
[example](#)
[hasTopConcept](#)
[hiddenLabel](#)
[historyNote](#)
[inScheme](#)

Properties (cont.)

[mappingRelation](#)
[member](#)
[memberList](#)
[narrowMatch](#)
[narrower](#)
[narrowerTransitive](#)
[notation](#)
[note](#)
[prefLabel](#)
[related](#)
[relatedMatch](#)
[scopeNote](#)
[semanticRelation](#)
[topConceptOf](#)



<http://www.w3.org/TR/skos-reference/>

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

RDF-Vokabulare, Namensräume in der Linked Data Cloud

These namespaces are used to describe entities in X number of documents

Namespaces (814 know namespaces)	Number of documents
http://xmlns.com/foaf/0.1/	1, 261, 382
http://purl.org/goodrelations/v1#	645, 201
http://blogs.yandex.ru/schema/foaf/	584, 781
http://www.w3.org/2002/07/owl#	451, 101
http://sites.wiwiss.fu-berlin.de/suhl/bi...	252, 361
http://rdfs.org/sioc/ns#	183, 227
http://www.w3.org/2003/01/geo/wgs84_pos#	157, 659
http://rdfs.org/sioc/types#	117, 599
http://semantic-mediawiki.org/swift/1.0#	64, 281
http://xmlns.com/foaf/spec/	50, 514
http://auswiki.org/index.php/Special:URI...	25, 887
http://usefulinc.com/ns/doap#	19, 127
http://purl.org/dc/dcmitype/	17, 766
http://www.radarnetworks.com/shazam#	8, 061
http://wecowi.org/view/Spezial:URIResolv...	6, 967
http://www.radarnetworks.com/2007/09/12/...	6, 113
http://www.geonames.org/ontology#	5, 474
http://purl.org/rss/1.0/	4, 606
http://smw.ontoware.org/2005/smw#	4, 459
http://www.w3.org/1999/02/22-rdf-syntax-...	4, 346

„Nutzung von RDF-Vokabularen“

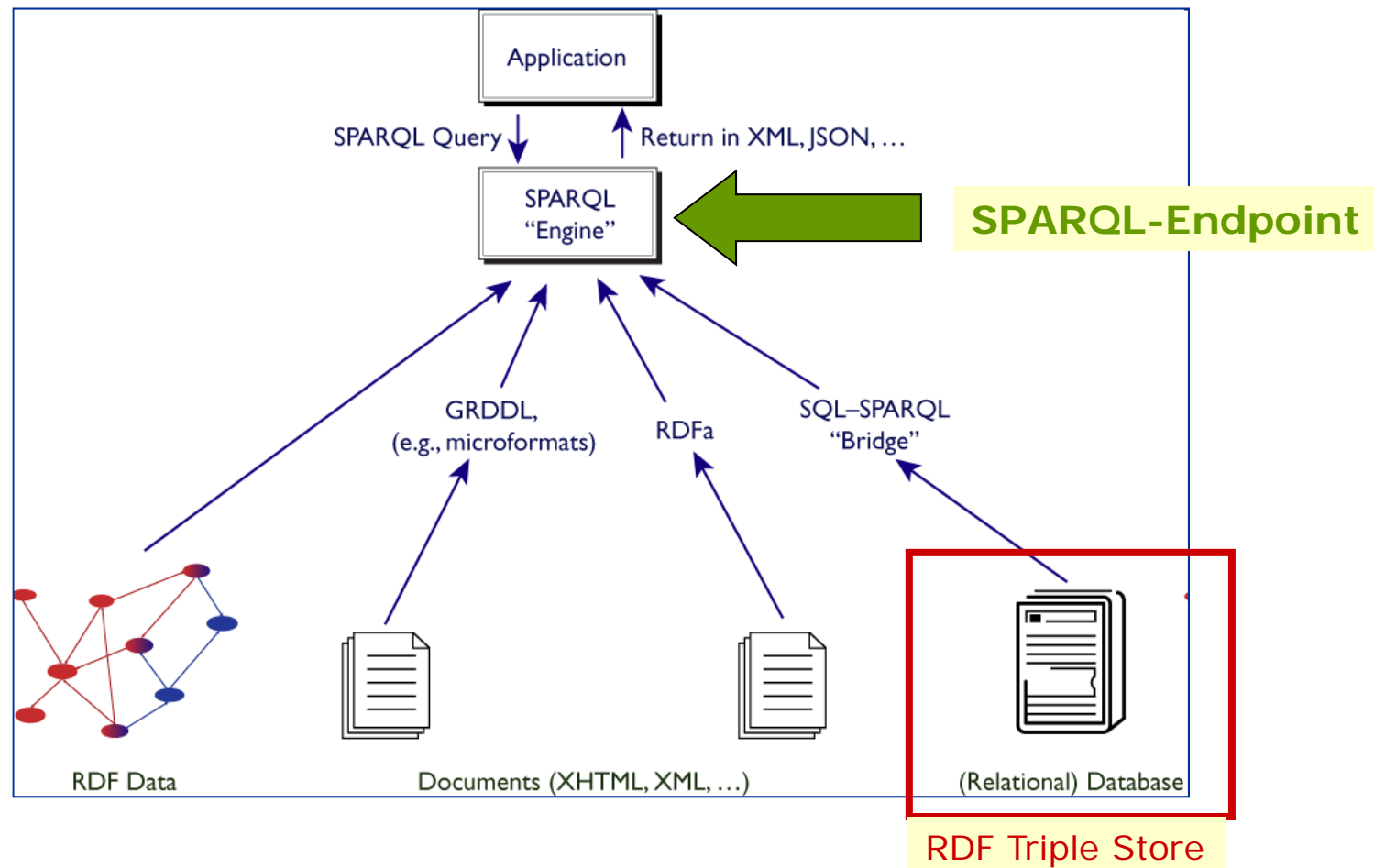
```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:openid="http://xmlns.openid.net/auth#"
  xmlns:bio="http://vocab.org/bio/0.1/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:wot="http://xmlns.com/wot/0.1/"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:dc="http://purl.org/dc/terms/"
  xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xhv="http://www.w3.org/1999/xhtml/vocab#"
  xmlns:rel="http://vocab.org/relationship/"
  xmlns:sioc="http://rdfs.org/sioc/ns#"
  xmlns:air="http://www.daml.org/2001/10/html/airport-ont#"
  xmlns:geo="http://www.w3.org/2003/01/geo/wgs84_pos#"
  xmlns:rss="http://web.resource.org/rss/1.0/"
>
  <foaf:PersonalProfileDocument
    rdf:about="http://www.ivan-herman.net/foaf.rdf">
    <foaf:primaryTopic>
      <dc:Agent rdf:about="http://www.ivan-herman.net/me">
        <bio:event>
          <bio:Birth>
            <bio:date
```

<http://pingthesemanticweb.com/stats/namespaces.php>

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Exkurs: „Semantic Web/Linked Data“ - Intro

SPARQL : Abfrage von RDF-Daten



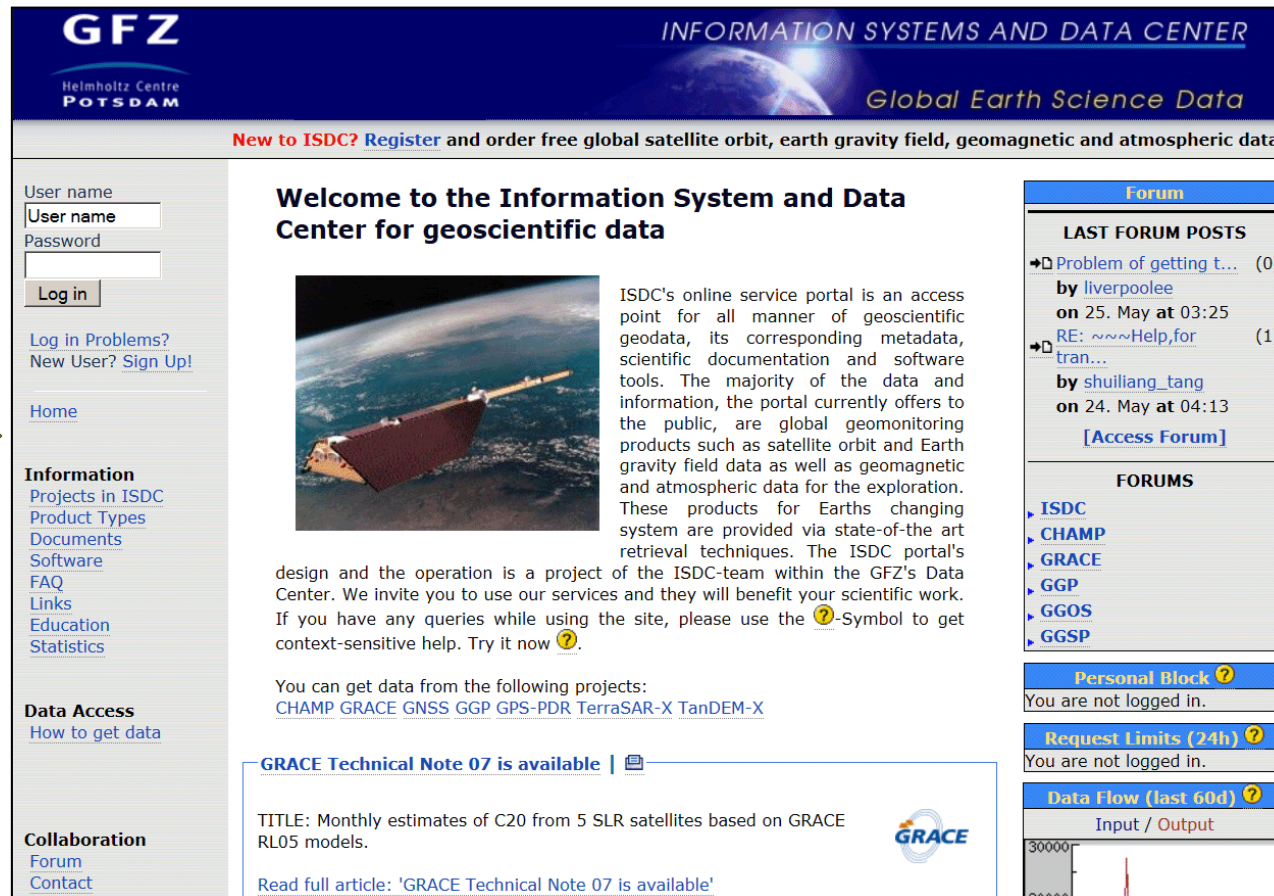
■ Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Semantische Verknüpfung von Forschungsdaten

Use Case I SDC

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Use Case ISDC: Information System and Data Center am GFZ

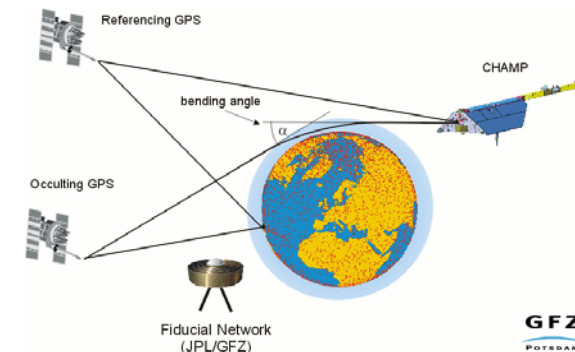
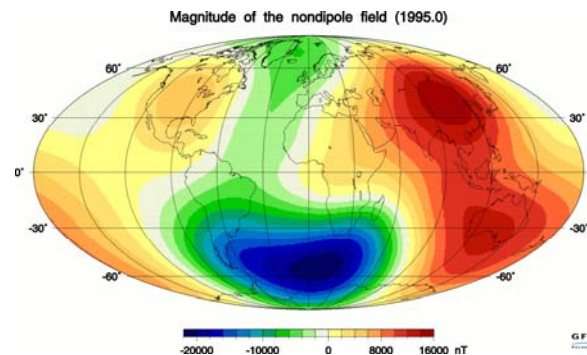
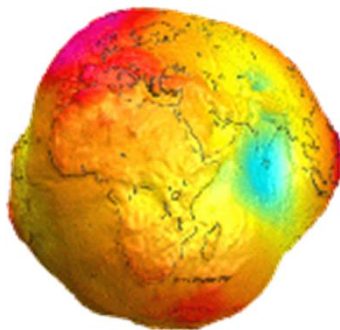
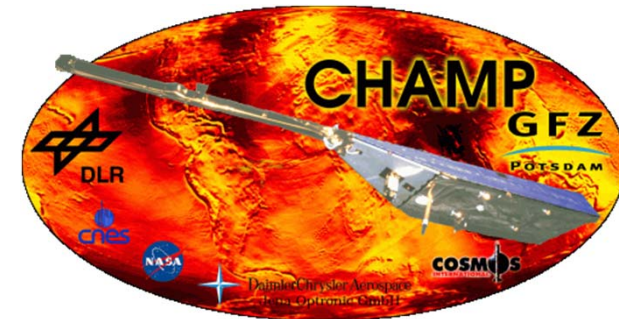


The screenshot shows the homepage of the Information System and Data Center (ISDC) at GFZ. The header includes the GFZ logo (Helmholtz Centre Potsdam) and the text 'INFORMATION SYSTEMS AND DATA CENTER' and 'Global Earth Science Data'. A navigation bar contains a link for 'New to ISDC? Register' and a promotional message about free global satellite orbit, earth gravity field, geomagnetic, and atmospheric data. The main content area features a 'Welcome to the Information System and Data Center for geoscientific data' section with an image of a satellite in orbit and a detailed description of the portal's services. A sidebar on the left provides navigation options like 'Log in', 'Home', 'Information', 'Data Access', and 'Collaboration'. A right sidebar contains a 'Forum' section with 'LAST FORUM POSTS' and 'FORUMS' lists, and status indicators for 'Personal Block' and 'Request Limits'. A 'Data Flow' graph is also visible at the bottom right.

<http://isdc.gfz-potsdam.de/>

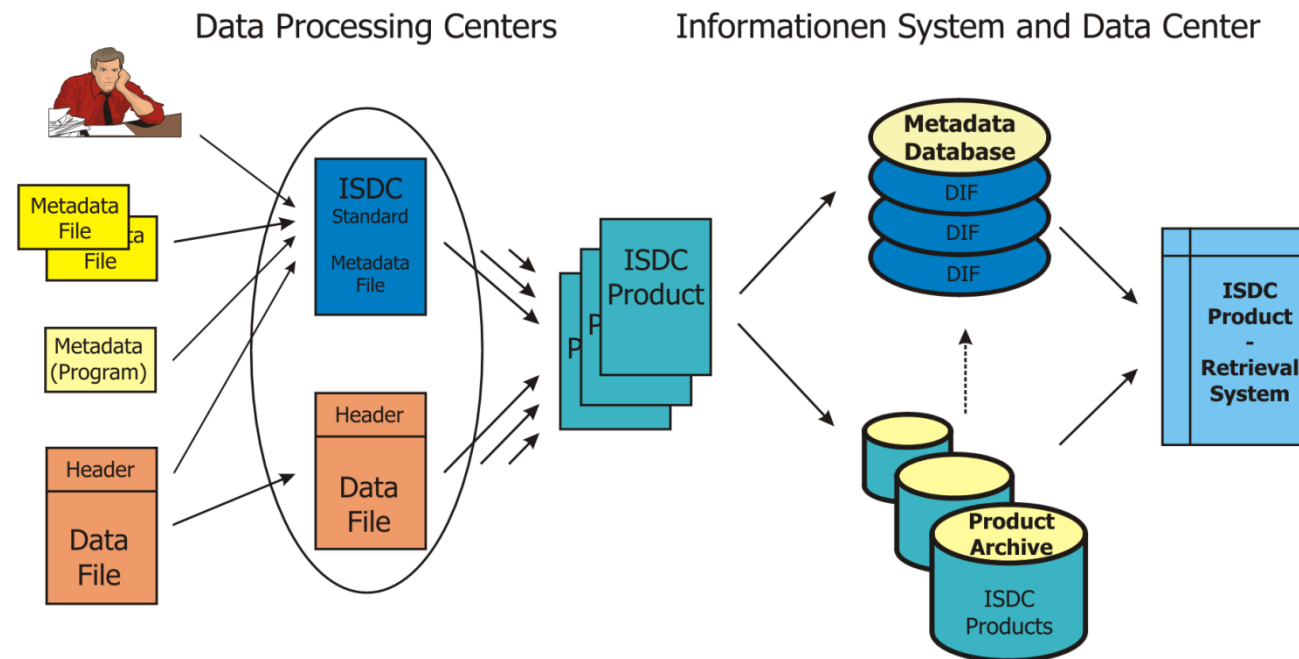
Use Case ISDC: Forschungsdaten, Datenprodukte, ...

- Partners: GFZ, DLR, NASA, CNES
- Lifetime: 04/2000 – 09/2010
- Main objectives:
 - Global Earth **gravity field** (models)
 - Global Earth **magnetic field** (models)
 - **Atmospheric/Ionospheric sounding** (GPS occultations)
 - Monitoring surface heights (GPS altimetry experiment)



Use Case ISDC: Forschungsdaten, Metadaten, XML

ISDC Product Philosophy and Metadata Processing



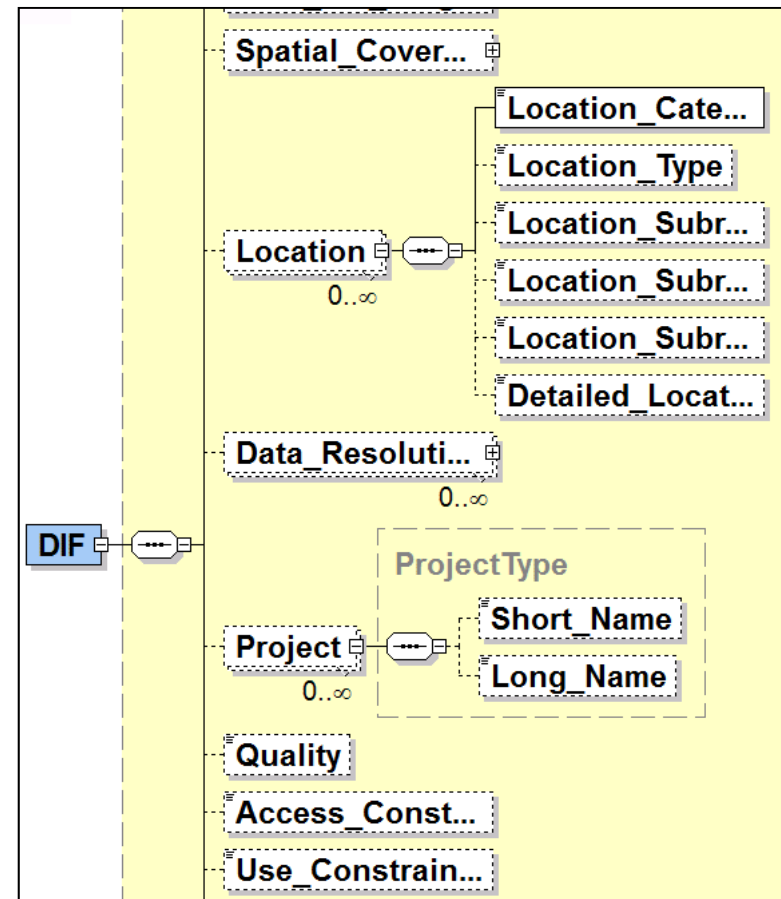
ISDC Standard for Product Type and Data Product Metadata:
XML Product Type DIF* (V. 9) + XML Data Product DIF(s)**

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Use Case ISDC: Forschungsdaten, Metadaten, XML

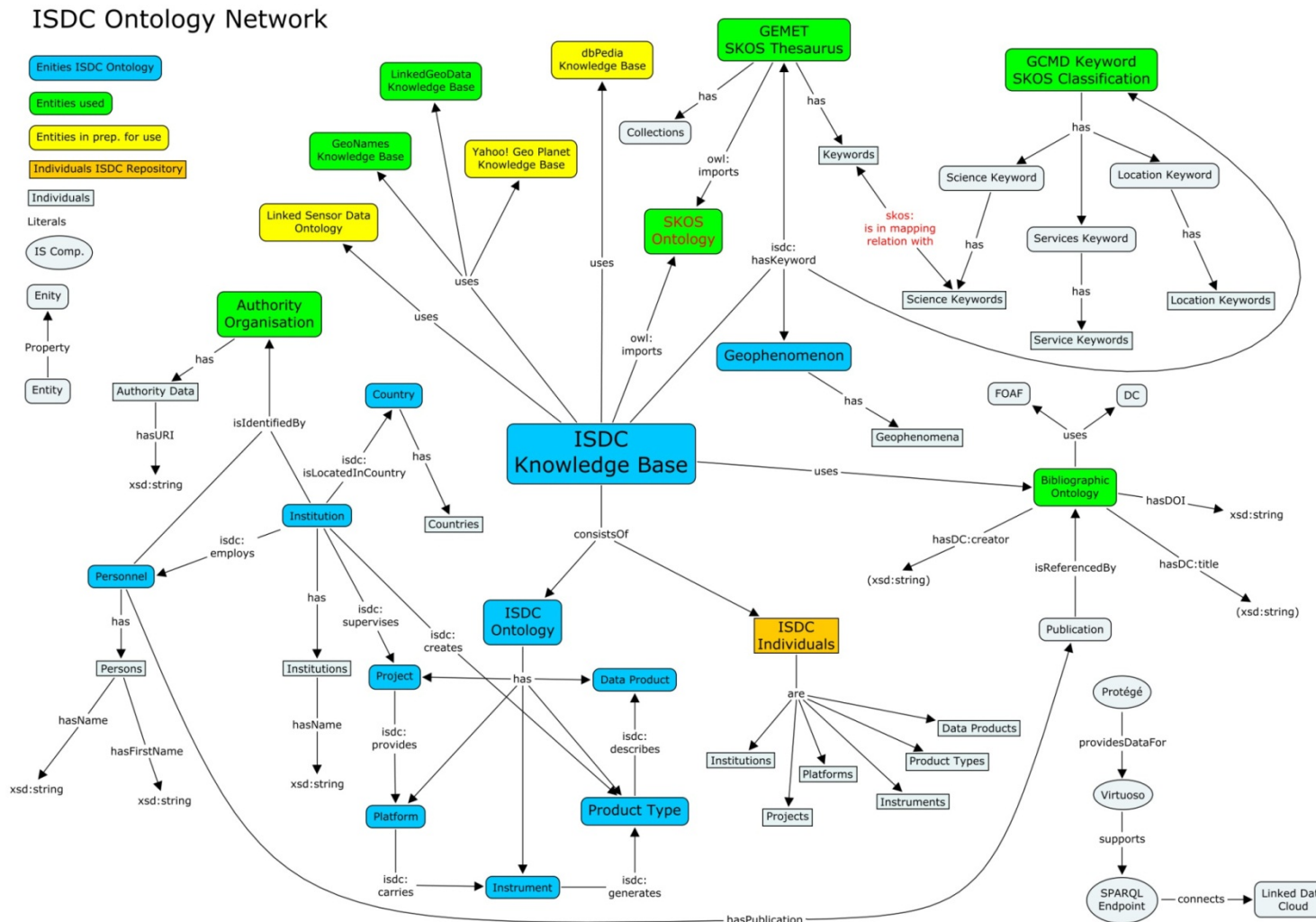
Ausgangspunkt:

Metadatenstandard **DIF**
(**D**irectory **I**nterchange **F**ormat)



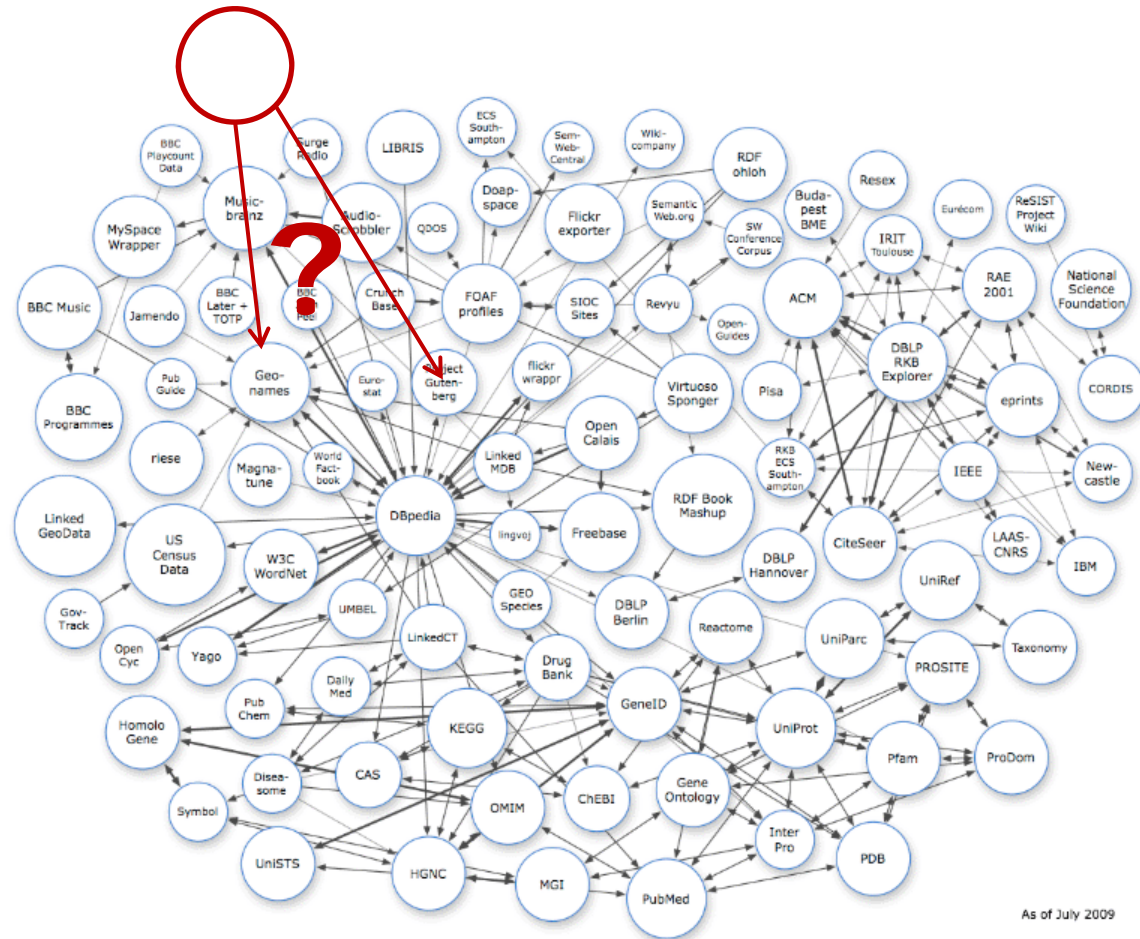
<http://gcmd.nasa.gov/User/difguide/difman.html>

Use Case ISDC: Entwurf ISDC Ontology Network



Use Case ISDC: Zielsetzung

Semantisch präzise Verknüpfung mit relevanten Kontextinformationen in der Linked Data Cloud



As of July 2009

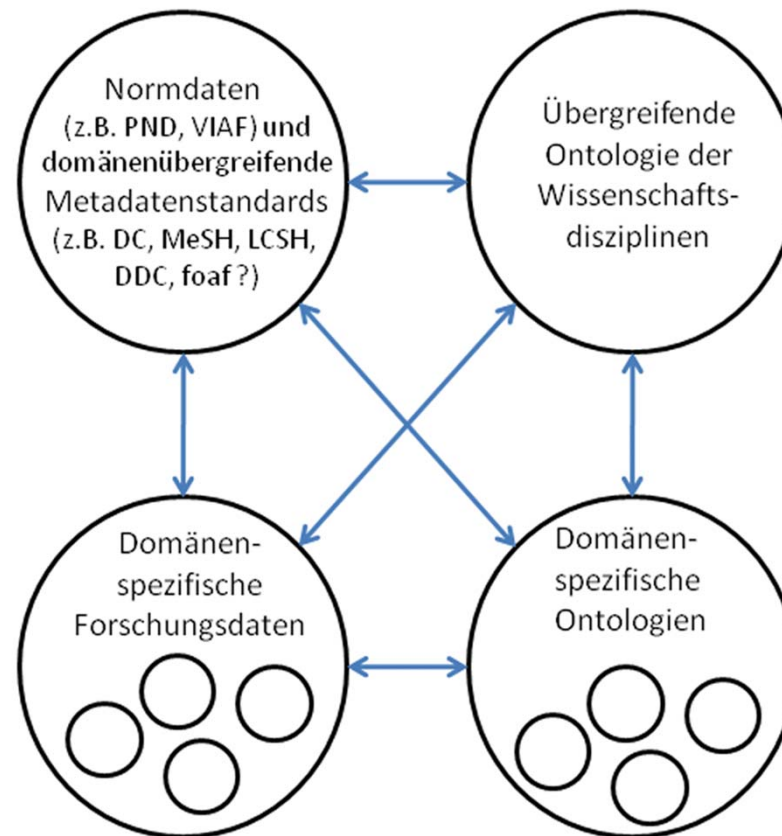
Linked Data für Forschungsdaten: Herausforderungen

Vertrauenswürdigkeit
(Semantische) Präzision
Stabilität (URIs)

...

Linked Data für Forschungsdaten: Herausforderungen

Spezifische Komponenten erforderlich für „Linked Research Data“ ?




Linked Data für Forschungsdaten: Herausforderungen



Aspekt: Identifiers - Bedeutung von Authorities

Suchen

Feld auswählen: Index auswählen:

Suchbegriffe:

Neher, Günther 
VIAF ID: 37628791 (Personal)
Permalink: <http://viaf.org/viaf/37628791>

  100 1 [_ja Neher, Günther](#)

<http://viaf.org>

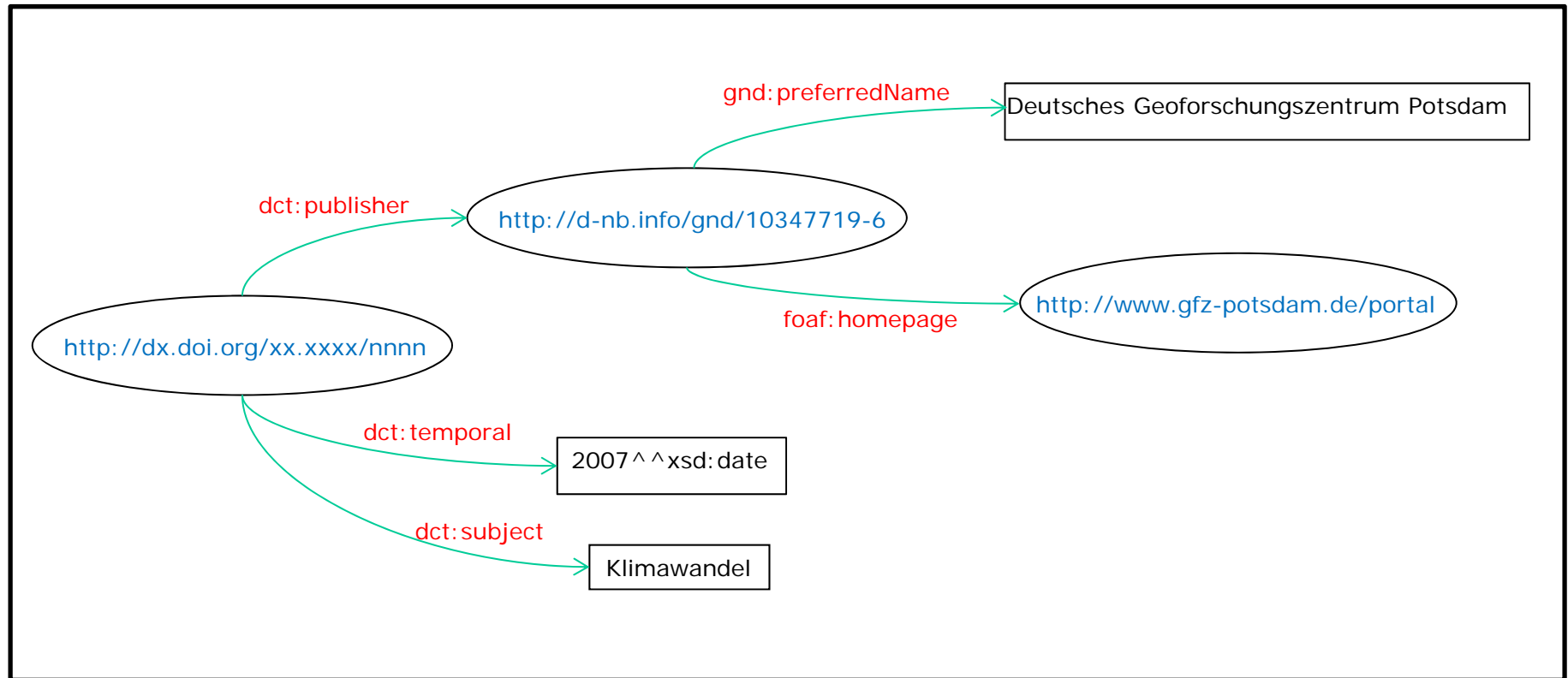
```
-<rdf:RDF xml:base="http://viaf.org/">
- <rdf:Description rdf:about="http://viaf.org/viaf/37628791">
  -<!--
    The primary entity (formerly a viaf:NameAuthorityCluster)
  -->
  <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
  <rdf:type rdf:resource="http://rdvocab.info/uri/schema/FRBRentitiesRDA/Person"/>
  <foaf:name>Neher, Günther</foaf:name>
  <owl:sameAs rdf:resource="http://d-nb.info/gnd/114798192"/>
</rdf:Description>
- <!--
  Controlled and alternate headings from various authority files
-->
- <skos:Concept rdf:about="http://viaf.org/viaf/sourceID/DNB%7C114798192#skos:Concept">
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/DNB"/>
  <skos:prefLabel>Neher, Günther</skos:prefLabel>
  <foaf:focus rdf:resource="http://viaf.org/viaf/37628791"/>
</skos:Concept>
<!--Deprecated identifiers-->
- <rdf:Description rdf:about="http://viaf.org/viaf/37628791/#skos:Concept">
  - <rdfs:comment>
    This concept identifier is deprecated. See the foaf:focus link for the primary entity.
  </rdfs:comment>
  <foaf:focus rdf:resource="http://viaf.org/viaf/37628791"/>
</rdf:Description>
- <rdf:Description rdf:about="http://viaf.org/viaf/37628791/#foaf:Person">
  - <rdfs:comment>
    This primary entity identifier is deprecated. See the owl:sameAs link for the new identifier.
```

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Linked Data für Forschungsdaten: Herausforderungen

Aspekt: Identifiers - Bedeutung von Authorities

Beispiel:




Linked Data für Forschungsdaten: Herausforderungen

Aspekte: Vokabular-Mappings, Domänen-Modelle, Vertrauenswürdigkeit

=> relevante W3C-Aktivitäten: Beispiele

- Library Linked Data
- Semantic Sensor Network
- Provenance
- VoID – Vocabulary of Interlinked Datasets (<http://www.w3.org/TR/void/>)

 W3C[®] Incubator Group Report

Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets

W3C Incubator Group Report 25 October 2011

This Version:
<http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025/>

Latest Published Version:
<http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset/>

Authors
Antoine Isaac, Europeana and Vrije Universiteit Amsterdam, Netherlands
William Waites, University of Edinburgh (School of Informatics), UK
Jeff Young, OCLC Online Computer Library Center, Inc., US
Marcia Zeng, Kent State University, US (W3C Invited Expert)

See also [translations](#).

<http://www.w3.org/2005/Incubator/lld/>

[Incubator Activity](#) > W3C
Semantic Sensor Network
Incubator Group

This XG started on 4 March 2009 and closed on 28 June 2011. See the [Final Report](#).

The mission of the Semantic Sensor Network Incubator Group, part of the [Incubator Activity](#), is to begin the formal process of producing ontologies that define the capabilities of sensors and sensor networks, and to develop semantic annotations of a key language used by services based sensor networks.

See the [charter](#) for more information.

- [News](#)
- [Deliverables](#)
- [Meetings](#)
- [Minutes](#)
- [About the Semantic Sensor Network XG](#)

<http://www.w3.org/2005/Incubator/ssn/>

W3C Provenance Incubator Group Wiki
(Redirected from [Main Page](#))

Welcome to the Provenance ^{Main Page} Incubator Group Wiki.

Contents [\[hide\]](#)

- [1 Mission and Charter](#)
- [2 Final Report](#)
- [3 About Provenance](#)
- [4 Group Reports](#)
- [5 Timeline of Activities](#)
- [6 Meetings and Discussions](#)
- [7 Twitter and social media](#)
- [7.1 Liaisons with Other Groups](#)
- [8 Contact](#)

Mission and Charter

The **mission** of the Provenance Incubator Group, part of the [Incubator Activity](#), was to provide a state-of-the art understanding and develop a roadmap in the area of provenance for Semantic Web technologies, development, and possible standardization. See

<http://www.w3.org/2005/Incubator/prov/>

■ Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Eine Aufgabe für
Information Professionals ?

Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

Eine Aufgabe/Perspektive für Information Professionals ?

Ref: American Geophysical Union, AGU-Fall Meeting 2011 Session ED22:

"Distributing Science Data for Re-Use"

Description:

*This session will focus on services to **distribute scientific research data** so that it can be used by educators, the general public or scientists from other disciplines for uses other than its original intended purpose. Submissions that address underlying technology used to allow interested parties to browse and interact with the data, **standards for documenting data, policies on data sharing**, or that show benefits to **scientific data re-use** are encouraged, as are any submissions that discuss remaining barriers to re-use and ways to overcome them.*
<http://fallmeeting.agu.org/2011/scientific-program/session-search/241>

Ref: American Geophysical Union, AGU-Fall Meeting 2011 Session IN29:

"Semantic Solutions and Linked Data for Science"

Description:

*As science problems become increasingly interdisciplinary, science solutions need to **integrate data from multiple disciplines**. Common language and terminology usage varies across disciplines inhibiting meaningful sharing. One approach that is gaining traction is leveraging different levels of **encodings of meanings of terms**. Approaches range from 'lightweight' linking as seen in many **linked data** approaches to **deeper integration** as seen in many **ontology-supported approaches**. This session seeks to bring together a range of research approaches leveraging all types of semantic solutions. We invite submissions describing ontologies, infrastructure, tools and methods to support semantically-enhanced approaches to science problems.*

<http://fallmeeting.agu.org/2011/scientific-program/session-search/542>

Ref: American Geophysical Union, AGU-Fall Meeting 2011 Session IN10:

"Cross-Cutting Interoperability: Enabling Something New"

Description:

*With Earth science research moving onto collaborative, distributed, and interdisciplinary paradigms, interoperability has become key to enabling progress. Interoperability issues are no longer isolated topics. Increasing complexity in data systems have presented inter-related interoperability issues. The increasing use of observational and model data in conjunction with in-situ data requires more interoperability. **Data discovery, access, workflows, processing, and collaboration** also require cohesively **addressing multiple areas of interoperability**. Progress has been made yet we are still far from achieving true interdisciplinary interoperability. This session seeks experiences on addressing cross-cutting interoperability issues.*

<http://fallmeeting.agu.org/2011/scientific-program/session-search/523>



JA !

- Metadaten für Forschungsdaten - eine Aufgabe für Information Professionals!?

</ENDE>

Danke für Ihre Aufmerksamkeit