

# Inhalt

- Einführung in Metadaten
- Was sind eigentlich Metadaten?
- Beispiele aus der Geo-Praxis
- Syntax und Semantik
- Ontologien
- Thesauri und Indexierung
- Technik: Hierarchische Kataloge
- Divergenz zwischen Angebot und Nachfrage
- Neue Anforderungen
- Digitale Fundgruben
- Technik: REST mit Sinn verbinden
- Ausblick

Beschreibung/Identität   Status/Weitergabe   Geometriedaten   Sachdaten   Dokumentation

**Fachthema:** Verwaltungsgrenzen  
**Thema:** Forstamtsgrenzen

**Datensatz:** forst\foa\_gr.shp

**Kurzbeschreibung:** Abgrenzung der Forstamtsflächen mit Stand der Begehung

**Nutzungsmöglichkeit:** enthält Telefonnummern und Emailadressen; Arbeitsgrundlage für die Forstämter; Information für Gemeinden

**Verarbeitungsprozeß:**

Meine Geschäftsdaten

- Kennzahlen
  - Aufträge
  - Verkäufe
  - Lager
  - Profitmarge
- Dimensionen
  - Geschäftseinheit
    - Deutschland
    - USA
    - Brasilien
  - Produkte
    - Software
    - Hardware
  - Periode
    - 2009
    - 2010
    - 2011
    - Q1
    - Q2

# Einführung in Metadaten



- Als Metadaten oder Metainformationen bezeichnet man allgemein Daten, die Informationen über andere Daten enthalten.
- [...] Eine allgemeingültige Unterscheidung zwischen Metadaten und gewöhnlichen Daten **existiert allerdings nicht**, da die Bezeichnung eine Frage des Standpunkts ist!

# Metadaten in INSPIRE



*Der Begriff "Metadaten" an sich wird nicht definiert, sondern lediglich Inhalte beschrieben:*

## RICHTLINIE 2007/2/EG DES EUROPÄISCHEN PARLAMENTS UND DES RATES

- Metadaten umfassen Angaben zu folgenden Aspekten:
  - a) Entsprechung der Geodatensätze mit den in Artikel 7 Absatz 1 vorgesehenen Durchführungsbestimmungen;

## Artikel 7 Absatz 1

*(1) Durchführungsbestimmungen, mit denen technische Modalitäten für die Interoperabilität und, wenn durchführbar, die Harmonisierung von Geodatenätzen und -diensten festgelegt werden und die eine Änderung dieser Richtlinie durch Hinzufügung neuer nicht wesentlicher Bestimmungen bewirken, sind gemäß dem in Artikel 22 Absatz 3 genannten Regelungsverfahren mit Kontrolle zu erlassen. Bei der Ausarbeitung der Durchführungsbestimmungen sind die einschlägigen Nutzeranforderungen, bestehende Initiativen und die internationalen Normen zur Harmonisierung von Geodatenätzen sowie Durchführbarkeits- und Kosten-Nutzen-Erwägungen zu berücksichtigen. Einschlägige Normen, die von Organisationen des Völkerrechts festgelegt worden sind, um die Interoperabilität oder Harmonisierung von Geodatenätzen und -diensten sicherzustellen, werden in die in diesem Absatz genannten Durchführungsbestimmungen einbezogen, und gegebenenfalls werden dort die bestehenden technischen Mittel angegeben.*

# Metadaten in INSPIRE



- b) Bedingungen für den Zugang zu Geodatenständen und -diensten und deren Nutzung sowie gegebenenfalls entsprechende Gebühren;
- c) Qualität und Gültigkeit der Geodatenstände;
- d) für die Schaffung, Verwaltung, Erhaltung und Verbreitung von Geodatenständen und -diensten zuständige Behörden;
- e) Beschränkungen des Zugangs der Öffentlichkeit gemäß Artikel 13 sowie die Gründe für solche Beschränkungen.

# Was sind Metadaten eigentlich?



Typische Metadaten zu einem Buch (Objekt) sind beispielsweise:

- der Name des Autors
- die Auflage
- das Erscheinungsjahr
- der Verlag
- die ISBN ...

# ...digitale Metadaten?



Zu den Metadaten einer digitalen Datei zählen unter anderem:

- der Dateiname
- die Zugriffsrechte
- das Datum der letzten Änderung
- das Dateiformat (odt, shp, png, xml, rdf, rss)

→ *Aber: da geht viel mehr, weil alles vernetzt ist!*

# Syntax und Semantik

*Syntax und Semantik sind Disziplinen der Semiotik...*

- Die **Syntax** beschreibt Geodaten rein **formal**
- Während sich die **Semantik** auf deren **Bedeutung** bezieht.



# Beispiele aus der Geo-Praxis

*Metadaten zu einem **Orthophoto** enthalten folgende Informationen:*

- Räumlicher Ausschnitt
- Koordinatensystem
- Projektion
- Format,  
Zugriffsmöglichkeiten
- Datum der Aufnahme
- Auflösung des  
Originalbildes
- Farbkanäle
- Aufnahmegerät
  - Digital
  - Analog
- Bearbeitungsschritte
  - Ausschnitt, Entzerrung
  - Schattenaufhellung
  - Kontrastanpassung
  - Farbanpassung
  - etc.

# Beispiele aus der Geo-Praxis

*Metadaten zu einem **Stau-Informationsdienstes** beschreiben folgende Aspekte:*

- Räumlicher Ausschnitt
- Koordinatensystem
- Projektion
- Format,  
Zugriffsmöglichkeiten
- Ursprung der  
geometrischen Grundlage
- Aktualität der geometrischen  
Grundlage
- Aufnahmeart
  - Datenerhebung durch  
Verkehrsüberwachung  
(amtlich)
  - Meldungen durch Autofahrer  
(freiwillig, verifiziert)
  - Prognosestatus
- Berücksichtigung von Baustellen
- etc.

# Pragmatik

*Syntax und Semantik sind Disziplinen der Semiotik...*

Die **Syntax** beschreibt Geodaten rein **formal**, während sich die **Semantik** auf deren **Bedeutung** bezieht.

*Syntax und Semantik werden durch die Lehre von den Beziehungen der Zeichen zu den Zeichenbenutzern (der **Pragmatik**) vollendet.*

Uns fehlen **Beziehungen!**

# Thesauri und Indexierung

Ein Thesaurus oder Wortnetz ist ein kontrolliertes Vokabular, dessen Begriffe durch Relationen (***Beziehungen***) miteinander verbunden ist.

- systematisch geordneten Sammlung von Begriffen, die in thematischer Beziehung zueinander stehen.
- Polyhierarchische Indexierung (→ ***für Kataloge***)

Eine gemeinsame **Bedeutung** wird vorausgesetzt.

# Thesauri und Indexierung

Ein Thesaurus oder Wortnetz ist ein kontrolliertes Vokabular, dessen Begriffe durch Relationen (*Beziehungen*) miteinander verbunden ist.

- systematisch geordneten Sammlung von Begriffen, die in **thematischer Beziehung** zueinander stehen.
- Polyhierarchische Indexierung (→ *für Kataloge*)

Eine gemeinsame **Bedeutung** wird vorausgesetzt.



# Ontologien

Die Ontologie ist die Lehre vom ***Sein der Dinge***.

- Eine Ontologie beschreibt die **Beziehungen** von Sinnzusammenhängen (***Bedeutungen***) und bringt sie in einen definierten und kontrollierten Kontext.

# Ontologien

Die Ontologie ist die Lehre vom ***Sein der Dinge***.

- Eine Ontologie beschreibt die Beziehungen von Sinnzusammenhängen (***Bedeutungen***) und bringt sie in einen definierten und kontrollierten Kontext.
- Akzeptieren und Anwendung gemeinsamer Ontologie oder eines Sinnzusammenhangs ermöglicht Kommunikation.

# Ontologien

Die Ontologie ist die Lehre vom ***Sein der Dinge***.

- Eine Ontologie beschreibt die Beziehungen von Sinnzusammenhängen (***Bedeutungen***) und bringt sie in einen definierten und kontrollierten Kontext.
- Akzeptieren und Anwendung gemeinsamer Ontologie oder eines Sinnzusammenhangs ermöglicht Kommunikation.



# Komplexe Technologie und Formate

## Die Situation auf Seite des Anbieters:

- Geodaten werden durch den kleinsten gemeinsamen Nenner beschrieben (Dublin Core, FGDC, ISO 19115).
- Metadatenerfassung und Pflege sind zusätzliche, (oft unbeliebte) Aufgaben der Datenanbieter.
- Bereitstellung über komplexe Mechanismen (CSW) und Formate (ISO 19139 und spezielle Profile).
- Metadaten-Qualität ist weitgehend auf Syntax beschränkt.
- Interaktionsmöglichkeiten mit den Anwendern sind minimal.

# Technik: Hierarchische Kataloge

## • Dublin Core

## • OGC CSW 2.0.2

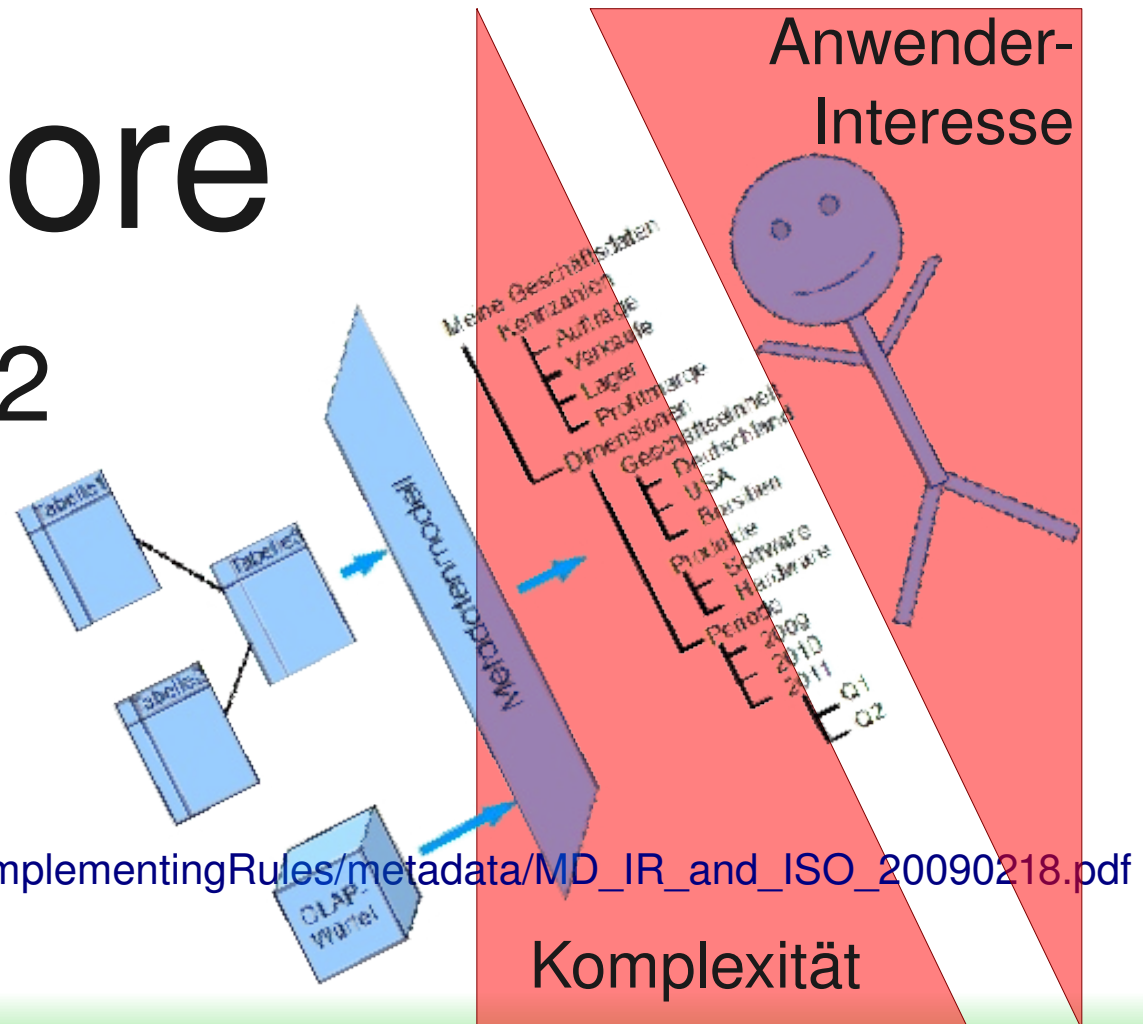
## • ISO 19115

## • ISO 19119

## • ISO 19139

• [http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/metadata/MD\\_IR\\_and\\_ISO\\_20090218.pdf](http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/metadata/MD_IR_and_ISO_20090218.pdf)

• INSPIRE



# Grenzen aktueller Möglichkeiten

## Suche nach **OGC Capabilities mit Google:**

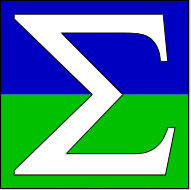
- Liefert unstrukturiert Capabilities von Kartendiensten, der Microsoft XBox, etc.

## **BKG-Portalsuche** nach "administrative Grenzen Deutschland":

- Metadatendienst WSV 9323 Treffer, BKG 0 Treffer

## Suche im **OGC CSW Portal der FAO:**

- Liefert Metadaten zu Geodaten, aber keinen Link oder Möglichkeit diese zu beziehen.



# Divergenz zwischen Angebot und Nachfrage

## Die Situation auf Seite des Anwenders:

- Anwender verstehen die "Sprache" der Anbieter nicht.
- Beschreibungen der Geodaten und -Dienste reicht nicht aus, um fachliche Themen zu finden.
- Dienste-Angebot ist technisch zu kompliziert, inhaltlich inkompatibel und unzuverlässig.
- Web-Portale erfüllen nicht die Anforderungen der Nutzer.
- Die Interaktion zwischen Anbieter und Nutzer ist minimal.
- Metadaten zeigen oft ins Leere, sind unvollständig **vernetzt**.

# Neue Anforderungen

- Angebot erweitern und Suche vereinfachen.
- Sortieren und "stöbern" (browsing) ermöglichen.
- Option, Ergebnisse bewerten zu können.
- Einordnung in feste Kategorien durch Kommentierungen (Tagging) ergänzen.
- Erstellung und Pflege der Metadaten **automatisieren**.
- Interaktion von Anbieter, Broker und Nutzer deutlich vereinfachen.

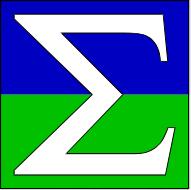
# Konzepte aus dem Web 2.0

## Das Potential von Metadaten im Web 2.0:

- REST-Architekturen ermöglichen flexible, einfache Erhebung, Pflege und Suche von Metadaten.
- Metadaten gehören in offene "Ablagen" oder "Eimer"
- Die Aktualisierung syntaktischer Metadaten können über Nachrichtendienste wie [GeoRSS](#) automatisiert werden.
- Die Erzeugung und Pflege von qualitativen ***und*** inhaltlichen Metadaten erfolgt durch Gesellschaften.

# Digitale Fundgruben

- Kataloge müssen durch flexible "**Ablagen**" (**Repositories**) ersetzt werden, einer Kombination aus Wiki, Portal, Verzeichnis, Datenbank und Lesezeichen.
- Ablagen ermöglichen **Tagging** (Datenkennung, Identifizierung) und Indizierung durch Suchmaschinen
- In aktuellen INSPIRE Architekturdiagrammen wurde der proprietäre "Service-Bus" durch "Internet" ersetzt.



# Filter on the way out!



# Digitale Fundgruben

- Kataloge müssen durch flexible "Ablagen" (Repositories) ersetzt werden, einer Kombination aus Wiki, Portal, Verzeichnis, Datenbank und Lesezeichen.
- Ablagen ermöglichen Tagging (Datenkennung, Identifizierung) und Indizierung durch Suchmaschinen
- In aktuellen INSPIRE Architekturdiagrammen wurde der proprietäre "Service-Bus" durch "**Internet**" ersetzt.
- Das **REST Architektur-Paradigma** beschreibt, wie das Internet effektiver genutzt werden kann: Addressierbar, unveränderlich, zustandslos und vernetzt.

# Technik: REST mit Sinn *verbinden*

## Vier Konzepte:

- die Ressource
- ihr Name (URL)
- ihre Repräsentation
- ihre Verbindungen (Links, Verweise)

## Vier Eigenschaften:

- Adressierbarkeit
- Zustandslosigkeit
- Verbundenheit
- Wohldefinierte Operationen

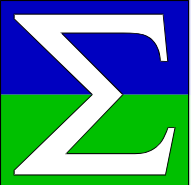
# Ausblick

*Aktuelle Katalog-Technik wird dem Potential digitaler Daten nicht gerecht. Technologien rund um "Ablagen" sind besser geeignet.*

- **Das "Ablagen"-Konzept muss in die INSPIRE Richtlinie eingearbeitet werden.**

*Das "Internet" ist bereits bestens für die Arbeit mit digitalen Geodaten gerüstet. Es muss lediglich "richtig" genutzt werden, hier: RESTful.*

- **REST-Paradigmen müssen durch die Standards des OGC berücksichtigt werden.**



## Veranstaltungshinweis:

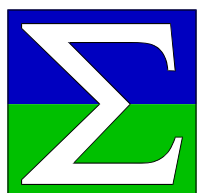
# Fragen & Diskussion



**6. bis 9. September 2010**

Die internationale  
Open Source Konferenz der  
Geo-Branche

Mit freundlicher Unterstützung von:



# Metaspatial

Dieser Foliensatz ist lizenziert  
Creative Commons Share Alike Lizenz 3.0  
Copyright: Arnulf Christl 2010