

Digital Skills und Nachhaltigkeit

**Vorlesung im Modul 10-201-2333 im
Wahlbereich Bachelor GSW, im Modul 10-
202-2330 im Master und Lehramt Informatik
sowie im Modul 10-202-2309 im Master
Informatik**

Wintersemester 2019/20

Prof. Dr. Hans-Gert Gräbe

<http://bis.informatik.uni-leipzig.de/HansGertGraebe>

Ausgangspunkt

- Beschreibungsformen (Plural) und Wirklichkeit (Realität)
- Widersprüchlichkeit der Welt (als von uns wahrgenommener Realität)
- Verschiedenheit des Begriffs *Widerspruch* in Beschreibungsformen und in Handlungsvollzügen
- Beschreibungen und Kontextualisierungen
 - Modell- und Begriffsbildung
 - Begriffe sind eine Form kooperativer Praxen von Menschen und damit selbst *konkret-historisch zu kontextualisieren*.
- Begriff *Weltbild* für den komplexen Zusammenhang des modellhaften Bezugs *im Modell* auf Wirklichkeit.

Welt ist Wirklichkeit für uns und damit Wirklichkeit im Prozess begrifflicher Erfassung.

Ausgangspunkt

Was sind Daten?

- Daten als spezifische Beschreibungsform
- Daten zu erfassen bedeutet stets Auswahl, anderes *nicht* zu erfassen.
- Daten als Link zwischen Welt und Wirklichkeit.
- Was aber sind dann *objektive Daten*?
 - Spezifischer Reflex eines positivistischen Wissenschaftsverständnisses.
 - Gebrauch und Missbrauch: Ein solches Verständnis (von Wissenschaft) ist eine wichtige kulturelle Errungenschaft der Menschheit, die aber ebenfalls *konkret-historisch zu kontextualisieren* ist.
- Daten sind damit ebenfalls eine Form kooperativer Praxen von Menschen.

Begriff des **digitalen Universums** als durch Verarbeitung von digitalen Daten eher technisch geprägter innergesellschaftlicher Handlungsraum mit vager Abgrenzung.

- Aufgreifen eines verbreiteten Buzz-Worts.
- „Im Jahre 2020 wird sich das digitale Universum auf 44 Billionen Gigabyte belaufen“ (EMC Digital Universe with Research & Analysis by IDC. The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things. April 2014).
- Bezug zur zentralen These – es wird mit einer Raummetapher gearbeitet, mit welcher der digitale Wandel aus einer spezifischen Dichotomie heraus analysiert wird.

Zentrale These: Der digitale Wandel wird geprägt durch eine schnell wachsende „Welt der digitalen Daten“, durch deren Analyse und Aufbereitung Einfluss auf realweltliche Prozesse genommen wird.

Zur Kritik dieses Ansatzes

- Mit dieser Fassung wollen wir uns auf Fragen konzentrieren, wie aktuell ablaufende Strukturierungsprozesse im digitalen Universum und realweltliche Prozesse zusammenspielen und sich gegenseitig beeinflussen.
- Konzept der Gegenüberstellung von „realweltlicher“ und „digitaler“ Realität ist insgesamt problematisch, da Handlungen im digitalen Universum sowohl aus realweltlichen Praxen heraus motiviert sind als auch Einfluss auf realweltliche Praxen haben.
- Konzept betont aber, dass viele realweltliche Wirkzusammenhänge mit technischen Prozessen in diesem Raum interagieren und deshalb eine solche Abstraktion sinnvoll erscheint.

Nutzung digitaler Handlungsräume

- Fragen privater digitaler Handlungsräume können nur sinnvoll diskutiert werden, wenn der Nutzer über einen **Account** an einem Rechner „eingeloggt“ ist. Das gilt auch für mobile Endgeräte, obwohl dort die technische Bindung an einen Account (über SIM-Karte und eigene Sicherheitseinstellungen) weniger sichtbar ist.
- Mit einem solchen Account ist eine **digitale Identität** verbunden, der Handlungen im Internet zugeordnet werden, über welche die üblichen rechtlich-sozialen Konstrukte der *rechtlichen Zurechenbarkeit von Handeln* in den digitalen Bereich übertragen werden.
 - Die private Zuordnung von Handlungsfolgen ist eine *Säule der bürgerlichen Rechtsordnung*.
 - Die technischen Möglichkeiten im digitalen Universum können die Zurechenbarkeit rechtlicher Verantwortung *erleichtern* oder *erschweren*.
 - Möglichkeit *anonymen Handelns*. Aber: Spuren von Handeln sind grundsätzlich einer forensischen Analyse zugänglich. Das gilt auch für Handeln im Netz.

Realweltliche und digitale Identitäten

- Die Zuordnung einer digitalen Identität zu einer realen Person erfolgt über eine **Authentifizierung**, die als (allerdings technisch präkonditionierter) *privater Akt* erscheint.
 - Setzt aber einen Authentifizierer als technische Gegenseite und damit einen übergeordneten rechtlichen Kontext voraus. Dieser Zuordnungsprozess wird dennoch öffentlich als privat postuliert.
- Private digitale Handlungsräume sind nur über die Bindung an eine digitale Identität gestaltbar.
 - Die Rückbindung einer digitalen Identität an ein bürgerliches Rechtssubjekt ist selbst ein sozio-technisch institutionalisierter Prozess.
 - Diese Rückbindung wird besonders einfach, wenn dem bürgerlichen Rechtssubjekt die Signatur eines technischen Artefakts aus dem digitalen Universum einfach zugeordnet werden kann.

Handeln im digitalen Universum

- Handlungsräume sind sozial determiniert. Digitale Handlungsräume können durch **Autorisierung** konstituiert und zugewiesen werden.
- Bei der Gestaltung von Handlungsräumen im digitalen Universum sind Subjekte in hohem Maße auf technische Dienstleistungen und damit auf externe Institutionen angewiesen, deren *Vertrauenswürdigkeit* sie angemessen einschätzen müssen.
- Ordnungsrechtliche Regelungen für Handeln im digitalen Universum existieren erst in Ansätzen, so dass *angemessenes praktisches Handeln* sowie *kooperative Gestaltung* auf *vertragsrechtlicher* Basis Hauptformen der Ausformung privater Handlungsräume sind.
- Ein *angemessenes* Verständnis der technischen Bedingtheiten, Möglichkeiten und Restriktionen des digitalen Universums ist für die qualifizierte Gestaltung eigenen digitalen Handelns unerlässlich.
- Soziales Handeln konstituiert die intersubjektiven Verhältnisse eines Subjekts.

Zum Begriff des Handlungsraums

These: Der Begriff des Handlungsraums im heutigen Verständnis ist eine kulturelle Errungenschaft der bürgerlichen Gesellschaft.

- Handlungsräume als „Raum im Raum“ kontextualisieren kooperative Gestaltungsmöglichkeiten in einem „äußeren Raum“.
- *Meine* Handlungsräume sind identitätskonstituierend, die Handlungsvollzüge in diesen Räumen Basis für das ICH als bürgerliches Rechtssubjekt.
- Erst auf dieser Basis sind Abgrenzungen anderer Begriffe wie *Umwelt, Handeln* in einer Umwelt, *kooperatives Handeln* und damit letztlich Begriffe wie *Subjekt, Privatheit* und *Identität* sinnvoll zu fassen.
- Gemeinschaftliche Handlungsräume können zu „kooperativen Subjekten“ im Sinne der bürgerlichen Rechtsordnung verdichtet werden.

Privates Handeln und (digitale) Identität

Privates Handeln setzt einen *Begriff des Ich*, einer eigenen *Identität* voraus.

- Digitale Identität, multiple digitale Identität und Rollen
 - Ist Identität teilbar?
- Abstrakte Identität, textuelle Repräsentation
 - Zuordnungsmechanismen, etwa Webseite und Login
- Authentifizierung
 - Passwort, andere Authentifizierungsformen
- Autorisierung
 - *Ich* als Subjekt und als Objekt von Autorisierung
- Potenzielle und reale Zuordnung. Begriff der *Session*.

Digitale Identitäten

- Digitale Identität, Abstrakte Identität, textuelle Repräsentation
 - Webseite, Login, mobile Endgeräte
 - Begriff der Session (nicht nur auf Webseiten)
 - Authentifizierung und Autorisierung

Wir werden im Weiteren unter einer *digitalen Identität* ein unter einer textuellen Repräsentation `<name@rechnername>` *authentifiziertes* und im Rahmen einer Session *autorisiertes* **realweltliches bürgerliches Subjekt** verstehen, das zeitlich begrenzt Handlungen im digitalen Universum vornimmt.

Digitale Identitäten und Rollen

Der Rollenbegriff der Informatik

- Als Rolle bezeichnet man in der Informatik ein Bündel von notwendigen *Erfahrungen, Kenntnisse und Fähigkeiten*, über die ein Mitarbeiter verfügen muss, um eine bestimmte *Aktivität* durchzuführen.
- Rollen sind dabei durch *Rollenbeschreibungen* innerhalb eines *Rollenmodells* definiert.
- Eine Rolle wird mit *Aktivitäten* und *Verantwortlichkeiten* verbunden.
- Für die Ausübung einer Rolle sind *Qualifikationsmerkmale* erforderlich.
- Eine Person kann mehrere Rollen inne haben. Mehrere Personen können jeweils die gleiche Rolle inne haben.

Internet Basics

Wir wollen im Weiteren den Begriff der *Rolle* als partielle Identität zu Grunde legen, wenn wir nun die technischen Gegebenheiten des Agierens digitaler Identitäten (genauer: *a/s* digitale Identitäten) betrachten wollen.

Im Internet werden *Beschreibungen* ausgetauscht

- Auch z.B. Bilder sind Beschreibungen, die dem Computer Anweisungen geben, wie das Bild zu rendern ist.
- Austausch von *Beschreibungen* zwischen Computern erfolgt, indem diese in *Pakete* vorgegebener Struktur und Größe zerlegt werden.

Paketübertragung im Internet, das OSI 7-Schichten-Modell

- <http://de.wikipedia.org/wiki/OSI-Modell>
- Schichten und Protokolle
- Protokolle und Sprache

Zur Vergabe digitaler Identitäten

- *Digitale Identität* = *authentifiziertes* und im Rahmen einer Session *autorisiertes* realweltliches bürgerliches Subjekt, das zeitlich begrenzt Handlungen im digitalen Universum vornimmt.
- Solche digitalen Identitäten fallen nicht vom Himmel, sondern müssen zwecks privater Zuordnung der Folgen von Handeln in die bürgerliche Rechtsordnung eingebettet werden.
 - Marktwirtschaft: Ordnungsrechtliche Rahmen und vertragsrechtliche Ausgestaltung in einem hierarchischen sozio-technischen System.
- Wer authentifiziert und autorisiert?
 - Rechner, Rechnernetze, Rechnernamen
 - Registrar, Provider, Host

Das Internet als Welt von Fiktionen

- Bitströme und **Datenpakete**
 - Im „Internet“ sind keine Bitströme unterwegs, sondern Datenpakete, die in Endgeräten über die 4 unteren Ebenen des OSI-Stacks aus Bitströmen erzeugt und wieder zurücktransformiert werden.
 - Fiktion der universell vernetzten Endgeräte und Realität der Netzausfälle
- Das Mausphänomen
 - Werkzeuge und deren Gebrauch. Der Löffel.
 - Fiktionen im Alltag. Diskussion.

Fiktion als gesellschaftlich gestützter, garantierter und aufrecht erhaltener *Konsens* einer verkürzenden *Sprechweise* über eine *gesellschaftliche Normalität*.

Digital Skills und Nachhaltigkeit

**Vorlesung im Modul 10-201-2333 im
Wahlbereich Bachelor GSW, im Modul 10-
202-2330 im Master und Lehramt Informatik
sowie im Modul 10-202-2309 im Master
Informatik**

Wintersemester 2019/20

Prof. Dr. Hans-Gert Gräbe

<http://bis.informatik.uni-leipzig.de/HansGertGraebe>

Das Internet als Welt von Fiktionen

- Bitströme und **Datenpakete**
 - Im „Internet“ sind keine Bitströme unterwegs, sondern Datenpakete, die in Endgeräten über die 4 unteren Ebenen des OSI-Stacks aus Bitströmen erzeugt und wieder zurücktransformiert werden.
 - Fiktion der universell vernetzten Endgeräte und Realität der Netzausfälle
- Das Mausphänomen
 - Werkzeuge und deren Gebrauch. Der Löffel.
 - Fiktionen im Alltag. Diskussion.

Fiktion als gesellschaftlich gestützter, garantierter und aufrecht erhaltener *Konsens* einer verkürzenden *Sprechweise* über eine *gesellschaftliche Normalität*.

Das Internet als Welt von Fiktionen

- Fiktionen sind eine spezifische Form des Umgangs mit einer steigenden Komplexität von Welt.
- Fiktionen in diesem Sinne sind kein neues Phänomen.
- Fiktionen und Mythen
 - Ein *Mythos* ist in seiner ursprünglichen Bedeutung eine Erzählung. Im religiösen Mythos wird das Dasein der Menschen mit der Welt der Götter oder Geister verknüpft. Mythen erheben einen Anspruch auf Geltung für die von ihnen behauptete Wahrheit. ... Das Ensemble aller Mythen eines Volkes, einer Kultur, einer Religion wird als *Mythologie* bezeichnet. (Wikipedia)

Das Internet als Welt von Fiktionen

Komplexität und Taktfrequenzen in der Gesellschaft:

- Ein Takt (oder eine Taktung) dient dem Aufprägen einer Periodizität auf einen Ablauf oder der Synchronisation von Vorgängen. Der Systemtakt in einem Computer bestimmt die Arbeitsgeschwindigkeit vieler Komponenten. (Wikipedia)
- Taktung ist auch für die Koordinierung und Synchronisation gesellschaftlichen Handelns unerlässlich.
- Entwicklung von Komplexität und Taktraten von Computerchips siehe <http://www.pc-erfahrung.de/prozessor/cpu-historie.html>
- Moore's Law (1965) besagt, dass sich die Komplexität integrierter Schaltkreise mit minimalen Komponentenkosten regelmäßig verdoppelt; je nach Quelle werden 12 bis 24 Monate als Zeitraum genannt.

Daten und Informationen Syntax, Semantik, Pragmatik

Informationen = interpretierte Daten

Daten = formalisierte Informationen

Beides (Formalisierung und Interpretation) sind nur in einem *speziellen Kontext* natürlicher, technischer oder sozialer Gegebenheiten – einem *Kontext* (oder Pragmatik) – „gültig“ und setzen damit eine „funktionierende Fiktion“ voraus.

Vergleiche auch das Konzertbeispiel in der ersten Vorlesung.

Syntax, Semantik, Pragmatik im OSI-Schichtenmodell

Pragmatisch kontextualisiertes Wechselspiel von (formalisierter) *Syntax* und (zu formalisierender) *Semantik* auf verschiedenen Ebenen am Beispiel des OSI-Stacks.

<https://de.wikipedia.org/wiki/OSI-Modell> und nächste Folie.

Jede Schicht geht von einer Fiktion (= gesellschaftlicher Normalität) und ihrer als formalisierter Syntax gegebenen sprachlichen Repräsentation aus, die auf der vorhergehenden Schicht praktisch hergestellt wird, und realisiert auf dieser Basis eine weitere Pragmatik durch dafür entwickelte spezielle Sprechweisen (Semantik), die ihrerseits für den Gebrauch auf der nächsten Schicht zu formalisieren ist.

Syntax, Semantik, Pragmatik im OSI-Schichtenmodell

Erläuterung dieses Gedanken:

Schicht 1: Syntax = modulierte Wellen, Semantik = Bitfolgen (erste Fiktion), Pragmatik = Verschiedenheit der Übertragungsmedien

Schicht 2: Syntax = Bitfolgen, Semantik = Frames (zweite Fiktion), Pragmatik = Steuerung der Übertragungsgeschwindigkeit der Bitfolgen, Ergänzung von Prüfsummen zur Fehlererkennung

Schicht 3: Syntax = Frames, Semantik = Datenpakete (dritte Fiktion), Pragmatik = Routing und Organisation der Weiterleitung von Paketen über mehrere Knoten

Usw.

Beschreibungen und Interpretationen

- Information als interpretierte Daten?
 - Messwerte als Daten?
- Sprache ist voraussetzungsreich. Ein Beispiel:
 - Am 8.11. wurde an der Station Leipzig-Flughafen um 17 Uhr eine Temperatur von 16°C gemessen.
 - Am `<a type="Datum">8.11.</a>` wurde an der *Station* `<a type="Ortsangabe">Leipzig-Flughafen</a>` um `<a type="Uhrzeit">17 Uhr</a>` eine `<a type="physikalischeGroesse">Temperatur</a>` von `<a type="Temperatur">16°C</a>` *gemessen*.
 - Dinge und deren Bezeichnungen.

RDF Basics (1)

Beispiel: Beschreibung unserer Lehrveranstaltungen

```
@prefix od: <http://od.fmi.uni-leipzig.de/model/> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
@prefix odr: <http://od.fmi.uni-leipzig.de/rooms/> .  
@prefix odp: <http://od.fmi.uni-leipzig.de/personal/> .  
  
<http://od.fmi.uni-leipzig.de/w19/SWS.IADW.1>  
  a od:Vorlesung ;  
  od:beginsAt "11:15" ;  
  od:dayOfWeek "donnerstags" ;  
  od:endsAt "12:45" ;  
  od:locatedAt odr:Hs_19 ;  
  od:servedBy odp:Graebe_HansGert ;  
  rdfs:label "Vorlesung ..." .
```

- Identifier und Literale. Namensräume.

RDF Basics (2)

Auflösung in Drei-Wort-Sätze

Subjekt Prädikat Objekt .

W19:SWS.IADW.1	a	od:Vorlesung .
W19:SWS.IADW.1	od:beginsAt	"11:15" .
W19:SWS.IADW.1	od:dayOfWeek	"donnerstags" .
W19:SWS.IADW.1	od:endsAt	"12:45" .
W19:SWS.IADW.1	od:locatedAt	odr:Hs_19 .
W19:SWS.IADW.1	od:servedBy	odp:Graebe_HansGert .
W19:SWS.IADW.1	rdfs:label	"Vorlesung ..." .

Mehr dazu siehe <http://od.fmi.uni-leipzig.de/>

Interpretationen im Internet der Dinge

- Industrie 4.0 und das Internet der Dinge (IoT – Internet of Things)
- Fiktion: Im Internet gibt es keine Dinge, sondern nur *Repräsentationen* von Dingen, genau wie Repräsentationen von Personen.
- Beschreibungen als relationale Komplexe zwischen Repräsentationen von realen Dingen oder auch nur Bedeutungskomplexen.
- Auch diesen Dingen und Bedeutungskomplexen müssen „digitale Identitäten“ als textuelle Repräsentationen zugeordnet werden, um über sie im digitalen Universum Sätze formulieren zu können.

RDF Basics (3)

Konzeptionelle „Zutaten“:

- UTF-8 als **einheitliche Zeichenbasis** für URIs und Literale.
 - Best Practise: URIs nur aus ASCII-Zeichen, keine Umlaute oder Ähnliches
- URI als „digitale Identitäten“ von Ressourcen, *zeigen* auf Ressourcen
 - Wie digitale Identitäten von Personen sind dies **textuelle Repräsentationen der „Dinge“** in den im Internet kursierenden Textfragmenten.
- Für Computer sind URIs einfach Zeichenketten, für Menschen ist es hilfreich, wenn die URI bereits eine Ahnung von der Semantik des Gezeigten vermittelt.
 - Best Practise: „sprechende Namen“ als URIs

RDF Basics (4)

RDF – Resource Description Framework

- Konzept zum Aufschreiben von Geschichten über „die Welt“ als *Mengen* von Drei-Wort-Sätzen
 $\langle \text{Subjekt} \rangle \langle \text{Prädikat} \rangle \langle \text{Objekt} \rangle .$
- Subjekt und Prädikat müssen URIs sein, als Objekt kann eine URI oder ein Literal (Typ `rdfs:Literal`) stehen. Literale können Typ- und Sprachmarkierungen tragen.
- Es gibt verschiedene Notationen für dieselbe Menge von RDF-Sätzen (Turtle, rdf/xml, json, ntriples) und Werkzeuge, diese Notationen ineinander umzuwandeln.
 - Redland RDF libraries <http://librdf.org/>
- Mustersuche als mächtiges Konzept der Analyse derartiger Mengen. SPARQL als Anfragesprache.

RDF Basics (5)

Beispiel einer Anfrage an den SPARQL Endpunkt

<http://od.fmi.uni-leipzig.de:8892/sparql>

Einfache Anfrage nach allen LV mit URI-Präfix SWS.IADW

PREFIX od: <http://od.fmi.uni-leipzig.de/model/>

SELECT distinct ?l ?name ?d ?b ?r

from <http://od.fmi.uni-leipzig.de/w19/>

WHERE {

 ?l a od:LV .

 ?l rdfs:label ?name .

 ?l od:beginsAt ?b .

 ?l od:dayOfWeek ?d .

 ?l od:locatedAt ?r .

filter regex(?l,'SWS.IADW') .

}

RDF Basics (6)

- **Turtle-Notation** – fasst alle Sätze zum selben Subjekt zusammen. Menge von Prädikat-Objekt-Paaren kann als Menge von Schlüssel-Wert-Paaren (key – value) betrachtet werden, welche dieses Subjekt beschreiben.
 - Aber: ein Schlüssel kann mehrere Werte haben!
 - Besonders verbreitete von Menschen lesbare Notation.
 - Subjekt-zentrierte Sicht, die spezifische Sichtweisen auf „MEINE Welt“ – wie früher thematisiert – gut bedient.
 - Computer arbeiten lieber mit Tripel-Mengen.
- Interpretiert man Subjekte und Objekte als Knoten und Prädikate als Kanten eines Graphen, so beschreibt eine Menge von RDF-Sätzen einen **RDF-Graphen** (und umgekehrt).
 - Ein Bild sagt oft mehr als tausend Worte.

RDF – Sätze und Muster

Sätze folgen dabei Mustern:

1) Turtle: Sammle alle Sätze mit demselben Subjekt → Eigenschaften eines Individualsubjekts als Schlüssel-Wert-Paare.

- Schlüssel-Wert = Eigenschaft und Ausprägung

2) Sammle alle Sätze mit demselben Prädikat

A od:beginsAt B

- od:beginsAt nicht nur eine URI (*Syntax*), sondern auch ein zweistelliges Prädikat mit einer gewissen *Semantik*, die in allen Sätzen mit diesem Prädikat als *Instanziierungen des Prädikats* transportiert wird.

3) Andere Muster sind möglich, SPARQL als genereller Standard einer Anfragesprache für Mustersuche.

Siehe die Datei *Queries.txt* im Ordner des Kurses.

RDF – Beschreibungen von Beschreibungen

- **Selbstähnlichkeit des Konzepts:** Auch Beschreibungen von Beschreibungen können als RDF-Sätze formuliert werden. Insbesondere kann man RDF verwenden, um RDF zu beschreiben.
 - Eine URI, die in einem Satz als Prädikat auftritt, kann in einem anderen Satz als Subjekt oder Objekt auftreten.
Beispiel:
 `od:beginsAt rdfs:domain od:LV .`
 `od:beginsAt rdfs:range rdfs:Literal .`
- Damit können auch *Begriffe und Konzepte* RDF-basiert beschrieben werden. → **Universalien**
 - Was sind Universalien? Ideen aus Platons Ideenhimmel oder institutionalisierte Konventionen, d.h. „Fiktionen“ im Sinne der Vorlesung?

RDF – Grundlegende Beschränkungen

- Mengensemantik, Reihenfolge der Sätze spielt keine Rolle
 - Dies ist in anderen Ansätzen, etwa dem XML-basierten TEI der Text Encoding Initiative anders, das in den Digital Humanities eine zentrale Rolle spielt.
- Problem der Kontextualisierung. In welchem raumzeitlichen Kontext ist der Satz gültig? Hier gibt es mehrere Ansätze:
 - Aufbohren zu Quadrupeln mit vierter Komponente als URI auf die Provenienz.
 - Wird der Satz als Instanziierung eines zweistelligen Prädikats interpretiert, so ergibt sich der Kontext oft implizit aus dem Prädikat.
 - Oft ergibt sich der Kontext allgemeiner aus dem Namensraum des Prädikats und steht damit als (explizites oder implizites) Modell für eine ganze Klasse von Begrifflichkeiten. Damit verlagert sich die Frage aber nur in die Beschreibung des Modells und damit eine Abstraktionsebene nach oben.

RDF – Zentrale Konzepte

Zusammenfassung der zentralen Konzepte

- Zentrale Idee: Speichere textuelle Beschreibungen auf uniforme Weise als Tripel und verwende Standardkonzepte und -werkzeuge zur Verwaltung dieser Daten
- *Ressourcen*: URI, HTTP access
 - URI = Unique Resource Identifier
 - Damit kann auf einen weltweit verteilten Datenbestand auf uniforme Weise über ein gängiges Protokoll zugegriffen werden.
- *Resource Descriptions*: Gib auf Anfrage an die HTTP-Adresse ein nützliches Stück Information im RDF-Format zurück, das mit anderen solchen Informationseinheiten zu neuen RDF-Sätzen kombiniert werden kann.
- Betreibe *RDF Triple Stores* als Teil einer weltweiten verteilten Datenspeicher-infrastruktur, etwa <http://od.fmi.uni-leipzig.de/>
- (Verteilte) Anfragesprache SPARQL
 - Stelle *SPARQL Endpunkte* auf RDF Triple Stores bereit

RDF – Sprachformen und Praxen

Verfahrenswissen → Verfahrensweisen

- Korrespondenz zwischen der Kohärenz der Sprachform und der Kohärenz der Praxen
- Die Etablierung kohärenter Praxen als Verfahrensweisen erlaubt es, Prädikate zu substantivieren. „Zum Beginn der Schulstunde werden die Pausenbrote weggepackt“. Erst nach einer solchen Transposition des Prädikats an die Subjekt-Position werden normative Sätze möglich.
 - Parallelen zum Konzertbeispiel in der ersten Vorlesung.
- Grundweisheiten der Informatik:
 1. Eine Funktion kann erst aufgerufen werden, nachdem sie definiert wurde.
 2. Eine Funktion, die definiert wird, aber nicht aufgerufen, deutet auf einen Designfehler hin.
- *Verfahrenswissen* ist die Beschreibungsform, *Verfahrensweise* die Ausführungsform.

Die Linked Open Data Cloud

Noch einmal zur Grundidee

- Auf diese Weise entsteht eine weltweit vernetzte dezentrale *offene* Datenbank, die Linked Open Data Cloud, in der alle *öffentlichen* Informationen freizügig und maschinenlesbar verfügbar sind.
 - Der Spieß des Datensammelns wird umgekehrt und an einer Welt gebaut, in der ein *Kernbestand öffentlich verfügbarer Informationen* als Public Domain eine wesentliche kulturelle Konstituante ist.
 - Siehe <http://lod-cloud.net/>
 - Wachsen der LOD Cloud: <http://lod-cloud.net/versions/>
- Kontext von Industrie 4.0 und allen großen Datenprojekten, auch Google ist ohne diese Bemühungen nicht denkbar.
- Basis für innerbetriebliche und zwischenbetriebliche Informationssysteme wie ERP und CRM.

Namensräume und Begriffswelten

Kommunikation wird durch die Einführung von **Namensräumen** als URI-Präfix unterstützt.

Namensräume erlauben es, überlappungsfrei URIs zu generieren.

- Damit können Beschreibungen erzeugt werden, welche die *Fiktionen* MEINE Welt, MEINE Begriffe, ICH-Kern, Welt und Wirklichkeiten, Wirklichkeitskonstruktion bedienen, ohne diese gedanklich transzendieren zu müssen.
- Aber wir wollen ja mehr: Kooperation mit konkreten Anderen.
- Semantik = pragmatisch kontextualisierte Bildung von *Modellen* als Basis für gemeinsame *Verfahrensweisen*.
- Wir brauchen Sprache, um über die Modelle selbst sprechen zu können, und damit Wege der Formalisierung von Semantik.

Konzeptualisierungen und Begriffswelten

Aber: **Das Turmbau von Babel Phänomen**

- Was bedeutet es, dass jeder kommunikative Zusammenhang hierbei zusammen mit den eigenen Verfahrensweisen auch seine eigenen Modelle und damit seine eigene Sprache entwickelt und spricht?
- Welche Konzepte können Übersetzungsleistungen unterstützen?
- **Ontologien** (oder Vokabulare): WIR einigen uns auf die Verwendung gemeinsamer Namensräume (foaf, skos, org, sioc usw.) für *spezielle* Zwecke und damit auf gemeinsame *Teilmodelle* von Welt.
 - Phänomen der Kohärenz privater und kooperativer Sprachpraxen.
- Wie geht das genau?
 - **Inhalt:** Pragmatisch kontextualisierte Bildung von *Modellen* als Basis für gemeinsame *Verfahrensweisen*.
 - **Form:** *Ontologisierung* als pragmatisch kontextualisierte Semantifizierung von Syntax.
 - Beispiele auf der nächsten Folie

Vokabulare und Konzeptualisierungen

- **Beispiel foaf: Friend of a Friend**
 - foaf: <<http://xmlns.com/foaf/0.1/>>
 - Weiterleitung auf <http://xmlns.com/foaf/spec/>
 - Wir studieren das dort entwickelte Modell und die Beschreibungsformen von Semantik und Syntax
- **Beispiel skos: Simple Knowledge Organization System**
 - skos: <<http://www.w3.org/2004/02/skos/core#>>
 - Weiterleitung auf tabellarische Übersicht
<https://www.w3.org/2009/08/skos-reference/skos.html>
 - Ganz am Ende der Seite drei Referenzen
- **Beispiel org: The Organization Ontology**
 - org: <<http://www.w3.org/ns/org#>>
 - Weiterleitung auf eine Turtle-Datei. Herunterladen und anschauen
 - rdfs:seeAlso <<https://www.w3.org/TR/vocab-org/>>

Vokabulare und Konzeptualisierungen

- Sozial ein extrem schwieriger Prozess, aber das ist **der Kern semantischer Technologien**: Die *Institutionalisierung* maschinenlesbarer gemeinsamer Begriffswelten *als sozialer Prozess*.
- Damit verbunden sind *Modellbildungen*, Bedingtheiten (Kontextualität verschiedener Wirklichkeiten) und der Prozess des Transzendierens von Kontexten, wenn Ontologien nicht wie ursprünglich vorgesehen angewendet werden.
 - Miteinander sprechen – Vereinbarung von Ontologien
 - Weiterentwicklung von Ontologien
 - Große Datenbanken von Ontologien: <http://prefix.cc> oder <http://lov.okfn.org> (Linked Open Vocabularies)
 - Kreativität im kooperativen Kontext. Formalisierungserfordernis, um Informationen als Daten auszutauschen. Noch einmal das Konzertbeispiel.

Zwei Beispiele

- DBPedia – Extrahiert strukturierte Information aus Wikipedia
 - DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web. ... We hope that this work will make it easier for the huge amount of information in Wikipedia to be used in some new interesting ways. ...
 - Beispiel: <http://dbpedia.org/page/Leipzig>
- Linked-Data-Service der Deutschen Nationalbibliothek
 - Die Deutsche Nationalbibliothek erstellt einen Linked-Data-Service, der langfristig eine Nutzung der gesamten nationalbibliografischen Daten inklusive aller Normdaten durch die Semantic-Web-Community erlaubt. Sie ist bestrebt, durch diesen Datenservice einen Beitrag zur Informationsinfrastruktur weltweit zu leisten und damit eine Voraussetzung für moderne kommerzielle und nicht-kommerzielle Webdienstleistungen anzubieten.
 - <http://www.dnb.de/lds>

Schema.org

- Anderer Zugang: <http://schema.org> - Googles Ontologisierung der Welt und Einbau in Webseiten statt Aufbau einer verteilten Datenbank wie in der Linked Open Data Cloud.
- Schema.org und Microdata: <https://schema.org/docs/gs.html>
 - itemscope, itemtype und itemprop und die Verbindung zu RDF.
- Auszeichnung von Webseiten mit diesem Markup erhöht deren Sichtbarkeit bei Google.

Googles Knowledge Graph

- **Googles Knowledge Vault:** Ansatz, durch supervised learning aus untersuchten Webseiten entsprechende Fakten als Googles Wissensbasis zu extrahieren.
 - Enthielt 2014 über 1.6 Milliarden Fakten, die mit einem probabilistischen Konfidenzwert bewertet sind.
- **Google Knowledge Graph:** Konsolidierung und Anreicherung mit strukturierten Fakten aus Freebase (2007 gegründet, 2010 von Google aufgekauft), Wikipedia und Wikidata.
 - Enthielt 2016 über 70 Mrd. Fakten.
 - Ende 2015 wurde die Google Knowledge Graph API veröffentlicht, über die Webentwickler auf den Bestand zugreifen können.
 - <https://www.sem-deutschland.de/seo-firma/seo-glossar/knowledgege-graph/>

Wolfram Alpha

- Ebenfalls Suchmaschine, die auf Fakten aufbaut, die aus eigener Recherche gewonnen wurden. Zusammen mit *Mathematica* als Compute Engine lassen sich komplexere Präsentationen und Visualisierungen erstellen. Ziel ist die Vernetzung von mathematischem Wissen und Allgemeinwissen.
- <https://www.wolframalpha.com>
 - Beispiel „Leipzig“.