

Kreativität und Technik

**Vorlesung im Modul 10-201-2334
im Wahlbereich Bachelor GSW
sowie im Modul 10-201-2333
im Bachelor Informatik**

Sommersemester 2016

Prof. Dr. Hans-Gert Gräbe

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

Das Internet als Welt von Fiktionen

Fiktion der universellen Ende-zu-Ende-Verbindung und deren Realisierung als **skalenfreies Netz**:

Beschreibungsebene:

- $v(k) = c \cdot k^{-a}$ – Anteil der Knoten mit k Nachbarn (v wie Valenz)
- Beispiel mit $a=3$: $v(1)=0.832$, $v(2)=0.104$, $v(3)=0.031$, $v(4)=0.013$, $v(5)=0.007$, $v(6)=0.004$, ...
- Gegenüber einem Zufallsnetz (eigenständiger Vergleichsbegriff!) fällt der Anteil von Knoten mit vielen Verbindungen (Hubs) langsamer.
- Empirie: Typisches Phänomen eines organisch wachsenden Netzes, in dem Knoten nacheinander angelegt und Knoten mit vielen Valenzen systematisch bevorzugt werden.

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

Das Internet als Welt von Fiktionen

Fiktion der universellen Ende-zu-Ende-Verbindung und deren Realisierung als **skalenfreies Netz**:

Phänomenologische Ebene:

- *small world effect*: Mittlere kürzeste Weglänge zwischen zwei Knoten ist sehr gering.
- Hohe Resistenz gegenüber *zufälligen* Störungen: Fallen zufällig gewählte Knoten aus, so bleibt das Netz lange zusammenhängend. Netz bleibt robust, wenn derartige Störungen zügig beseitigt werden.
- Hohe Empfindlichkeit gegenüber *gezielten* Angriffen auf Hubs – das Netz zerfällt dann schnell in nicht miteinander verbundene Teilnetze.

Daten und Informationen Syntax, Semantik, Pragmatik

Informationen = interpretierte Daten

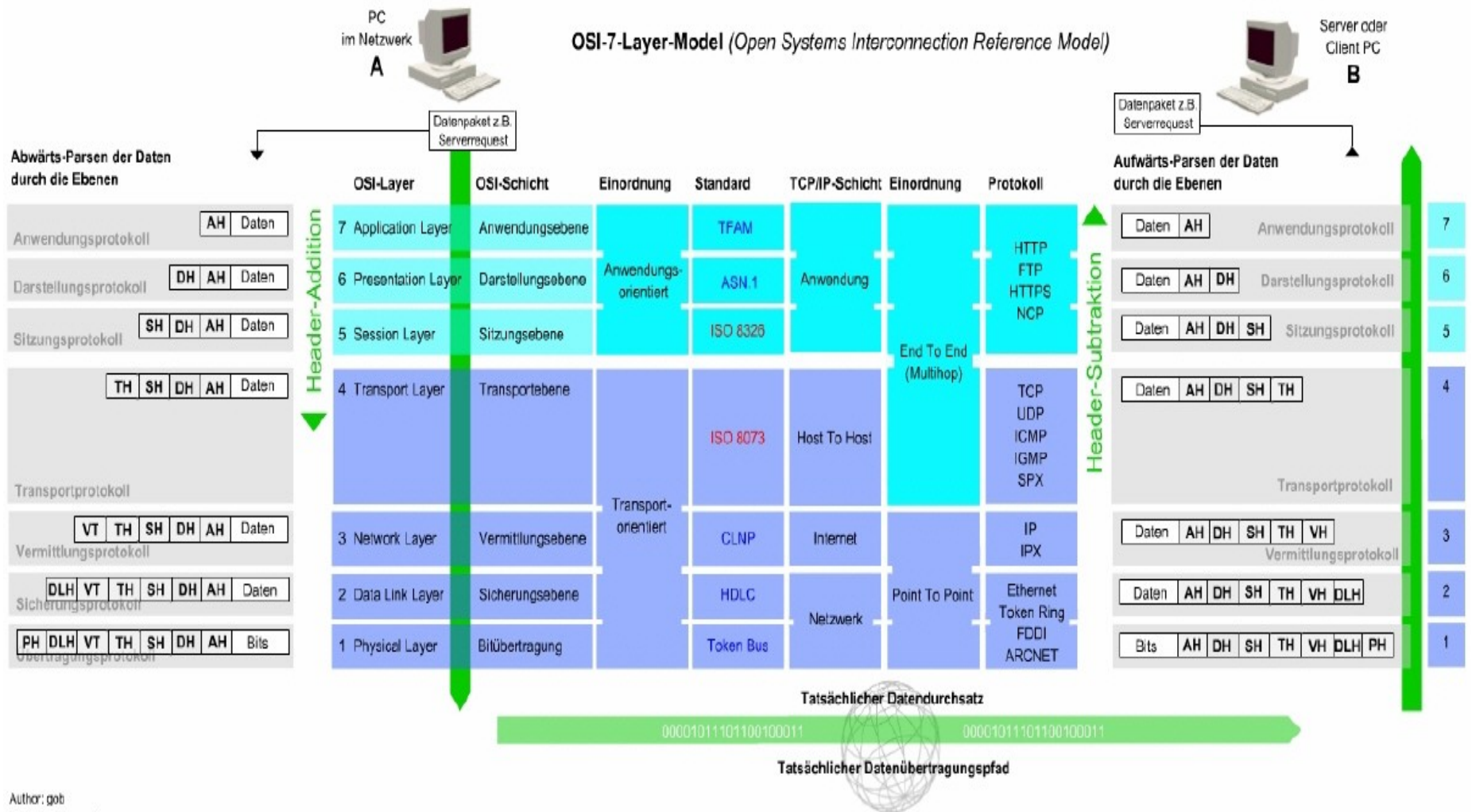
Daten = formalisierte Informationen

Einschub – Anmerkungen zum Seminar am 3.5.2016:

Pragmatisch kontextualisiertes Wechselspiel von (formalisierter) Syntax und (formalisierter) 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.

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/s16/BIS.KT.1>  
  a od:Vorlesung ;  
  od:beginsAt "11:15" ;  
  od:dayOfWeek "dienstags" ;  
  od:endsAt "12:45" ;  
  od:locatedAt odr:Hs_19 ;  
  od:servedBy odp:Gräebe_HansGert ;  
  rdfs:label "Vorlesung ..." .
```

- Identifier und Literale.

RDF Basics (2)

Auflösung in Drei-Wort-Sätze

Subjekt – Prädikat – Objekt .

```
S16:BIS.KT.1 a od:Vorlesung .  
S16:BIS.KT.1 od:beginsAt "11:15" .  
S16:BIS.KT.1 od:dayOfWeek "dienstags" .  
S16:BIS.KT.1 od:endsAt "12:45" .  
S16:BIS.KT.1 od:locatedAt odr:Hs_19 .  
S16:BIS.KT.1 od:servedBy odp:Graebe_HansGert .  
S16:BIS.KT.1 rdfs:label "Vorlesung ..." .
```

Mehr dazu in der Datei *Kurs.ttl* im Material-Ordner.

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
 <Subjekt> <Prädikat> <Objekt> .
 - Subjekt und Prädikat müssen URIs sein, als Objekt kann eine URI oder ein Literal (Typ `rdf: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.

RDF Basics (5)

- **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: Im Gegensatz zu Excel-Tabellen kann ein Schlüssel mehrere Werte haben!
 - Besonders verbreitete von Menschen lesbare Notation.
 - Subjekt-zentrierte Sicht, die spezifische Sichtweisen auf „die 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 Basics (6)

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
 - Access to worldwide distributed data in a unified way
- *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
 - Siehe Beispieldatei *Queries.txt* im Kursordner