

Semantic Web



Prof. Dr. Gerd Stumme
Dipl.-Math. Robert Jäschke

Wintersemester 2005/06



ENDOWED CHAIR OF THE HERTIE FOUNDATION
Knowledge and Data Engineering
DEPARTMENT OF MATHEMATICS & COMPUTER SCIENCE

Organisatorisches

Vorlesung

- Beginn: 1. November 2005
- dienstags 14.15 - 15.45, Raum 0443
- Die Folien sind teilweise in Englisch, der Tafelanschrieb auf Deutsch

Übungen

- Mittwochs, 14.15-15.45
- Beginn: 9. November 2005
- wird als Präsenzübung abgehalten (s. nächste Folie)

Prüfung

- Die Prüfung wird je nach Teilnehmerzahl mündlich oder schriftlich abgehalten.

Organisatorisches

Präsenzübung bedeutet

- **selbständiges Bearbeiten** des Übungsblattes in Kleingruppen à 3-4 Personen unter Betreuung eines Mitarbeiters,
- **kein prinzipielles Wiederholen** des Vorlesungsstoffs,
- **kein Vorrechnen** der Musterlösung etc. (Diese wird später zur Verfügung gestellt.)

Nötig dafür:

- selbständige Vorlesungsnachbereitung **vor** der Übung,
- Mitbringen des Skriptes ,
- eigene Aktivität entfalten.

Organisatorisches

Warum ein neues Übungskonzept?

- aktives Erarbeiten des Vorlesungsstoffes bringt mehr
- Zusammenhänge im Stoff erkennen
- strukturiertes Denken und selbständiges Arbeiten lernen
- Teamarbeit lernen
- Erklären lernen (als Tutor und als Teilnehmer)
- Klausurtraining ;-)
- *Ihr Studium der ... haben Sie abgeschlossen. Zu Ihren persönlichen Stärken zählen Sie Eigeninitiative, Kommunikations- und Kooperationsbereitschaft, Teamarbeit.* (Typischer Anzeigentext)

Organisatorisches

Sprechstunden nach Absprache:

Prof. Dr. Gerd Stumme (Vorlesung): stumme@cs.uni-kassel.de 0561/804-6251
Dipl.-Math. (Übungen): jaeschke@cs.uni-kassel.de 0561/804-6254

FG Wissensverarbeitung, FB Mathematik/Informatik
Raum 0439, Wilhelmshöher Allee 73

Informationen im Internet: <http://www.kde.cs.uni-kassel.de>

Hier ist u.a. folgendes zu finden:

- aktuelle Ankündigungen
- Folienkopien
- Übungsblätter
- Literaturempfehlungen
- Termine



ENDOWED CHAIR OF THE HERTIE FOUNDATION
Knowledge and Data Engineering
DEPARTMENT OF MATHEMATICS & COMPUTER SCIENCE

Themenübersicht

- A. Motivation
- B. Wissensprozesse
- C. Künstliche Intelligenz
- D. Wissensrepräsentation
- E. Semantic Web
- F. Beschreibungslogiken



A. Motivation

G. Schröder (CEBIT 2000):

- Wir befinden uns am Übergang von der Industrie- in die Wissens- und Informations-Gesellschaft.
- Wir müssen den Menschen [...] die notwendigen Hilfestellungen geben, damit sie sich Zugang verschaffen können in die Wissens- und Informationsgesellschaft.

Wolf-Michael Catenhusen (BMBF, 1999):

- Deutschland steht mitten in einem weltweiten Prozess, in dem der Rohstoff Information, seine Erzeugung, Speicherung, und Verarbeitung, die Gewinnung und der richtige Umgang mit Wissen eine strategische volkswirtschaftliche und gesellschaftliche Größe geworden ist.

Einführung in die Wissensverarbeitung, Kassel, WS 2005/06

7

A. Motivation

Wissen ist ein wichtiger Erfolgsfaktor für Unternehmen:

- **Lebenszyklus** für Produkte wird immer kürzer
 - schnellere Umsetzung neuer Erkenntnisse in innovative Produkte
- **Lean Management** erfordert Entscheidungen auf allen organisatorischen Ebenen
 - erforderliches Wissen muss auf allen Ebenen verfügbar sein
- Unternehmen operieren mehr und mehr **global**
 - geographisch verteilte Unternehmensbereiche
 - virtuelle Unternehmen
- **Concurrent Engineering** erfordert die Integration von Wissen aus unterschiedlichen Unternehmensbereichen

Einführung in die Wissensverarbeitung, Kassel, WS 2005/06

8

A. Motivation

- Anteil **wissensintensiver** Geschäftsprozesse nimmt zu (in den Industrienationen)
 - hoher Freiheitsgrad bei der Abwicklung
 - hoher Anteil des Aufwands besteht aus
 - Suche nach Wissen
 - Auswahl relevanten Wissens
- Wissen wird selbst zum **Produkt**
 - neue Arten spezialisierter Serviceunternehmen entstehen,
 - z.B. Informationsbroker
- **Wert** eines Unternehmens hängt nicht mehr allein von seinen materiellen Vermögenswerten ab, sondern auch vom im Unternehmen verfügbaren **Wissen**,
 - z.B. Beratungsunternehmen

A. Motivation

Wie ist die Ist-Situation?

- „Je spezieller/kritischer mein eigenes Know-How, desto wichtiger bin ich für das Unternehmen.“
 - Wissen ist Macht
 - Wissen wird nicht geteilt
- **Weggang** eines Mitarbeiter führt zu **Know-How-Verlust** für Unternehmen
- **langwierige** Einarbeitung neuer Mitarbeiter
- **Fehler** werden **wiederholt** gemacht
 - niemand spricht über eigene Fehlschläge
- Entscheidungen werden **ohne** Ausnutzung allen relevanten Wissens getroffen

Was ist Wissen und wie kann es repräsentiert werden?

Kaffeemaschine Um eine Kaffeemaschine auf umweltfreundliche Weise zu entkalken, füllt man die Maschine mit Wasser und gibt 2 TL Zitronensäure (aus der Apotheke) hinzu. Man schaltet die Maschine ein und läßt die Mischung durchlaufen. Danach zweimal klares Wasser durchlaufen lassen.

Aus: Gewusst wie - 10000 praktische Tips für alle Tage. ADAC Verlag, München 1993

Übung:

- a) **Wie kann dieses Wissen repräsentiert werden, damit es von einem Computer verarbeitet werden kann?**
- b) **Welche Aspekte sind dabei Ihrer Ansicht nach wichtig?**

→ 10 Minuten

Daten, Information, Wissen

Begriffsdefinitionen

Es folgen Definitionen der Begriffe

- Daten,
- Information und
- Wissen

aus Sicht verschiedener Disziplinen, wie

- Philosophie,
- Psychologie und
- Betriebswirtschaftslehre.

(Dieser Teil wurde übernommen von der Vorlesung Wissensmanagement, SS 2004, des Instituts AIFB, Universität Karlsruhe.)

Webster's New Encyclopaedic Dictionary

- **Data:** factual information (as measurements or statistics) used as a basis for reasoning, discussion, or calculation
- **Information:**
 1. the communication or reception of knowledge or intelligence
 2. a) knowledge obtained from investigation, study, or instruction
 - b) knowledge of a particular event or situation (news)
 - c) fact, data
 - d) a signal or mark put into or put out by a computing machine

Wissen in der Philosophie

Wissen hat etwas mit Wahrheit zu tun (dies geht bis zu **Plato** zurück)

- nach **Kant:** „Ein sowohl subjektiv als auch objektiv zureichendes Fürwahrhalten.“
- nach **Fries:** „Fürwahrhalten mit vollständiger Gewissheit.“
- nach **Husserl:** „Evidenz davon, dass ein gewisser Sachverhalt gelte oder nicht.“
- nach **Sodian:** „begründeter, wahrer Glauben über einen Sachverhalt, d.h. wahrer Glauben, der durch die Verarbeitung entsprechender Informationen zustande gekommen ist.“

- **Knowledge:**
 1. understanding gained by actual experience
 2. a) the state of being aware of something or of having information
 - b) range of awareness or information
 3. the act of understanding: clear perception of truth
 4. something learnt and kept in mind

Wissen vs. Glauben vs. Meinen:

- **Wissen** = subjektiv und objektiv zureichendes Fürwahrhalten
- **Glauben** = subjektiv zureichendes, objektiv unzureichendes Fürwahrhalten
- **Meinen** = subjektiv und objektiv unzureichendes Fürwahrhalten

Die Vorlesung folgt diesem Ansatz.

Semiotische Sichtweise

- **Daten** beinhalten lediglich eine **syntaktische Dimension**
 - z.B. R2D2, 1500 ...
- **Informationen** beinhalten **Syntax** und **Semantik** (Form und Inhalt)
 - z.B. R2D2 (Form) ist der Name eines Roboters (Inhalt)
 - auf dem Typschild eines Motors bedeutet 1500 (Form) z.B. die Anzahl Umdrehungen pro Sekunde (Inhalt)
- **Wissen** beinhaltet darüberhinaus eine **pragmatische Dimension**, ist also mit einem Ziel oder Zweck verbunden.

Daten	= Zeichen + Syntax
Information	= Daten + Bedeutung
Wissen	= Internalisierte Information + Fähigkeit, sie zu nutzen [Devlin 1999]

Daten, Information, Wissen

Wissen in der Wissenspsychologie

- Wissen ist etwas, was jemand für wahr hält.
- Wissen regelt die
 - Gliederung und Interpretation eintreffender Informationen,
 - Erzeugung adäquater Reaktionen.
- Wissen ist die Basis des
 - Wahrnehmens,
 - Handelns und
 - Sprechens.

Wissen in der BWL

Es gibt auch hier viele verschiedene Definitionen des Begriffs Wissen

Author	Data	Information	Knowledge
Wiig, 1993	-	Facts organised to describe a situation or condition	Truths, beliefs, perspectives, judgements, know-how and methodologies
Spek & Spijkervet, 1997	Not yet interpreted symbols	Data with meaning	The ability to assign meaning
Davenport & Prusak, 1998	A set of discrete facts	A message meant to change the receiver's perception	Experience, values, insights, and contextual information
Quigley & Debons, 1999	Text that does not answer questions to particular problem	Text that answers the questions who, when, what, or where	Text that answers the questions why or how

siehe nächste Folie

Daten, Information, Wissen

(Davenport & Prusak, 1998)

Data:

Data is a set of discrete, objective facts about events. In an organizational context, data is most usefully described as structured records of transactions. Data by itself has little relevance or purpose.

Information:

Information is a message, usually in the form of a document or an audible or visible communication. Information is meant to change the way the receiver perceives something, to have an impact on his judgment and behavior. It must inform; it's data that makes a difference.

Knowledge:

Knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. In organizations, it often becomes embedded not only in documents or repositories but also in organizational routines, processes, practices, and norms.

Entspricht der semiotischen Sichtweise

Überblick über die weitere Vorlesung

- Dies waren viele Antworten auf die Frage: Was ist Wissen?
- Die Prozesse, in denen Wissen verarbeitet wird, sind das Thema von Kapitel B.
- Wie wird Wissen nun repräsentiert? Dies ist eine zentrale Frage der Forschungsrichtung **Künstliche Intelligenz** (Artificial Intelligence, KI, AI). In Kapitel C folgt eine kurze Geschichte der KI.
- In Kapitel D diskutieren wir verschiedene **Formalismen zur Repräsentation von Wissen**.
- In Kapitel E (welches den Großteil der Vorlesung einnimmt) betrachten wir die Anwendung von Wissensrepräsentation im **Semantic Web**, dem WWW der 2. Generation. Wir lernen unterschiedlich ausdrucksstarke Repräsentationen kennen: **XML, RDF, OWL (Ontology Web Language)**.
- In Kapitel F wird dann der formale Unterbau von OWL angesprochen: **Beschreibungslogiken**.