

# Semantic Integration

## Einführung

---

Carola Eschenbach, Özgür L. Özçep  
Universität Hamburg, Department Informatik  
AB Wissens- und Sprachverarbeitung (WSV)  
Sommersemester 2006

---

- **Einordnung in Studienplan**
- **Motivation: Semantic Web**
- **Die Lösung: Ontologien**
- **Die Aufgabe: Semantic Integration**  
Einführung und Vorstellung der Literatur
- **Organisatorisches**
- **Terminüberblick**

---

## Einordnung in Studienplan

---

### Profil Intelligente Systeme

- **Wissensverarbeitung**
- **Sprachverarbeitung**
- **(Bildverarbeitung)**

### Aktuelle Anwendungen

- **Semantic Web**
- **Datenbanken**

---

## Semantic Web (1)

---

Lit: T. Berners-Lee, J. Hendler, and O. Lassila: The semantic web. *Scientific American*, 2001. (<http://www.scientificamerican.com/>)

### Syntaktisches WWW: Sammelsurium an verlinkten Ressourcen

- Semantische Information bisher nur von Menschen erfassbar
- Bsp.: Link mit Aufschrift „curriculum vitae“
- Bsp.: Web-Page einer Universität: Nicht nur Seite mit Schlüsselwörtern wie „Geschäftszeit, Universität, Studienplan, Student“; sondern: Das Geschäftszimmer hat Montags, Donnerstags von 16-19 Uhr geöffnet und Frau so-und-so ist dort zu erreichen...
- Maschinen können Information nicht interpretieren / verstehen, daher auch komplexere Aufgaben, die das Erfassen der Bedeutung erfordern, nicht ausführen

## Semantic Web (2)

### Semantic Web: Erweiterung des WWW um semantische Information

- Maschinen werden zumindest in naher Zukunft nicht wirklich natürliche Sprache verstehen lernen können, aber ...
- sie können uns einen Teil der Verstehensleistung abnehmen, wenn wir ihnen das nötige Wissen durch semantische Annotation zur Verfügung stellen.
- Semantic Web = WWW (von heute) + semantische Annotation + KI

### Bsp.-Anwendung:

Beauftrage einen Software-Agenten, das Buch mit dem Titel „Handbook on Ontologies“ herausgegeben von Studer et al möglichst günstig in einem Online-Buchladen zu kaufen und an die eigene Adresse liefern zu lassen

## Semantic Web (3)

- Finde ein geeignetes Online-Shop und vollführe die Transaktion durch Angabe der vom Onlineshop geforderten Informationen
- Wie ist die Gesamtarchitektur zu gestalten?
- Was muss der Agent „wissen“, „kennen“, „verstehen“?
  - Finden eines geeigneten Onlineshops
  - Was ist ein Titel? Was ist ein Herausgeber? Was ist ein Handbuch
  - Herausgeber  $\Leftrightarrow$  Editor; Handbook  $\Leftrightarrow$  Handbuch? Adresse  $\Leftrightarrow$  shipTo?
  - Was heißt günstig: Im Preis Zollgebühren eingebunden? Porto?
  - Siehe die Lösung für dieses Szenario in Sitzung 9

### 9. Sitzung (30.05.06)

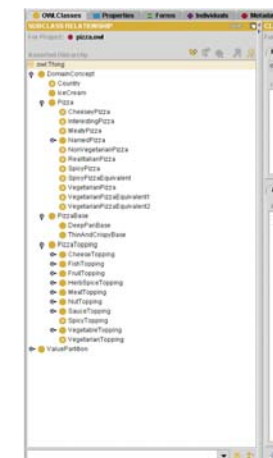
Mark H. Burstein & Drew V. McDermott: Ontology Translation for Interoperability Among Semantic Web Services 2005, *AI Magazine*, 26 (1), 71-82

## Ontologien: Begriff

- **Wie erreicht man das „Verstehen von Inhalten“ im Netz?**  
Nutze Ontologien, um die Bedeutung der semantischen Annotationen zu spezifizieren
- **Ontologiebegriff**
  - Philosophisch (sprachunabhängig): Welche Dinge gibt es da draußen und welche Struktur haben sie?
  - Beispiel: Gibt es neben raumzeitlichen Dingen auch Entitäten, die Eigenschaftsausdrücken entsprechen und in welchem Sinne existieren sie? (-> Universalienstreit)
  - Informatisch (sprachabhängig): Formale Beschreibung eines Anwendungsbereichs zum Zwecke des Gebrauchs durch verschiedene Anwendungen - verfasst in einer Sprache, die zum (automatischen) Schließen benutzt werden kann

## Ontologien: informatisch

Beispiel für eine informatische Ontologie: Pizza-Ontologie in OWL im Protege-Tool



## Tools für Ontologiekonstruktion

---

- 5. Sitzung (02.05.06)** Natalya Fridman Noy, Ray W. Ferguson, Mark A. Musen: The Knowledge Model of Protégé-2000: Combining Interoperability and Flexibility. In: *Rose Dieng, Olivier Corby (Eds.): Knowledge Acquisition, Modeling and Management, 12th International Conference, EKAW 2000, Juan-les-Pins, France, October 2-6, 2000, Proceedings*. Lecture Notes in Computer Science 1937 Springer 2000, 17-32
- In der Stanford University (Stanford Medical Informatics) entwickeltes, frei verfügbares (Registrierung erforderlich) Tool zum Editieren von Ontologien
  - Enthält Plugins zum erstellen von OWL-Ontologien und zum Mergen/Mapping
  - Framebasiert

## Webontologien

---

### Komponenten von Ontologien

- Terme zur Bezeichnung der relevanten Objekte, Konzepte in einer Domäne und ihrer Beziehungen
- Beschreibung der Bedeutung der Terme in der einen oder anderen Form
- Eventuell Mittel zur Beschreibung von Hintergrund-/weltwissen

### Webontologiesprache: OWL

- Entstanden aus DAML-OIL (das Produkt aus dem in Europa entwickelten OIL und in Amerika entwickelten DAML-ONT ist)
- Als Standard von W3C vorgegeben
- Logikbasiert (Beschreibungslogik)
- Syntax ähnelt RDF (Resource Description language)
  - Kann als Ontologiebeschreibungssprache aufgefasst werden
  - Graphenbasiert; Format: (Subjekt Prädikat Objekt)

## Sprachen für Ontologien: OWL und DL

---

### 3. Sitzung (18.04.06)

Ian Horrocks, Peter F. Patel-Schneider & Frank van Harmelen: From SHIQ and RDF to OWL: the making of a Web Ontology Language. *Journal of Web Semantics*, 2003, 1(1), 17-26 (.html)

Franz Baader, Ian Horrocks & Ulrike Sattler: Description Logics as Ontology Languages for the Semantic Web. *Mechanizing Mathematical Reasoning*, 2005, 228-248 (.pdf)

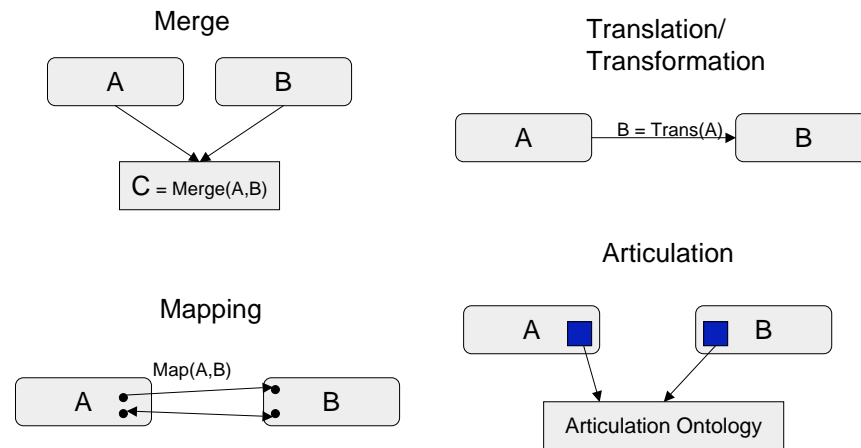
- Beschreibungslogiken als ideales Mittel zur Repräsentation von Ontologien
  - Sprechen über Konzepte und Rollen
  - Ausdrucksmächtig, dabei berechnungstechnisch handhabbar
- OWL (Ontology Web Language); Beziehung zu DL

## Die Lösung und die resultierende Problematik

---

- Es gibt nicht **die** Ontologie für denselben Anwendungsbereich, sondern viele verschiedene, die integriert werden müssen.
- **Integration:**
  - Ausmachen der Überlappungsbereiche der zu integrierenden Ontologien
  - Setze die Konzepte in den Ontologien in Beziehung (z.B. Äquivalenz, Subsumption)
  - Konsistenz- und Redundanzprüfung
- Das Ziel der Integration kann je nach Anwendung unterschiedlich sein: Mergen, Mapping, Articulation

## Semantic Integration: Terminologie



## Semantic Integration: Terminologie

### Wichtige Begriffe

- **Combining (Kombinieren):** Nutzung von zwei oder mehr Ontologien für eine Anwendung, bei der die Beziehung der Ontologien untereinander relevant ist.
- **Merging (Mischen):** Das Erstellen einer neuen Ontologie aus zwei sich überlappenden Ontologien
- **Aligning (Angleichen, Auf-eine-Linie-Bringen):** Zwei oder mehr Ontologien in Übereinstimmung bringen
- **Mapping:** Äquivalente Konzepte/Rollen aus zwei Ontologien aufeinander abbilden
- **Articulation:** Spezifikation der Übereinstimmungen zweier angeglichener Ontologien. (Betrifft in einer Lesart nur Teile der Ontologien)
- **Translating:** Repräsentationsformat einer Ontologie ändern
- **Transforming:** Semantik einer Ontologie leicht ändern

## Diskussion und Klärung der Begriffe

- 2. Sitzung (11.04.06):** Diskussion über die Begriffe und ihren Gebrauch in den jeweiligen Artikeln;  
erstes Lesen der Artikel!

## Semantic Integration: Mismatches auf Sprachebene

- Integration von Ontologien erschwert durch eventuelle Nichtübereinstimmung der Ontologien (mismatch) in der einen oder anderen Form
- Zwei Typen: sprachlicher und ontologischer Mismatch
- Mismatches auf sprachlicher Ebene
  - Syntax  
Bsp.: `<rdfs: Class ID = „Stuhl“>` vs. (defconcept Stuhl)
  - Logische Repräsentation: Ausdruck für „Disjunktheit“ vs. „Hilfskonstruktion mit Negation und Unterklassenbildung“
  - Die Semantik atomarer Ausdrücke: „A equalTo B“
  - Ausdrucksmächtigkeit der Sprache: Negation vorhanden?

## Semantic Integration: Mismatches auf Ontologieebene

---

- **Mismatch in der Konzeptualisierung**
  - Überdeckung: „Universitätsangestellte“ (auch technisches ode Verwaltungspersonal enthalten)
  - Granularität: Automobil-Ontologie: Modellierung von Trucks? Untertypen?
- **Mismatch in der Explikation**
  - Terminologie
    - Synonyme: „Automobil“ vs. „Wagen“
    - Homonyme: „Leiter“ in der Halbleiterphysik vs. Chorleiter
  - Modellierungsstil
    - Konzeptbeschreibung: Dissertation < Buch < wissenschaftliche Veröffentlichung < Veröffentlichung vs.
    - Dissertation < wissenschaftliches Buch < Buch < Veröffentlichung
    - Paradigmen: Intervallbasierte vs. punktbasierte Zeitrepräsentation
  - Codierung: Datumskodierung

## Semantic mappings

---

### Drei Hauptfragen

- Wie findet man semantic mappings?
- Wie stellt man sie dar?
- Was macht man mit ihnen?

## Semantic mappings: Upper ontologies

---

### Zwei unterschiedlich motivierte Ansätze

- Upper ontologies
- Heuristikbasierte Verfahren / maschinelle Lernverfahren

### Upper ontologies

- Vorgabe einer allgemeinen oberen Ontologie, die durch anwendungsspezifisch gestaltete Ontologien erweitert wird.
- Philosophischen Ontologiebegriff am nächsten
- Erleichtert das semantic mapping zwischen Ontologien mit gemeinsamer oberer Ontologie
- Im Seminar
  - SUMO
  - OntoMap
- Nicht behandelt: DOLCE, PSL, Cyc

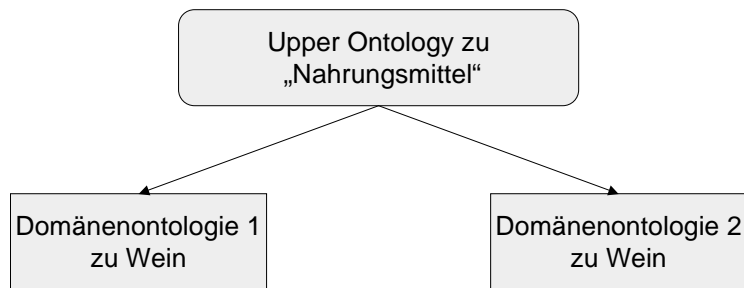
## SUMO & OntoMap

---

### 4. Sitzung (25.04.06)

- Ian Niles, Adam Pease: Towards a standard upper ontology. In: *Proceedings of the international conference on Formal Ontology in Information Systems - Volume 2001*, Ogunquit, Maine, USA, 2001, 2 - 9
  - Entwickelt vom IEEE Standard Upper Ontology Working Group
  - Entstanden als Merge verschiedener Ontologien
- Atanas Kiryakov, Kiril Iv. Simov & Marin Dimitrov: OntoMap: portal for upper-level ontologies. In: *Proceedings of the international conference on Formal Ontology in Information Systems - Volume 2001*, Ogunquit, Maine, USA, 2001, 47 - 58
  - S. <http://www.ontotext.com/>
  - Beschreibt Webportal, das Zugang zu verschiedenen Upper-ontologies bietet
  - Bietet Möglichkeiten, Upper-Ontologies in Beziehung zu setzen

## Bsp: Upperontologies



## Heuristische Verfahren

- Upper-Ontologies nicht immer verfügbar; Alternative: heuristisches Verfahren
- **Heuristische Verfahren nutzen**
  - Konzeptnamen und natürlichsprachliche Beschreibungen
  - Klassenhierarchie
  - Eigenschaftsdefinitionen
  - Instanzen von Klassen
  - Klassenbeschreibungen

## Prompt

### 8. Sitzung (23.05.06)

Noy, N.F. & Musen, M.A.: The PROMPT Suite: Interactive tools for ontology merging and mapping. In: *International Journal of Human-Computer Studies*, 2003, 59(6), 983 - 1024

- Framebasiertes semiautomatisches Tool zum Mergen und Mapping von Ontologien
- Als Plugin integriert in Protege
- Benutzt lexikalische und strukturelle Eigenschaften der Ontologien
- Unterstützt den Entwickler im Verarbeitungszyklus: (1) Anfangsvorschlag (durch Entwickler) -> (2) Wähle nächste Operation -> (3) Führe Updates durch -> (4) Konfliktfindung -> (5) Vorschläge für weiteres Verfahren -> gehe zu (2)
- ANCHORPROMPT ist ein darauf aufsetzendes Tool, das die strukturellen Eigenschaften der Ontologie in der Graphdarstellung ausnutzt

## Maschinelles Verfahren

### 11. Sitzung (20.06.06)

AnHai Doan, Jayant Madhavan, Pedro Domingos & Alon Y. Halevy: Learning to map between ontologies on the semantic web. In: *The Eleventh International WWW Conference Hawaii, US, 2002*, 662-673 (.pdf)

- Beschreibt die Konzeption des GLUE-Tools zum semantic mapping
- Basiert auf Verfahren des maschinellen Lernens
- Verwendet probabilistisches Modell, um zu einem Konzept einer Ontologie das ähnlichste in der anderen Ontologie zu finden
- Benutzt für verschiedenen Aspekte verschiedene Lerner, deren Ergebnisse von einem Met-Lerner zusammengefasst werden

## Formale Konzeptanalyse

### 12. Sitzung (27.06.06)

G. Stumme & A. Mädche: Ontology Merging for Federated Ontologies for the Semantic Web. In: E. Franconi, K. Barker, D. Calvanese (Eds.): *Proc. Intl. Workshop on Foundations of Models for Information Integration (FMII'01)*, Viterbo, Italy, Sept. 16-18, 2001. LNAI, Springer 2002

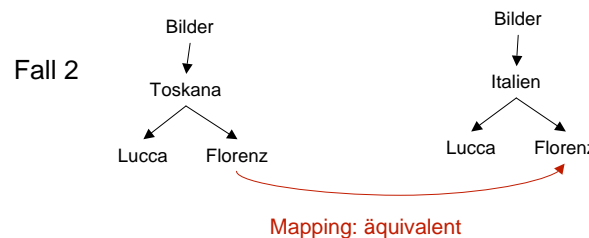
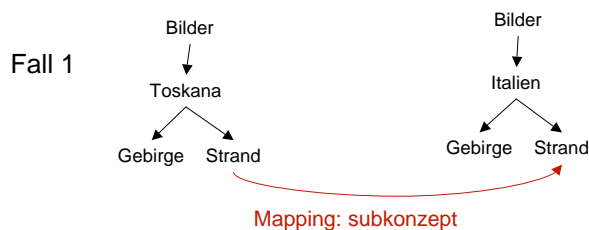
- Greift auf Formale Konzeptanalyse zurück (Konzeptbegriff hier ein anderer)
  - Mathematische Theorie von Begriffen verstanden als Tupel aus Begriffsexension und -intension.
- Input: Eine gemeinsame Menge von Textdokumenten für die aufeinander abzubildenden Ontologien
- Output: Konzeptverband mit Äquivalenz- und Subsumptionsbeziehungen
- Auf dem Konzeptverband aufbauend kann (nichtautomatisch) eine Ontologie konstruiert

## Sprachbasiertes Verfahren

### 13. Sitzung (4.07.06)

- P. Bouquet, L. Serafini, S. Zanobini & M. Benerecetti: An algorithm for semantic coordination. *Technical Report ????-??, IRST, 2003. Presented at the Semantic Integration workshop at ISWC-03 (International Semantic Web Conference) 2003.*
- P. Bouquet, L. Serafini & S. Zanobini. Semantic coordination: a new approach and an application. In: *Second International Semantic Web Conference*, volume 2870 of Lecture Notes in Computer Science, Springer Verlag, September 2003, 130-145.
- Sprachorientiertes Mapping-Tool, das auf WordNet zurückgreift
  - Benutzt SAT-basierten Theorembeweiser (Logik), um Beziehungen zwischen Konzepten verschiedener heterogener Wissensquellen (insbesondere Hierarchieklassen) zu errechnen

## Mehr als Struktur



In Fall 1 und 2 sind die Hierarchieklassen strukturell ähnlich, dennoch sind die mappings unterschiedlich

## Darstellung von semantic mappings

### Darstellung der mappings als

- Brückenaxiome (-> OntoMerge)
- durch Sichten (-> OIS)
- in einer eigenen Ontologie für mappings

### 6. Sitzung (09.05.06)

Dejing Dou, Drew McDermott & Peishen Qi: Ontology Translation on the Semantic Web. LNCS Journal of Data Semantics, Volume II, LNCS 3360, 2004, 35-57.

- Merging-Tool (Online-Tool OntoMerge)
- Mappings werden als Brückenaxiome (bridging axioms) (in einer prädikatenlogischen Sprache) formuliert und sind vom Entwickler vorzugeben
- OntoMerge ermöglicht auch die Übersetzung zwischen Domänenontologien, die semantisch aufeinander abgebildete Ontologien erweitern

### 7. Sitzung (16.05.06)

D. Calvanese, G. De Giacomo & M. Lenzerini: A framework for ontology integration. In: *Proc. of the First Semantic Web Working Symposium*, 2001, 303-316

- Semantic mappings werden als Anfragen/Sichten (Queries/Views) verstanden
- Lokale vs. globale Ontologie

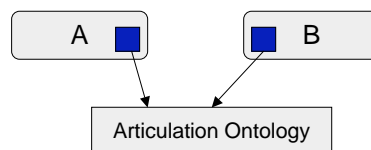
## Artikulation: eingeschränktes Mapping

---

### 10. Sitzung (13.06.05)

P. Mitra, G. Wiederhold & M. Kersten: A Graph-Oriented Model for Articulation of Ontology Interdependencies. In: *Proceedings Conference on Extending Database Technology 2000 (EDBT'2000)*, Konstanz, Germany, 2000.

- Allgemeine Algebra zur Komposition von Ontologien basierend auf Artikulationsregeln
- Statt global konsistenter Ontologie eine Artikulationsontologie



## Was ausgelassen wurde

---

- Weitere Anwendungen
- Schemaintegration (Datenbanken)

## Semantic Integration: Verwendete Literatur für den Themenüberblick

---

### Lit.:

- M. Klein: Combining and relating ontologies: an analysis of problems and solutions. In: IJCAI-2001, Workshop on ontologies and Information Sharing, 53-62, Seattle, WA, 2001
- N.F. Noy: Semantic Integration: A Survey of Ontology-Based Approaches. SIGMOD Record, Vol. 33, No. 4, December 2004, 65-70
- M. Uschold & M. Grüninger: Ontologies and Semantics for Seamless Connectivity. SIGMOD Record, Vol. 33, No. 4, December 2004, 58-64

## ... Der Studienführer

---

Seminare ... dienen der Erörterung ausgewählter wissenschaftlicher Probleme im Hauptstudium. Die Studierenden werden in der Arbeit nach wissenschaftlichen Grundsätzen und der Darstellung wissenschaftlicher Inhalte geschult. Sie erarbeiten selbständig die benötigte Literatur, gestalten einen Seminartermin durch Vortrag und Diskussion und liefern dazu eine schriftliche Zusammenfassung. Dabei ist die Zusammenarbeit in Kleingruppen von 2-3 Studierenden erlaubt. Eine Teilnahmebestätigung erhält, wer diese Leistungen erbringt und sich regelmäßig aktiv an den Diskussionen im Seminar beteiligt. Für einen Seminarschein ist darüber hinaus die Anfertigung einer Seminararbeit erforderlich. Zur Überprüfung der regelmäßigen Beteiligung kann eine Anwesenheitsliste geführt werden.

---

## Organisatorisches

---

## Vortragsvorbereitung

---

### Zur 2. Sitzung (11.04.06) Aufgabe bearbeiten:

- Übernommenen Aufsatz lesen unter der Fragestellung: Was wird im Aufsatz unter „Ontologie“ verstanden? Welche Form von Kombination heterogener Wissensquellen wird beschrieben? Was wird dabei unter „matching“, „bridging axiom“, „semantic map“, „merging“, „articulation“ bzw. „alignment“ etc. verstanden?

### rechtzeitig vor dem Seminartermin

- Lesen des Aufsatzes (zu finden neben weiterer Information unter [http://www.informatik.uni-hamburg.de/WSV/teaching/seminare/SemIntegr\\_SoSe06.php](http://www.informatik.uni-hamburg.de/WSV/teaching/seminare/SemIntegr_SoSe06.php))
  - Rücksprache mit den SeminarleiterInnen
    - zum Inhalt des Aufsatzes
    - zur Auswahl des zu präsentierenden Anteils
    - zum Aufbau des Vortrags
  - Klärung technischer Anforderungen und Kompatibilität
- ### im Vortrag
- Übersicht zum Vortrag, Einordnung in Gesamtzusammenhang, Begriffe

## Terminüberblick (1)

Sitzung	Datum	Thema	Literatur	Referent
1	04.04.06	Einführung	(siehe diesen Vortrag)	ÖLÖ
2	11.04.06	Diskussion	Jeweils die Aufsätze	alle
3	18.04.06	OWL und DL	Horrocks et al.(2003): From SHIQ and RDF to OWL: the making of a Web Ontology Language. Baader et al (2005): Description Logics as Ontology	
4	25.04.06	Upper Ontologies: SUMO und OntoMap	Kiryakov et al. (2001): OntoMap: portal for upper-level ontologies. Niles & Pease (2001): Towards a standard upper ontology.	
5	02.05.06	Ontologie-Tool	Noy et al (2000): The Knowledge Model of Protege-2000: Combining Interoperability and Flexibility.	
6	09.05.06	Briding-Axiome	Dou et al (2004): Ontology Translation on the Semantic Web.	

## Terminüberblick (2)

Sitzung	Datum	Thema	Literatur	Referent
7	16.05.06	OIS (Perspektivzugang)	Calvanese et al. (2001).: A framework for ontology integration. Calvanese et al. (2001): Ontology of Integration and Integration of Ontology.	
8	23.05.06	Mapping-Tool PROMPT	Noy et al. (2003).: The PROMPT Suite: Interactive tools for ontology merging and mapping.	
9	30.05.06	Anwendung/Architektur	Burstein & McDermott (2005): Ontology Translation for Interoperability Among Semantic Web Services.	
-	06.06.06	PFINGSTEN		
10	13.06.06	ONION-Tool (Ontologieartikulation)	Mitra et al. (2000): A Graph-Oriented Model for Articulation of Ontology Interdependencies.	
11	20.06.06	GLUE-Tool (Maschinelles Lernen)	Doan et al. (2002): Learning to map between ontologies on the semantic web.	

## Terminüberblick (3)

Sitzung	Datum	Thema	Literatur	Referent
12	27.06.06	Formale Konzeptanalyse	Stumme & Mädche (2001): Ontology Merging for Federated Ontologies for the Semantic Web.	
13	04.07.06	„Semantic coordination“ und WordNet	Bouquet et al. (2003): An algorithm for semantic coordination. Bouquet et al. (2003): A new approach and an application	
14	11.07.06	Resume und Ausblick	Oben aufgelistete Literatur	alle