



Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

Semantische Datenbanken

Abschlusspräsentation Bachelor Thesis

Theoretische Grundlagen; Aufbau und Nutzung einer
semantischen Datenbank

- ▶ Autoren: Mira Günzburger
Sven Osterwalder
- ▶ Betreuer: Prof. Dr. Jürgen Eckerle
- ▶ Experte: Jean-Marie Leclerc
- ▶ Datum: 16.01.2015

Herzlich Willkommen zu der Abschlusspräsentation unserer Bachelor Thesis. Das Thema der Thesis sind Semantische Datenbanken.

Den schwerpunkt haben wir dabei einerseitz auf die Analyse der theoretischen Grundlagen und andereseits auf den aufbau und die nutzung einer solchen semantischen Datenbank gelegt.

Wir haben uns ganz bewusst entschieden in dieser Präsentation auf die technischen Aspekte unsere Arbeit zu verzichten. Uns ist es wichtig euch in der folgenden viertelstunde einen Überblick über die Arbeit zu verschaffen. Und euch nicht mit Theorien und Techniken zu überschütten. Da es in dieser kurzen Zeit nicht möglich sein wird, diese ausführlich zu erklären. Bei Interesse für die Details sind wir ja am nachmittag an der Austellung für euch da.

Mira

Ablauf

- ▶ Motivation
- ▶ Wissensabbildung
- ▶ Ziel
 - ▶ Tutorial
 - ▶ Ontologie
 - ▶ Benutzerschnittstelle
- ▶ Fazit

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Nun einen kurzen überblick was euch in unserer Präsentation erwartet

- Motivation; also weshalb haben wir das Thema gewählt
- Wissensabbildung; wie ihr ja wisst gibt es verschiedene Arten der Wissensabbildung, uns ist es wichtig den unterschied hervorzuheben. Der Schwerpunkt liegt wie nicht anderst zu erwarten auf den semantischen Datenbanken
- Ziele: dann kommen wir schon konkret zu unserer Arbeit; wir möchten euch erklären was unsere Ziele waren, und auch konkret wie haben wir sie umgesetzt
- Fazit: zum schluss kommen wir schon zum Fazit, bei dem wir euch unsere wichtigsten Erfahrung mittelen werden

Mira

Motivation

- ▶ Beantworten von Fragen
 - ▶ Klassische Suchmaschine
 - ▶ Konzepte und Zusammenhänge → Semantisches Wissen

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Internet-Suchmaschinen sind heutzutage ein gängiges Mittel um an Wissen im Internet zu gelangen. Deren stetige Weiterentwicklung über das letzte Jahrzehnt macht sie zu einem mächtigen Instrument. Sie werden von vielen Personen zur täglichen Arbeit verwendet. Suchmaschinen sind heute meist so gehalten, dass in ein Suchfeld Begrifflichkeiten eingegeben werden. Suchmaschinen indexieren Inhalte mittels Stich- und Schlagworten. Deshalb muss der Suchende bereits eine mehr oder minder konkrete Vorstellung von den erwarteten Suchergebnissen haben.
Sven

Wissensabbildung

- ▶ Relationale Datenspeicherung
 - ▶ Objekteigenschaften und -verhalten eher schwer abbildbar
- ▶ Semantische Datenbanken
 - ▶ Wissensmodellierung auf Basis von Ontologien
- ▶ Expertensystem
 - ▶ Semantische Datenbank
 - ▶ Inferenzmaschine (Reasoner)
- ▶ Ontologie

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Der klassische Ansatz der Wissensabbildung, zum Beispiel in Form von UML, welchem die relationale Datenspeicherung zugrunde liegt, wird in der heutigen Informatik weitläufig eingesetzt und ist de facto Standard. Häufig geschieht dies in enger Verbindung mit der objektorientierten Programmierung. Experten aus einer Fachrichtung sind fähig diese Daten zu interpretieren und daraus Schlüsse zu ziehen. Es ist aber nicht möglich automatisch Fragestellungen zu beantworten, welche über reine Relationsverknüpfungen hinausgehen. Mit dieser Technik sind Objekteigenschaften und -Verhalten also eher schwer abbildbar. Eine andere Art Wissen zu repräsentieren sind semantische Datenbanken. Diese ermöglichen das Abbilden des Objektverhaltens und können mithilfe von Schlussfolgerungen die Rolle des Experten einnehmen.

Ontologien: eine formale Beschreibung des Wissens in einer Domäne in der Form von Konzepten der Domäne, deren Beziehung untereinander und der Eigenschaft dieser Konzepte und Beziehungen, sowie der in der Domäne gültigen Axiome und Prinzipien

Einleitung aus book: Semantische Datenbanken werden auf der Basis von Ontologien erstellt. Eine Ontologie beschreibt Sachwissen einer Wissens- bzw. Problem-Domäne. Sie

wird überall dort verwendet, wo Semantik zur Formulierung von Informationen benutzt wird.

Sven

Ziel

- ▶ Aufbau und Anwendung einer semantischen Datenbank
 - ▶ Theoretischer Teil
 - ▶ Tutorial
 - ▶ Praktischer Teil
 - ▶ Aufbau Ontologie
 - ▶ Benutzerschnittstelle

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

In dieser Bachelorthesis soll eine solche semantische Datenbank aufgebaut und angewendet werden. Die Arbeit wurde in zwei Teilen umgesetzt: Einem theoretischen und einem praktischen Teil. Der theoretische Teil zeigt in Form eines Tutorials auf, wie ein knowledge engineer bei der Wissensmodellierung vorgehen kann. Er nutzt dabei Ontologien als Basis, um eine semantische Datenbank aufzubauen. Im praktischen Teil soll eine solche Ontologie aufgebaut und per Benutzerschnittstelle zugänglich gemacht werden.
Mira

Tutorial

- ▶ Vorgehen Knowledge Engineer
 - ▶ Problemdomäne systematisch modellieren und formalisieren
- ▶ Aufbau
 - ▶ Theoretisches Hintergrundwissen zur Wissensmodellierung
 - ▶ Praktisches Beispiel Expertensystem
 - ▶ Gesammelte Erfahrungen

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Wie vorhin erwähnt, haben wir uns entschieden die theoretischen Grundlagen, welche wir erarbeiten mussten so abzubilden (niederzuschreiben) das sie auch für andere gut nutzbar sind. Aus diesem Grund haben wir uns für ein Dokument mit Tutorialcharakter entschieden. Im Tutorial zeigen wir auf wie wir eine Problemdomäne systematisch modellieren und formalisieren um eine Ontologie zu erhalten, welche danach in Form einer semantischen Datenbank abgespeichert wird.

Dabei war der Schwierigste Teil, die Überlegung wie wir den Aufbau vornehmen wollen. Ihr kennt sicher alle Tutorials, und wisst das dort das theoretische Hintergrundwissen eher spärliche vorhanden ist. Der Benutzer eines Tutorials will ja nur wissen, wie er vorgehen muss um sein Ziel zu erreichen. Da dies aber kein normales Tutorial ist, sondern in erster Linie eine Bachelor-Thesis betrachten wir die Wissensmodellierung aus drei Aspekten.

Einerseits erhält der Benutzer fundamentales Hintergrundwissen (welches aber für die Umsetzung nur bedingt notwendig ist)

Andererseits erklären wir anhand eines praktischen Beispiels das konkret vorgehen. Dem Benutzer ist es also möglich, diesem speziell gekennzeichneten Teil zu folgen, und ist so fähig eine semantische Datenbank aufzubauen und zu nutzen. Als dritten Aspekt haben wir uns entschieden unsere praktischen Erfahrungen in Form von Tips einzufliessen zu lassen.

Mira

Ontologie

- ▶ Problemdomäne Reisen
 - ▶ Ursprünglich Prolog
 - ▶ Klarer Rahmen
- ▶ Werkzeuge
 - ▶ Stanbol
 - ▶ Protégé
 - ▶ Stardog

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Wie Herr Osterwalder bereits erwähnt hat, werden semantischen Datenbanken auf der Basis von Ontologien erstellt. Eine Ontologie beschreibt sachwissen einer Problemedomäne, also eines klar definierten Ausschnitt der Welt.

Beim wählen der Problemedomäne sind wir schon auf die ersten Schweierigkeiten gestossen. Unsere Ursprünliche Gewählte Domäne, «das Erlernen der Programmierung am Beispiel von Prolog» hat uns vor grosse Probleme gestellt. Während der Modellierung haben wir es nicht geschafthen, den erwarteten Mehrwert von der Wissensmodellerung mittels Ontologien zu generien. Nach vielen gescheiterten Versuchen wurde uns bewusst, dass es wichtig ist eine Domäne zu wählen ,welche schlussfolgerungen mittels Inferenz erlaubt. Da dies ja ein wichtiger Teil ist um wissen mit semantik (also Bedeutung) zu versehen. Die Ursprünlich gewählte Problemedomäne hatten wir auf einer zu hohen Abstraktionsebene angesetzt.

-> Ontologien und Expertensysteme machen überall dort sin, wo ein Experte benötigt wird um das wissen zu Interpretieren. So kammen wir auf die idee der Domänen der Reiseplanung

Sven (oh sorry, jetzt hani das alles ufgschriebe und den gseh, dass das eigendli dini Folie isch! ;-)) i los jetzt dine, vlicht kasch jo öpis demit avo... ;)

Aufbau Ontologie

- ▶ Tripel:
 - ▶ Subjekt
 - ▶ Prädikat
 - ▶ Objekt
- ▶ Bestandteile
 - ▶ Klassen
 - ▶ Individuen
 - ▶ Eigenschaften
 - ▶ Beziehungen
 - ▶ Regeln

→ Semantische Datenbank

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Jetzt haben wir euch erklärt weshalb wir die Domäneneisepnaer gewählt haben und welche werkzeuge wir dabei verwendet haben. Aber was heisst es denn eigendlich «eine Ontologie zu Modellieren»? Eine Ontologie hat immer die Form von Tripeln. Diese beinhalten Subjekt, Prädikat und Objekt. Was ja sehr an einen simplen deutschen Satz erinnert und auch so verstanden werden kann.

Ontologien bestehen aus Klassen, Individuen, Eigenschaften, Beziehungen und Regeln. Eine Klasse stellt eine «Kategorie» dar (zb. Ausflugsziel). Das Individuum ein Konkretes Objekt einer Klasse (Seilpark). Eigenschaften weisen einem Individuum einen Wert zu (zb anzahl Teilnehmer) Beziehungen bestehen zwischen zwei Individuen (zb Seilpark hatStandort Balmberg . Und setzten diese in verbindung miteinander. Zum Schluss sind noch die Regeln. Regeln sagen aus, diese bringen erst den mehrwert von Ontologien richtig zur Geltung. Wenn wir zum Beispiel sagen, das Balmberg die Region Solothurn hat. Und dann die Regel festlegen, das ein Ausflug, welcher einen Standort hat, welcher eine Region hat, auch diese Region hat.... ;-) dann können wir daraus schliessen das Seilpark die Region Solothurn hat. Dies ist ein sehr simples beispiel um euch zu veranschaulichen wie Ontologien funktionieren. Ich möchte aber hier nicht weiter ins Detail gehen.

Solche Ontologien werden in Owl einer Ontologie abbildungssprache abgebildet, diese hat eine XML ähnliche schreibweise und kann so Hierarchies strukturen sehr gut abbilden (es gibt klassen, supklassen, Propertys und supprobertys...) Speicherung als RDF/XML in OWL Dateien

Mira

Ontologie abbilden

The screenshot shows the reiseplaner ontology editor interface. The top navigation bar includes tabs for Active Ontology, Entities, Classes, Object Properties, Data Properties, Annotation Properties, Individuals, OWLViz, DL Query, OntoGraf, and S. The main area is divided into several panels:

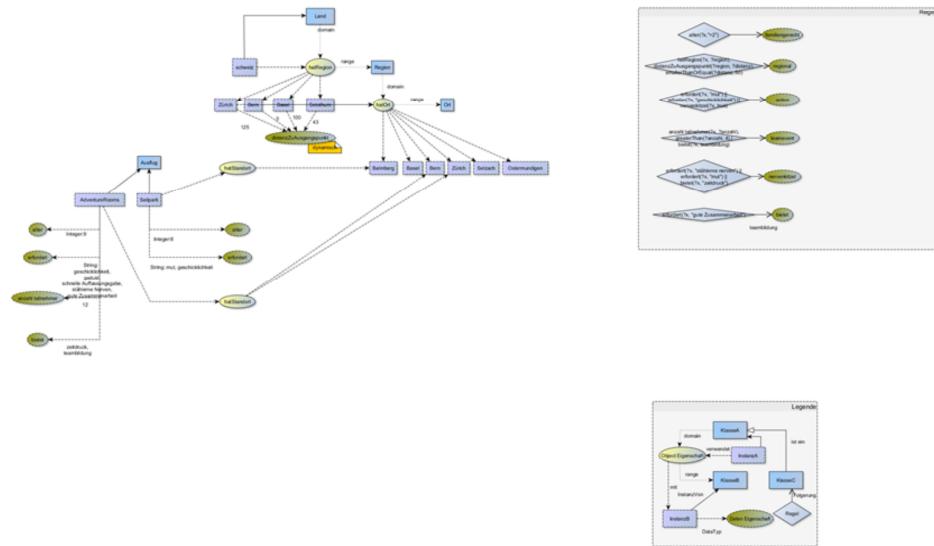
- Class hierarchy:** Shows a tree structure under "Gourmet". Root node: Thing. Subnodes include Ausflug, Jahreszeit, Land, Ort, Preissegment, Region, Restaurant, Gourmet, Landgasthaus, and Speziell.
- Annotations:** Shows annotations for the "Gourmet" class, including "Annotations: Gourmet" and "Description: Gourmet".
- Object property hierarchy:** Shows a tree structure under "Object property hierarchy". Root node: topObjectProperty. Subnode: gleicheRegion.
- Members:** Lists individuals: Bären_Ersigen and Signori.

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Wie Herr Osterwalder bereits erwähnt hat, haben wir zur Modellierung Protege verwendet. (unpraktisch owl zu schreiben). Auf diesem Bild sieht Ihr ein Ausschnitt unserer Modellierung mit den Klassen und Subklassen, gewiesen (in diesem Fall ermittelten Individuen) und einem Teil der Property's

Mira

Semantisches Netz

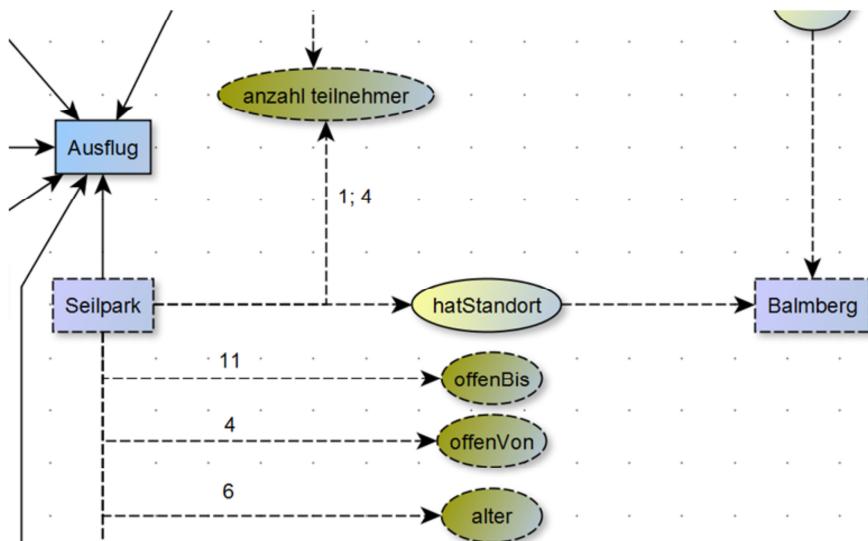


Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Wir die Gestaltung und die Modellerung war es für uns aber einfacher in semantischen netzten zu arbeiten. Oder besser gesagt einer Abgewandelten Form davon. Dies ist eine Graphische Darstellung von wissen
Speicherung Ontologie als semantisches Netz (bzw. so ähnlich)

Mira

Semantisches Netz Ausschnitt



Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Hier noch einen kurzen Blick auf einen auschnitt des Netzes, wo man das vorherige Beispiel grad abgebildet sieht...

Benutzerschnittstelle

- ▶ Technische Umsetzung
 - ▶ Backend
 - ▶ Graphdatenbank (Stardog)
 - ▶ Reasoner
 - ▶ REST-Schnittstelle
 - ▶ Frontend
 - ▶ Schritt-für-Schritt Assistent

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Sven

Assistent

OWL Reiseplaner Start Über Impressum

Schritt 2

Welche Kriterien soll dein Ausflug erfüllen?

Ausflug

Eigenschaften

- tages**
- kinder**
- kleinkinder**
- halbtags**
- romantik**
- teamevent**

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Fazit

- ▶ Viele neue Lerninhalte
- ▶ Umdenken
- ▶ Wechsel Problemdomäne
- ▶ Mächtig aber doch mit gewissen Einschränkungen
 - ▶ Beschränkt intelligent
 - ▶ Werkzeuge

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences

Beschränkt Intelligent: nur mit unseren Regeln intelligent; Nur simple Inferenz vorgegeben

Werkzeuge: Fehler/ Bastlig -> nur mit Kombination von zwei Werkzeugen nutzbar



?

Berner Fachhochschule | Haute école spécialisée bernoise | Bern University of Applied Sciences