

***„Ein Java-Rahmenwerk
zur kooperativen, anreizbasierten Erstellung
von Ontologien“***

Diplomarbeit

von

Thomas Much

Vortrag am 23.06.2006

Verantwortlicher Betreuer: Prof. Dr.-Ing. K. Böhm

Betreuender Mitarbeiter: Dipl.-Wirtsch.-Inf. Conny Kühne

Motivation (1)

The image shows a screenshot of the German Wikipedia homepage in a web browser window. The browser's address bar shows the URL <http://de.wikipedia.org/wiki/Hauptseite>. The page features the Wikipedia logo on the left, a navigation menu, and a search bar. The main content area is titled "Hauptseite" and includes a welcome message, a list of featured articles by theme, and a section for "Artikel des Tages" (Article of the Day) featuring the Argentine flag. A sidebar on the right contains a "In den Nachrichten" (In the news) section with several news items.

Hauptseite – Wikipedia

URL: <http://de.wikipedia.org/wiki/Hauptseite>

Navigation: [Artikel](#) [Diskussion](#) [Quelltext betrachten](#) [Versionen/Autoren](#) [Anmelden](#)

Hauptseite

Willkommen bei Wikipedia

Die Wikipedia ist ein Projekt zum Aufbau einer freien Enzyklopädie in mehr als 200 Sprachen. Jeder kann mit seinem Wissen beitragen. Seit Mai 2001 entstanden so 418.007 Artikel in deutscher Sprache. Gute Autorinnen und Autoren sind stets **willkommen**.

[Geographie](#) [Geschichte](#) [Gesellschaft](#) [Kunst](#) [Religion](#) [Sport](#)
[Technik](#) [Wissenschaft](#)

[Artikel nach Themen](#) · [Alphabetischer Index](#) · [Artikel nach Kategorien](#)

[Kontakt](#) · [Presse](#) · [Statistik](#) · [Andere Sprachen](#) · [Andere Ausgaben](#)

Artikel des Tages

 Die **Flagge Argentinien**s besteht aus drei gleich großen, horizontalen Streifen: oben hellblau (spanisch: *celeste*, deutsch „himmelblau“), in der Mitte weiß und unten wieder hellblau. In der Mitte der Flagge ist eine Sonne mit abwechselnd 16 geraden und 16 geflammten Sonnenstrahlen dargestellt. Es ist überliefert, dass Manuel Belgrano zu diesem Symbol inspiriert wurde, als er während der Schlacht von Paraná in den Himmel schaute. Das Sonnensymbol (*Sol de Mayo* = Maisonette; aber auch Inkasonne genannt) tauchte

In den Nachrichten

- Ehegattensplitting · Fußball-WM · Konflikt in Somalia
- Die **Internationale Walfangkommission** verabschiedete eine Erklärung, in der zum Ende des **Walfangverbotes** aufgerufen wurde. 
- Der Unternehmer **Ronald Lauder** erwirbt für 106,7 Mio. Euro das *Porträt der Adele Bloch-Bauer* von **Gustav Klimt**. Damit ist es das teuerste Gemälde der Welt.
- Die Bevölkerung **Kataloniens** hat dem neuen

Motivation (2)

Wikipedia:

- + Wissen wird kooperativ gesammelt
 - freiwillige Mitarbeit motivierter Autoren

Problem:

- Bestimmung und Sicherung der Qualität
 - Manipulation, Vandalismus und Unwissenheit verschlechtern die Qualität

Ein Lösungsansatz zur Qualitätsermittlung:

- ⇒ Bewertungen durch die Nutzergemeinschaft
- ⇒ Bereits bei Online-Händlern und Informationsportalen im Einsatz

Motivation (3)

More Music (Score:4, Funny)
by [Metabolife \(961249\)](#) on Saturday April 22, @ 11:41AM (#15180658)

Great, a 10GB would let you have 2,500 songs (according to apple's rating), this is great business for apple. Now instead of just getting 1,000 dollars from people to fill it, they can bump it up to 2,500 dollars! Brilliant!

[[Reply to This](#)]

o [Re:More Music](#) by Midnight Thunder (Score:2) Saturday April 22, @ 12:15PM

Rumor v.s. obvious. (Score:1, Redundant)
by [blanks \(108019\)](#) on Saturday April 22, @ 11:44AM (#15180672)
(<http://www.public-internet.net/> | Last Journal: [Friday October 14, @04:35PM](#))

Naturally as larger (smaller) hard drives become cheaper they will come up with ipods with more space, this isnt a rumor that they will be doing it, maybe a rumor that it will be happening soon.

Alle Kundenrezensionen

Durchschnittliche Kundenbewertung: ★★★★★

15 von 18 Kunden fanden die folgende Rezension hilfreich:

★★★★★ **Wirklich ein Bestseller**, 23. März 2006

Rezensentin/Rezensent: ["office15"](#) - [alle meine Rezensionen ansehen](#)

Wenn man oft laut lachen und gleichzeitig erkennen möchte, wieviel man schon von seiner eigenen Sprache verlernt oder sich falschem Deutsch, das von allen Seiten auf einen niederprasselt, angepasst hat, dann sollte man unbedingt dieses Buch lesen. Es ist einfach hervorragend. Ein lehrreicher Lese Genuss.

War diese Rezension für Sie hilfreich? JA NEIN ([Rezension nicht akzeptabel?](#))

29. März 2006 10:18

Danke Hauke!

JustMyCent (462 Beiträge seit 30.1.03)

Danke für das wiki.
Musste mal gesagt werden, die Idee war nämlich klasse!
Hat mir als Auswärtigen schon einige Male geholfen.
Das sollte es für jede Stadt geben.

Bewertung dieses Beitrags: ||||| ■

Beitrag bewerten:

[Erläuterung zum Bewertungssystem](#)

☐ **Danke Hauke!**

| Re: Danke Hauke!

| Re: Danke Hauke!

|||| ■ **JustMyCent**

Hauke Löffler

Wete

29.03.06 10:18

29.03.06 10:23

29.03.06 10:24

- Kann man Bewertungen auch zur Qualitätsermittlung von strukturierten Datenbeständen bzw. Ontologien nutzen?

Motivation (4)

Consensus Builder

Home | My Account | Recruit new user | Logout | Search | Administration | About... ckuehne (Conny Kuehne) · 6.00 Points

Topic-Type: Land

Topics				
Name	A	D	Dup	Actions
Italien	3	0	0	Agree Disagree Duplicate
Deutschland	3	0	0	Agree Disagree Duplicate
USA	2	1	0	Agree Disagree Duplicate
Belgien	1	0	0	Agree Disagree Duplicate
Germany	1	0	2	Agree Disagree Duplicate
				New topic
Help · Top				

Relationship-Types	
Relationship-Type	Topic-Type
enthält	Stadt
Help · Top	

User Top-10	
jörg	128.25 Points
bubble	116.00 Points
Mikey2	31.00 Points
Aminem	20.00 Points
nisu	20.00 Points
kBoehm	10.00 Points
ckuehne	6.00 Points
Anger	1.50 Points
bob_morris	1.00 Points

A service created by Jörg Schönebaum and Prof. Klemens Böhm
Professorship 'Data and Knowledge Engineering' at University of Magdeburg, Germany
Time needed to display this page: 0.172 sec

- + Hat die Nutzbarkeit von Bewertungen für strukturierte Datenbestände gezeigt
- Typ-Schema vom Administrator fest vorgegeben
- PHP-Prototyp für einen speziellen Anwendungsfall
 - ⇒ schlecht wartbar, erweiterbar und wiederverwendbar

Motivation (5)

Die Vorlesungsvideo Community

Sie befinden sich in
Thema: [Auditory system](#)
Segment: Opening

Details zu Segment "Opening"


[Abspielen](#)

Informationen über dieses Segment

Info	Wert
Vorlesung	Multimodal and multilingual dialogue systems
Ort	Language technologies seminar
Dozent	Professor Alex Waibel
Videodetails	zum ganzen Video
Erkannter Text	

Bewertbare Attribute

Info	Wert
Segmentname	Opening
Gehört zu Thema	Auditory system
Beginn	00:00:54
Ende	00:1:34
Segmentlänge	0:00:40
Schlagwörter	
Bewerten	Akzeptieren Ablehnen Kommentar Schlagwörter

[zurück](#)

Sie sind eingeloggt als: **kuehne** Erzeugungspunkte: **10** Stimmpunkte: **1000** Benutzergruppe: **Newbie**

Start
[alle Videos ansehen](#)
User Top10
Erfahren Sie wer die aktivsten Benutzer sind
Bisherige Stimmen
Schauen Sie Ihre bisherigen Stimmen an oder ändern Sie diese
neueste Stimmen ansehen
Übersicht der zuletzt im System eingegangenen Stimmen
Benutzergruppen
Benutzer nach Gruppen geordnet
Hilfe
[Logout](#)

- musste von Grund auf neu entwickelt werden, obwohl die Grundidee starke Ähnlichkeiten zum Consensus Builder aufweist

Aufgabenstellung

- Entwicklung eines Java-Rahmenwerks zur kooperativen Erstellung von Ontologien auf Grundlage von Bewertungen
 - ⇒ Basis für Systeme wie Consensus Builder oder die Video Community
 - ⇒ kein Prototyp
- austauschbares Anreizsystem (Punktevergabe, Rechte)
 - ⇒ um unterschiedliche Anreizmechanismen erforschen zu können
- flexible Protokollierung (Logging)
 - ⇒ um die Reaktionen der Benutzer auf unterschiedliche Anreize untersuchen zu können

- Zusätzlich Realisierung einer Beispiel-Anwendung (Prototyp) auf Basis des Rahmenwerks.

Gliederung

1. Fachliche Analyse
(Wissensrepräsentation, Bewertungen)
2. Entwurf
3. Implementierung
4. Auswertung
5. Fragen

Fachliche Analyse

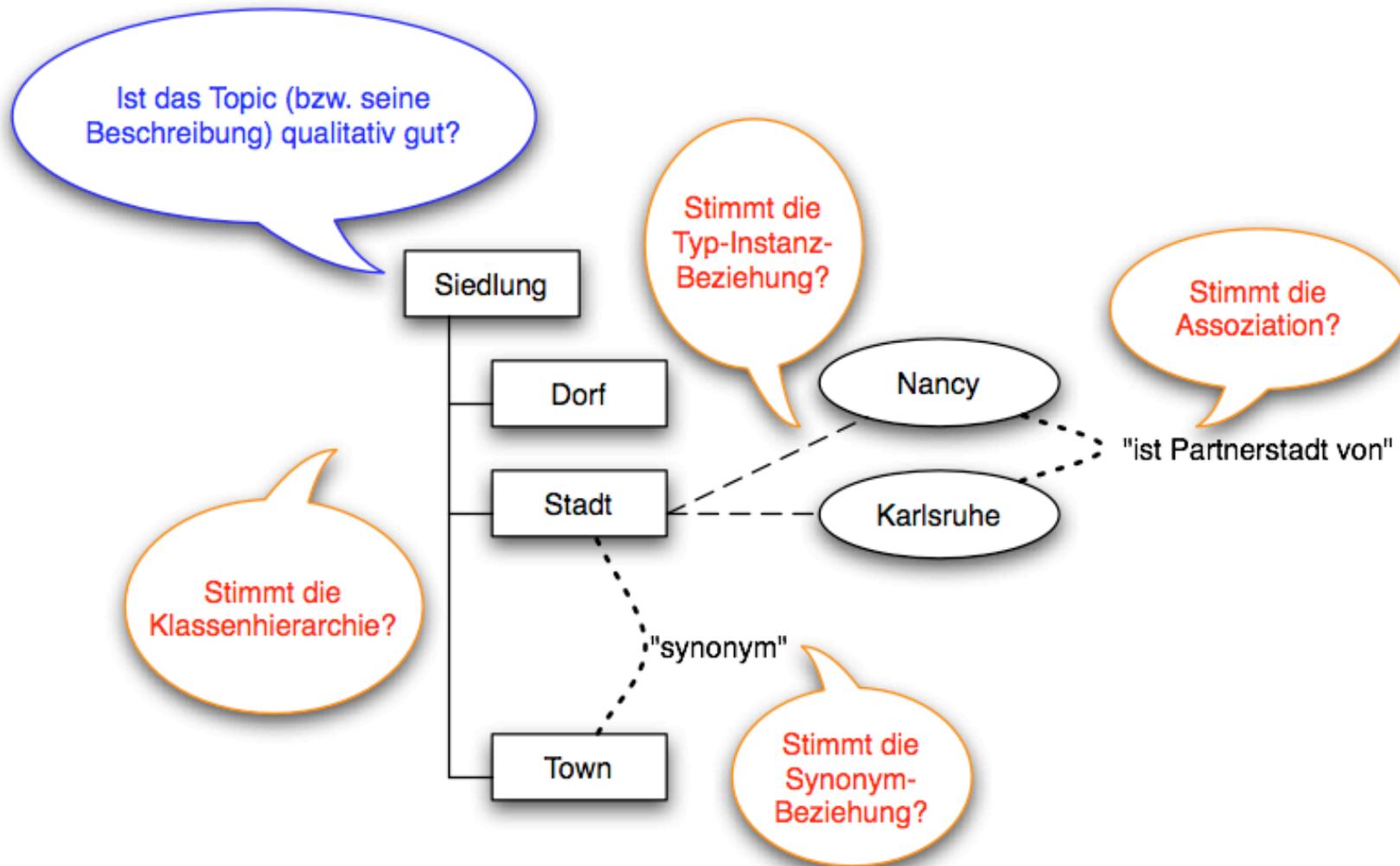
Wissensrepräsentation

- Formalisierung nötig, um Wissen (Konzepte, Begriffe) zu speichern
- Diverse Formen:
 - Thesaurus
 - Taxonomie / Klassifikation
 - Semantisches Netz
 - **Ontologie**

Ontologien

- *„Eine Ontologie ist eine explizite Spezifikation einer Konzeptualisierung“* (T. Gruber)
 - ⇒ Konzeptualisierung beschreibt Ausschnitt der Welt
 - ⇒ Ontologie definiert gemeinsames Vokabular dafür
 - ⇒ Nutzer verwenden das gemeinsame Vokabular korrekt („Ontological Commitment“)
- Ontologie $O := \langle C, R, I, A \rangle$
mit
 - Menge C von Konzepten (Typen bzw. Klassen)
 - Menge R von Relationen (z.B. „ist ein“, „ist Unterklasse von“)
 - Menge I von Instanzen der Elemente aus C
 - Menge A von Axiomen
- Wie kann man ein gemeinsames Vokabular kooperativ aufbauen?

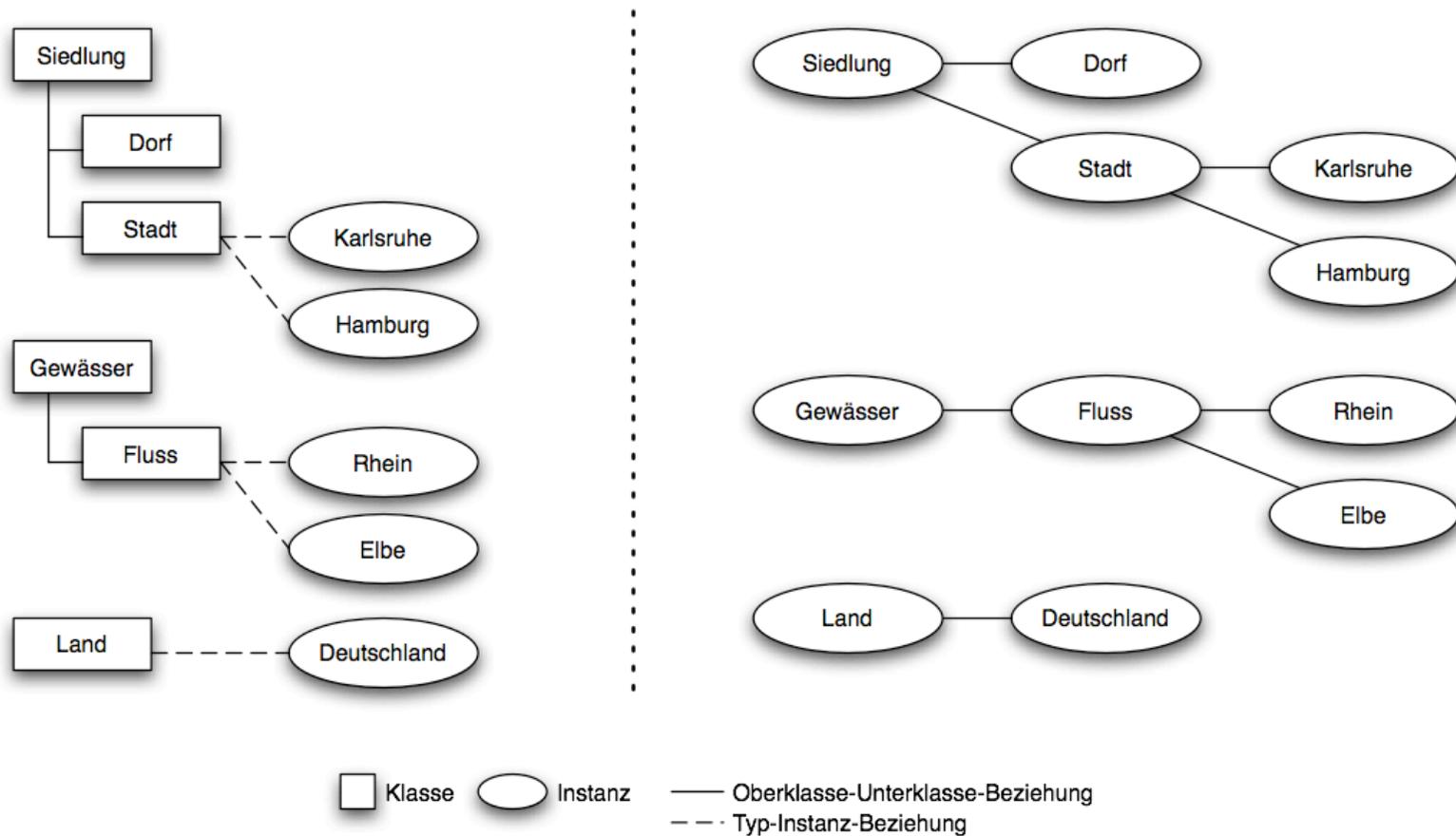
Bewertungen



Repräsentationsformen für Ontologien

- RDFS (RDF Schema)
- OWL (Web Ontology Language)
 - W3C-Sprachen für das semantische Web
 - ⇒ für die maschinelle Wissensverarbeitung geeignet
- **Topic Maps**
 - ISO-Standard
 - nicht so mächtig wie RDFS oder OWL
 - formales Datenmodell vorhanden, aber nicht so stark formalisiert wie RDFS bzw. OWL
 - dafür wesentlich leichter zu nutzen
 - ⇒ für die menschliche Wissensverarbeitung geeignet

Klassifikation



- Trennung Instanz \leftrightarrow Unterklasse selbst für Experten nicht immer klar entscheidbar
- Anforderung daher: Neues Rahmenwerk soll beide Arten der Klassifikation unterstützen.

Anforderungen

Folgende Module wurden als notwendig identifiziert:

- Ontologieverwaltung
- Benutzerverwaltung
 - ⇒ beides zum kooperativen Erstellen der Ontologie
- Bewertungs- und Anreizsystem
 - ⇒ damit Nutzer Bewertungen vergeben und sie durch Anreize belohnt (motiviert) werden können
- dynamische Rechtevergabe
 - ⇒ zum Durchsetzen der Anreize
- Protokollierung (Logging)
 - ⇒ zur Auswertung, wie sich unterschiedliche Anreizmechanismen auf das Verhalten der Nutzer auswirken

Entwurf

Entwurf

- Warum ist ein neues Rahmenwerk nötig?
 - ⇒ Vollständige Neuentwicklungen für unterschiedliche Forschungszwecke sollen entfallen
 - ⇒ **Wiederverwendbarer Rahmen** soll vorgegeben werden, nur die anwendungsspezifischen **Erweiterungen** müssen dann neu entwickelt werden
 - ⇒ Rahmenwerke (Frameworks) bieten dies: Ein halbfertiges Architekturgerüst, das für eine konkrete Anwendung angepasst werden kann.
- Voraussetzungen für Wiederverwendbarkeit und Erweiterbarkeit:
Lose Kopplung und *hohe Kohäsion* der Module
 - wohldefinierte Schnittstellen
 - klare Zuständigkeiten
 - ⇒ Modularität

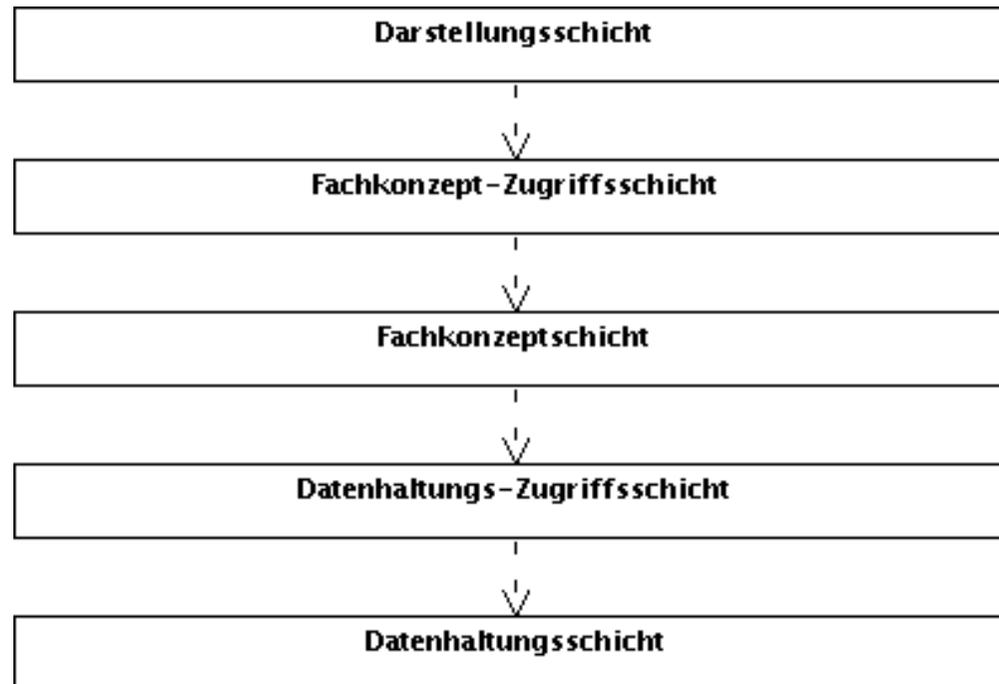
Rahmenwerke (Frameworks)

- Kehren den Kontrollfluss um (Inversion of Control)
 - ⇒ besitzen vollständige oder zumindest teilweise Ablaufhoheit
- Formen von Rahmenwerken:
 - **objektorientiertes Rahmenwerk**
 - ⇒ besteht aus konkreten Klassen
 - ⇒ Anpassung erfolgt durch Spezialisierung der Klassen
 - **komponentenbasiertes Rahmenwerk**
 - ⇒ Schnittstellen und deren Interaktionen werden spezifiziert
 - Mischformen
- Das zu entwickelnde Rahmenwerk
 - wird zunächst komponentenbasiert entworfen
 - soll aber auch als Standardimplementierung vorliegen
 - ⇒ Mischform

Einsatz des Rahmenwerks

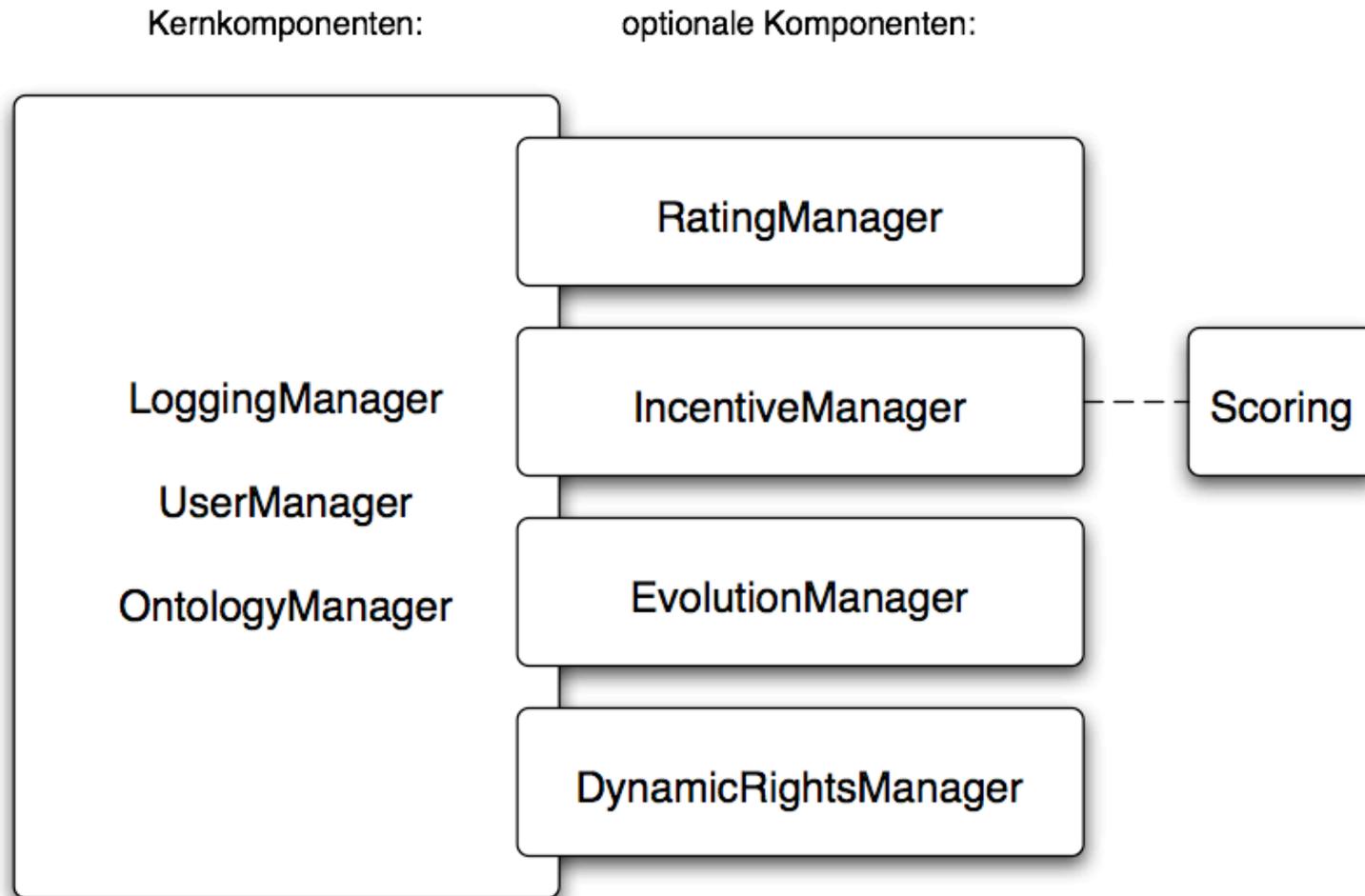
- typischerweise in einer Web-Applikation
 - ⇒ kooperatives Arbeiten auf einem zentralen Web-Server
- zusammen mit anderen Rahmenwerken, die sich z.B. um Darstellung und Steuerung kümmern
 - ⇒ neues Rahmenwerk ist nur im Fachkonzept angesiedelt
- Darstellung, Steuerung, Fachkonzept sollen modular austauschbar sein
 - ⇒ Mehrschicht-Architektur

Mehrschicht-Architektur



Implementierung

Kernkomponenten und optionale Komponenten



ConsensusFoundation

```
ConsensusFoundation cf;  
cf = new ConsensusFoundation( konfigurationsdatei );
```

```
cf.loggingmanager=de.uka.ipd.consensus.impl.LoggingManagerImpl  
cf.usermanager=de.uka.ipd.consensus.impl.UserManagerImpl  
cf.ontologymanager=de.uka.ipd.consensus.impl.OntologyManagerImpl  
cf.evolutionmanager=de.uka.ipd.consensus.impl.EvolutionManagerImpl  
cf.ratingmanager=de.uka.ipd.consensus.impl.SQLRatingManagerImpl  
cf.incentivemanager=de.uka.ipd.consensus.impl.IncentiveManagerImpl  
cf.dynamicrightsmanager=de.uka.ipd.consensus.impl.DynamicRights...  
  
cf.scoringlistener=de.snailshell.cf.MyScoringListenerImpl, ...  
  
cf.loglevel=debug
```

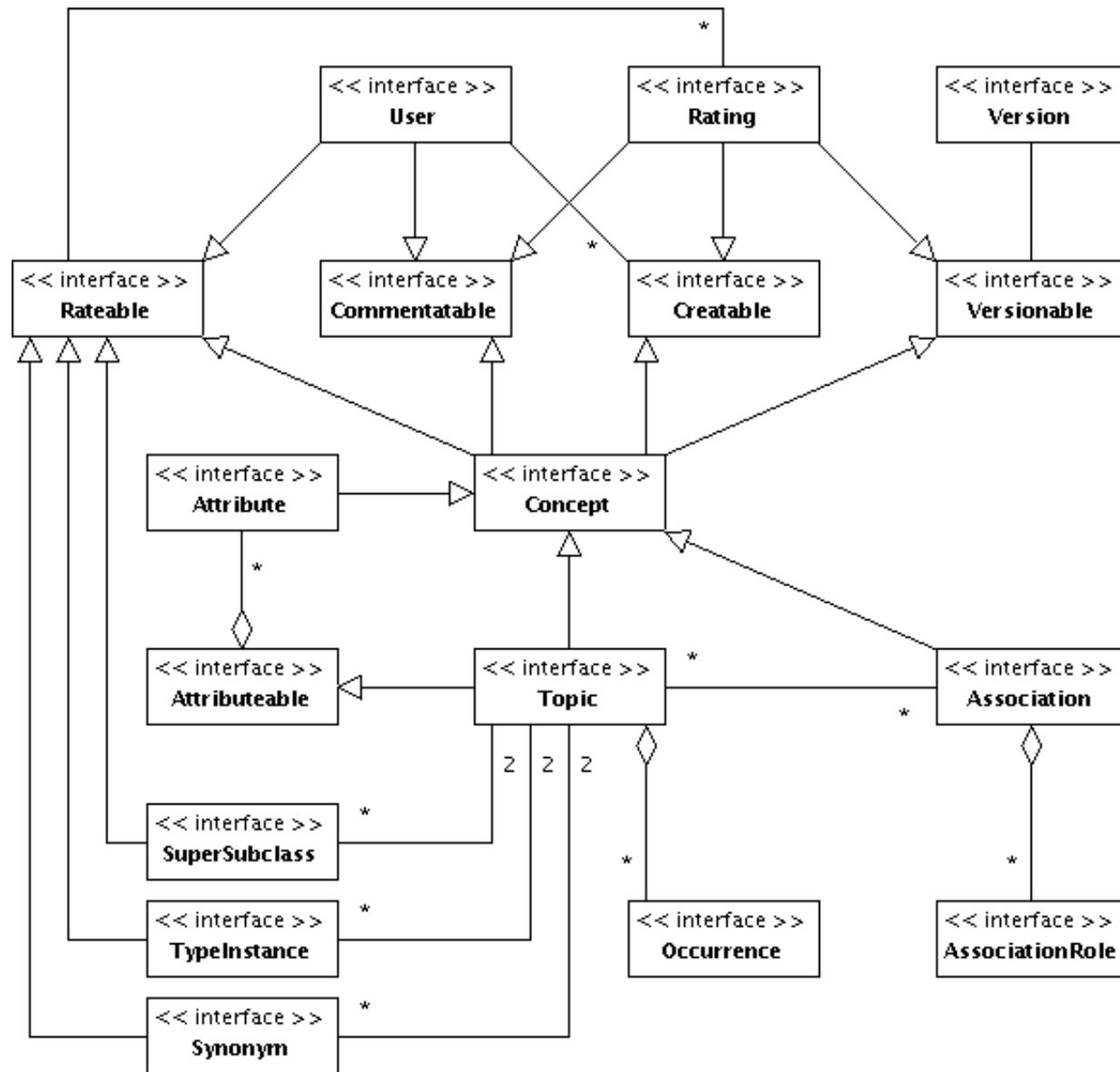
Punktevergabe

```
public class MyScoringListenerImpl implements ScoringListener {  
  
    public void score(ScoringEvent event) {  
        User  usr  = event.getUser();  
        double punkte = usr.getScore();  
  
        switch( event.getId() ) {  
            case ScoringEvent.TOPIC_CREATED:  
                punkte += 2.0;  
                break;  
  
            case ScoringEvent.CONCEPT_RATING_SET:  
            case ScoringEvent.CONCEPT_RATING_CHANGED:  
                Rating r = event.getRating();  
                if (r.getValue() != Rating.NEUTRAL) {  
                    punkte += 1.0;  
                }  
                break;  
        }  
  
        usr.setScore( punkte );  
    }  
  
    public void scored(ScoringEvent event) { }  
}
```

**Punktevergabe
(Scoring)**

Anreizberechnung

Datenmodell (Schema)



Auswertung der Bewertungen

```
Topic t    = ...;
User  usr  = ...;

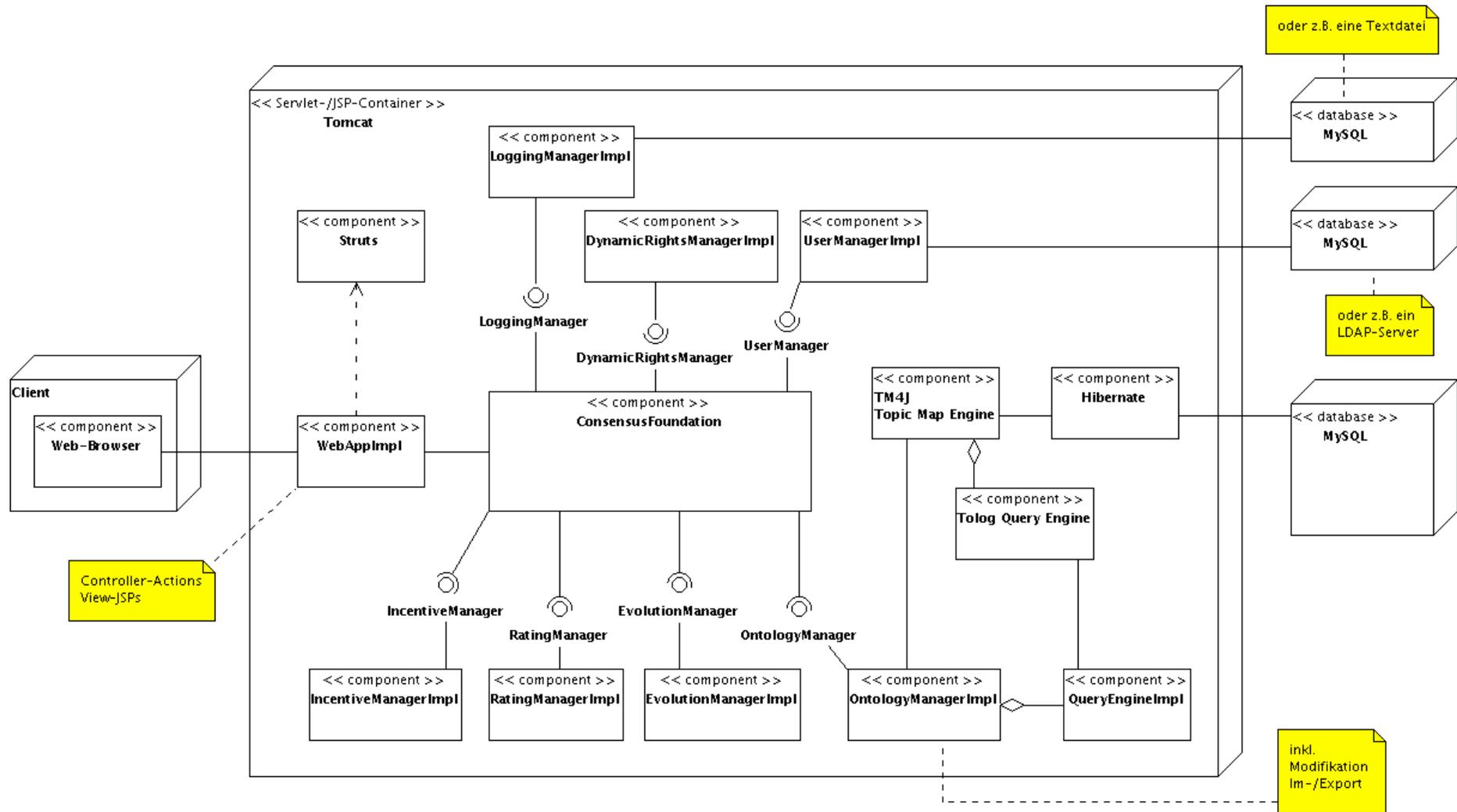
Collection alle           = t.getRatings();
Rating  bewertung        = t.getRating( usr );

int      anzahl          = t.getRatingsCount();
double   durchschnitt    = t.getAverageRatingValue();

Version  verRating       = bewertung.getVersion();
Version  verBewertetesTopic = bewertung.getRatedVersion();

Collection historie      = bewertung.getHistory();
```

Systemübersicht (Web-Applikation)



Auswertung

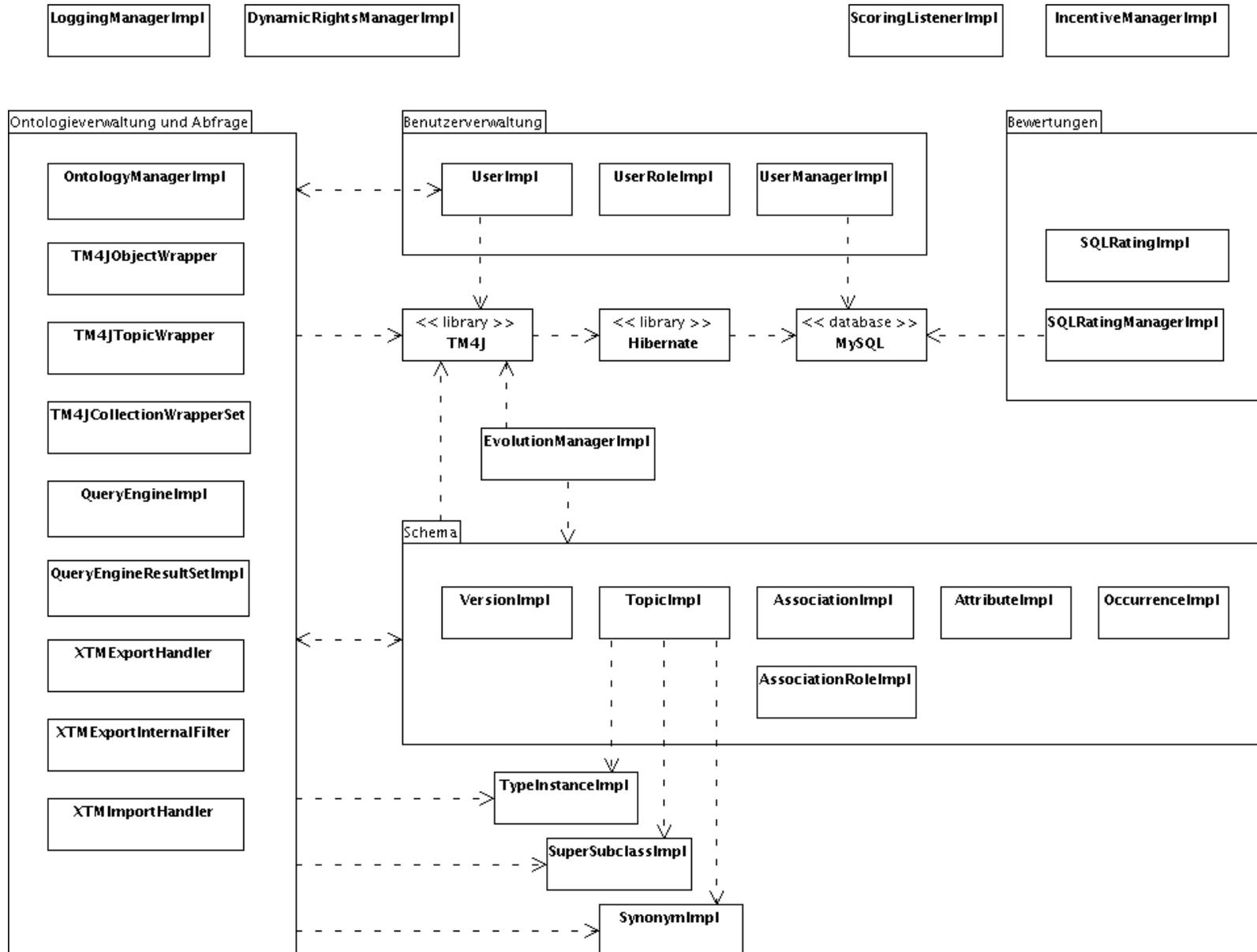
Auswertung

- Abhängigkeiten
- Architektur
- Vorteile von ConsensusFoundation

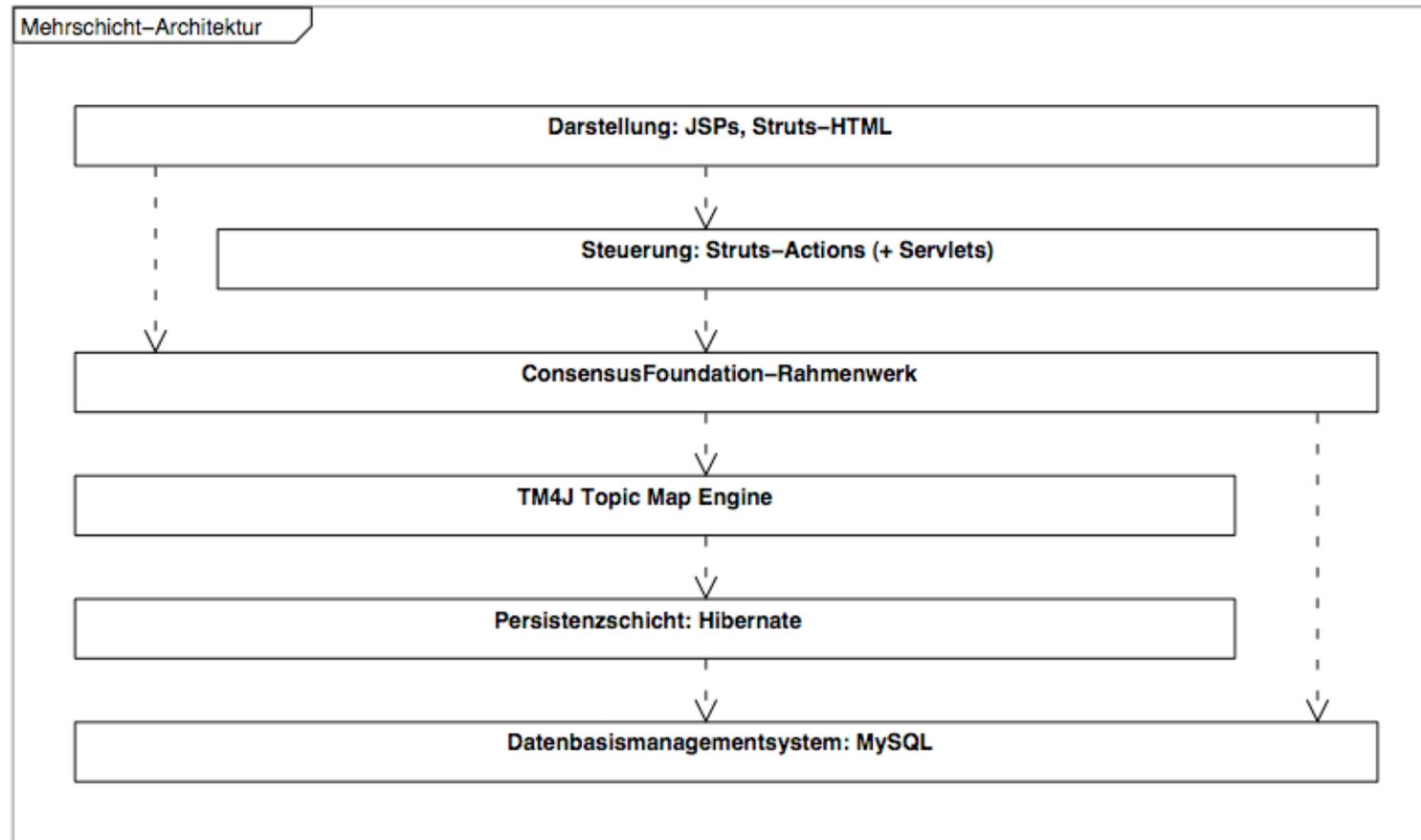
Abhängigkeiten (1)

- die Module werden unterschiedlich häufig ausgetauscht
- häufig ausgetauschte Module müssen lose gekoppelt sein
 - ⇒ wenige (idealerweise keine) Abhängigkeiten zur konkreten Implementierung
 - ⇒ Punktevergabe (Scoring), Protokollierung (Logging) und dyn. Rechtvergabe werden zur Erforschung unterschiedlicher Anreizmechanismen sehr häufig ausgetauscht werden
- selten ausgetauschte Module dürfen Abhängigkeiten zugunsten besserer Performanz aufweisen

Abhängigkeiten (2)



Architektur



- Schichtenmodell mit strikter Ordnung
- lineare Ordnung zugunsten der Performanz nicht umgesetzt

Vorteile von ConsensusFoundation

- gegenüber dem Consensus Builder
 - + wartbares, erweiterbares, wiederverwendbares Rahmenwerk
- gegenüber TM4J
 - + größere Funktionalität
(Benutzerverwaltung, Bewertungen, Anreize)
 - + trotzdem einfacher einsetzbar
- allgemein
 - + dynamische Rechtevergabe
 - + flexible Protokollierung
 - + zwei Arten der Klassifikation nutzbar

Demo & Fragen

ConsensusFoundation Demo
 Universität Karlsruhe (TH)
 Forschungsuniversität - gegründet 1825

Startseite - Topic-Liste - Topic-Hierarchie - Tolog-Suche - Meine Daten - Abmelden
 Admin (Benutzer: Thomas Much, angemeldet seit 11:25:52 11.05.2006; Punkte: 12.597.00)

Hierarchie der Topic-Instanzen

- Assoziationsstyp
 - entspringt in
 - ist Partnerstadt von
 - liegt an (der) / am
 - liegt in
- Einheit
 - Ganzzahl (0)
 - Einwohnerzahl (-1)
- Fahrzeug
- Figur
 - Dreieck
 - Viereck
 - Quadrat
- Fluss
 - Rhein
 - Elbe
- Land
 - Deutschland
 - Du
 - Disneyland
- Stadt
 - Entenhausen
 - Hamburg
 - Karlsruhe (1)
 - Nancy
- Town
- Universität
- Zahl
 - Ganzzahl (1)
 - Einwohnerzahl (-1)
 - Fliesstkommazahl

ConsensusFoundation Demo

Startseite - Topic-Liste - Topic-Hierarchie - Tolog-Suche - Meine Daten - Abmelden
 Admin (Benutzer: Thom)

Tolog-Suche

Im folgenden Formular können Sie eine Suchanfrage in Tolog 1.0-Syntax formulieren:

Vordefinierte Abfrage in das Eingabefeld übernehmen?

Suchanfrage:

```
select $A,$R,$P from
association($A),association-role($A,$R),role-player($R,$P)
```

Ausführen

Ergebnisse: 192

Nr	SA	SR	SP
1	Association "ConsensusFoundation created-by", id=x1m3p377m6-af	AssociationRole "ConsensusFoundation user account"	User "thmuch" (Much, Thomas, thomas@snailshell.de, login Sun May 21 17:01:39 CEST 2006)
2	Association "ConsensusFoundation created-by", id=x1m3p377m6-af	AssociationRole (untyped)	Topic (unnamed), id=x1m3p377m6-ae
3	Association "entspringt in", id=x1m3p377m6-aa	AssociationRole "ConsensusFoundation rating"	Topic (unnamed), id=x1m3p377m6-ae
4	Association "entspringt in", id=x1m3p377m6-aa	AssociationRole "ConsensusFoundation created-by"	User "thmuch" (Much, Thomas, thomas@snailshell.de, login Sun May 21 17:01:39 CEST 2006)
5	Association "entspringt in", id=x1m3p377m6-aa	AssociationRole "Fluss"	Topic "Rhein", id=x1m24p10eq-18
6	Association "entspringt in", id=x1m3p377m6-aa	AssociationRole "Land"	Topic "Deutschland", id=x1m24p10eq-1c

Topic "Stadt"

Angelegt von Ihnen selbst!
 Version "Sun Jun 11 22:29:49 CEST 2006"

Bewertungen: 1 (1)

Synonyme:
 • [Town](#) (1)

Beschreibung:
 Eine Stadt (von althochdt.: stat = Standort, Stelle; etymologisch eins mit Statt, Stätte; vgl. dagegen Staat) ist eine größere, zentralisierte, abgegrenzte Siedlung mit einer eigenen Verwaltungs- und Versorgungsstruktur im Schnittpunkt größerer Verkehrswe

Typen dieses Topics:
 • [Siedlung](#)

Attribute:
 Dieses Topic besitzt noch keine Attribute.

Assoziationen, in denen dieses Topic eine Rolle spielt:
 Dieses Topic spielt in keiner Assoziation eine Rolle.

Instanzen dieses Topics:
 • [Entenhausen](#) (-1)
 • [Hamburg](#)
 • [Karlsruhe](#) (1)
 • [Nancy](#) (0)

Occurrences
 • "<http://de.wikipedia.org/wiki/Stadt>" (löschen)

⇒ Download von <<http://www.snailshell.de/uni/dipl/>>

Vielen Dank für die Aufmerksamkeit! :-)

