

Softwaretechnikpraktikum 05

1 Recherchebericht

1. Begriffe:

(a) Semantisches Web

Domäne Unter einer Anwendungsdomäne, häufig auch verkürzend Domäne, versteht man in der Informatik und insbesondere in der Softwaretechnik ein abgrenzbares Problemfeld des täglichen Lebens oder - etwas spezieller - einen bestimmten Einsatzbereich für Computersysteme oder Software. Anwendungsdomänen stellen typischerweise sehr spezielle Anforderungen an ein technisches System, welches zur Bewältigung der domänenspezifischen Aufgaben und Probleme eingesetzt werden soll.

Metadata Als Metadaten oder Metainformationen bezeichnet man im Allgemeinen Daten, die Informationen über andere Daten enthalten. Bei den beschriebenen Daten handelt es sich oft um größere Datensammlungen (Dokumente) wie Bücher, Datenbanken oder Dateien. Eine allgemeingültige Unterscheidung zwischen Metadaten und normalen Daten existiert allerdings nicht. So werden auch Angaben von Eigenschaften eines Objektes (beispielsweise Personennamen) als Metadaten bezeichnet. Während der Begriff „Metadaten“ relativ neu ist, ist sein Prinzip unter anderem jahrhundertlang bibliothekarische Praxis.

Ontologie Unter einer Ontologie versteht man in der Informatik im Bereich der Wissensrepräsentation ein formal definiertes System von Begriffen und/oder Konzepten und Relationen zwischen diesen Begriffen. Zusätzlich enthalten Ontologien - zumindest implizit - Regeln.

Semantisches Web ist eine Erweiterung des gegenwärtigen World Wide Webs, durch die Informationen wohldefinierte Bedeutungen bekommen, so dass Computer und Menschen besser zusammenarbeiten können.

Wissensbasis In einem Expertensystem oder Wissensbasierten System ist die Wissensbasis der Bereich des Systems, der das Fachwissen in einer beliebigen Repräsentationsform enthält (zB. in Form von \rightarrow Ontologien).

(b) OWL / RDF

Attribut Das Merkmal einer Klasse, das entweder öffentlich zugänglich oder in der Klasse gekapselt ist. Attribute selbst können auch wieder komplexe Datentypen oder Klassen sein.

Entität Als Entität werden in der Informatik unterscheidbare, in der realen Welt eindeutig identifizierbare Objekte bezeichnet. Die Objekte können sowohl eine physische („real“) oder eine konzeptionelle („abstrakte“) Existenz haben.

Instanz eine abgeschlossene Einheit mit einer bestimmten Funktion : eine Ausprägung eines Objektes.

OWL (Web Ontology Language) ist eine Spezifikation des W3C, um Ontologien anhand einer formalen Beschreibungssprache erstellen, publizieren und verteilen zu können. Es geht darum, Terme einer Domäne und deren Beziehungen formal so zu beschreiben, dass auch Software (Agenten) die Bedeutung verarbeiten kann. OWL ist somit ein wesentlicher Bestandteil der Semantic Web-Initiative von Tim Berners-Lee. OWL basiert technisch auf der RDF-Syntax und historisch auf DAML+OIL, und geht dabei über die Ausdrucksmächtigkeit von RDF-Schema weit hinaus. Zusätzlich zu RDF und RDF-Schema werden weitere Sprachkonstrukte eingeführt, die es erlauben, Ausdrücke ähnlich der Prädikatenlogik zu formulieren.

RDF (Resource Description Framework) ist eine Spezifikation für ein Modell zur Repräsentation von Metadaten (Informationen über Webseiten und andere Objekte). In Verbindung mit RDF-Schema und der Web Ontology Language (OWL) soll RDF als grundlegendes Format zur Repräsentation von Taxonomien und Ontologien - also formalen Vokabularen im Allgemeinen - dienen.

XML (Extensible Markup Language) ist eine Metasprache für das Definieren von Dokumenttypen. Anders gesagt: XML liefert die Regeln, die beim Definieren von Dokumenttypen angewendet werden.

(c) Protégé 2000

Protégé ist ein Programm, das Benutzern erlaubt, Domains der Ontologie zu konstruieren, Dateieingabe einzustellen und Daten einzugeben. Es ist auch eine Plattform, die leicht erweitert werden kann, um graphische Bestandteile wie Diagramme und Tabellen, Mittel wie Ton, Bilder und Bildschirm, und verschiedene Speicherformate wie OWL, RDF, XML und HTML einzuschließen.

Plug-in ist ein Softwareprodukt zur Erweiterung von Software, die als Rahmenapplikation bezeichnet wird.

(d) Kundengespräch

Benchmark Ist ein Standard, der für Vergleiche verwendet wird. Es ist genormtes Testverfahren, mit dessen Hilfe man die Leistung von EDV-Systemen oder Systemklassen ermitteln und diese nach bestimmten Kriterien miteinander vergleichen kann.

BibTeX ist ein Programm zur Erstellung von Literaturangaben und -verzeichnissen in TeX- oder LaTeX-Dokumenten, entwickelt von Oren Patashnik.

Client (engl. „Kunde“) ist eine Anwendung, die in einem Netzwerk den Dienst eines Servers in Anspruch nimmt. Man spricht dann vom Client-Server-Prinzip.

Concurrent Versions System (CVS) bezeichnet ein Programm zur Versionsverwaltung von Quellcode. CVS vereinfacht die Verwaltung von Quellcode dadurch, dass es alle Dateien eines Software-Projektes an einer zentralen Stelle, einem so genannten Repository, speichert. Somit hilft es dabei, einen Überblick über die einzelnen Versionen der Dateien und die dazugehörigen Kommentare zu behalten. Es ist zudem auch noch begrenzt dazu in der Lage, besonders bei größeren Projekten, die Arbeit der einzelnen Entwickler eines Projektes zu koordinieren.

Datenserver/Datenbank ist die elektronische Form eines Karteikastens. Es handelt sich um eine Sammlung von Daten, die aus der Sicht des Benutzers zusammengehören, z. B. eine Personal-datenbank oder eine Lagerinventardatenbank.

Symbolic Data statistische Einheiten gekennzeichnet durch die mehrfachen Werte, die Veränderlichkeit ausdrücken, und durch die Verwandtschaftsstrukturen (logische Abhängigkeiten, taxonomies), definiert auf den Variablen.

Webserver ist im engeren Sinne ein Server-Dienst, welcher Dateien bzw. Daten über das HTTP-Protokoll zur Verfügung stellt, die sich über eine HTTP-URL adressieren lassen. Umgangssprachlich versteht man unter Webserver aber auch einen physisch vorhandenen Server, auf dem ein entsprechender HTTP-Dienst läuft.

2. Konzepte:

- Semantisches Web: Das semantische Web ist eine Erweiterung des World Wide Web. Zusätzlich zu der vom Menschen lesbaren Form liegen die im Semantischen Web enthaltenen Informationen auch formal in einer für Maschinen interpretierbaren Form vor.
- OWL (Web Ontology Language): OWL ist eine formale Beschreibungssprache um Ontologien zu erstellen. Sie ermöglicht die Beschreibung der Terminologien und Beziehungen innerhalb eines Wissensbereiches.

- Protégé 2000: Protégé ist ein Ontologie-Editor zur Repräsentation von Wissensbereichen. Protégé beinhaltet ein Plugin-Konzept, welches es über verschiedene Schnittstellen ermöglicht Anpassungen und Erweiterungen unterschiedlicher Art vorzunehmen:
 - (a) Tab Widget Plugin: Visualisierung der Informationen
 - (b) Slot Widget plugin: Gestaltung der slot widgets und deren Anpassung an Datentypen
 - (c) Back-end plugin: interne Datenverwaltung
 - (d) Import/Export plugin: import/export verschiedener Dateiformate
 - (e) Project Plugin: Manipulation an Projekten

3. Beschreibung der Rahmenapplikation:

- Quelle: <http://protege.stanford.edu/>
- Leistungsmerkmale:

Protege ist ein Werkzeug zum Erstellen von Ontologien, oder Wissensbasen, konform zum Open Knowledge Base Connectivity Protokoll (OKBC). Dabei wird eine einheitliche GUI für alle Arbeitsschritte und Darstellungen zur Verfügung gestellt. Mit Protege ist unter anderem Folgendes möglich:

 - (a) Modellierung von Ontologien über bestimmte Gebiete (domain)
 - (b) Erstellung von Werkzeugen zur Wissensgewinnung
 - (c) Erstellung von speziellen Instanzen, deren Daten und damit Wissensbasen
 - (d) Ausführung von Applikationen
- Use-Cases:
 - (a) Import von verschiedenen Datensammlungen über Bibliothekarische Bestände vom XML-Format (Bibltem.xsd)
 - (b) Import von Benchmarks im Symbolischen Rechnen („SD“)
 - (c) Verwaltung von Bibliotheksbeständen, Editieren von Büchern oder Zeitschriften, Organisation von Ausleihe.
 - (d) Literaturrecherche durch erweiterte Suchfunktionen (Stichwort-, Schlagwort-, Volltext-, Autorensuche)
 - (e) Darstellung von Klassen von Büchern (zB: Bücher von Autoren, die mit „A“ anfangen) durch Filterwerkzeuge.
- Äußerlich sichtbare Aspekte der inneren Logik
Zur Erstellung von Ontologien gibt es das Classes-Tab. Hier können Klassenhierarchien modelliert werden. Im Slots-Tab können den

Klassen entweder andere Klassen oder Datenstrukturen (Strings, Integer, Float...) zugeordnet werden. Diese Zuordnungen heissen Properties. Im Instances-Tab können Instanzen von Klassen mit den in Slots definierten Properties editiert werden. Im Forms-Tab kann man Formulare zur Einagabe im Instances-Tab modifizieren. Im Queries-Tab werden verschieden Such-Möglichkeiten angeboten.





