

Semantic Web für Menschen wie du und ich



benutzerfreundliche Anwendung semantischer Technologien

Tobias Kuhn

Institut für Informatik
Universität Zürich

16. KnowTalk

1. Februar 2011

KnowGravity Inc., Zürich

Einleitung

In diesem Vortrag geht es um die Kombination zweier Themen:

- ① Semantic Web und semantische Technologien im Allgemeinen
- ② Intuitive und natürliche Benutzerschnittstellen

Behauptung:

Punkt ② ist eine Voraussetzung für die erfolgreiche Anwendung von Punkt ①.

Was versteht man unter dem “Semantic Web”?

“Most of the Web's content today is designed for humans to read, not for computer programs to manipulate meaningfully”

“The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning”

(Tim Berners-Lee und andere, 2001)

“This simple idea, however, remains largely unrealized”

(Tim Berners-Lee und andere, 2006)

Semantic Web: Beispiel

Beispiel mit Bezug zum letzten KnowTalk zu [Produktpaketen](#):

- Mehrere Unternehmen bieten ähnliche Produktpakete an
- Web-Seiten mit Informationen zu den Produktpaketen

Wie verhalten sich Anbieter und Kunde?

- **Heute:**

- **Der Anbieter** definiert seine Angebote und bewirbt sie auf seinen Web-Seiten
- **Der Kunde** muss die Web-Seiten der Anbieter durchforsten

- **Semantic Web:**

- **Der Anbieter** definiert sein Angebot aufgrund dynamischer Regeln, und veröffentlicht das Angebot in standardisierter und computer-interpretierbarer Form
- **Der Kunde** benutzt ein Programm, das die passenden Angebote automatisch findet, vergleicht, und evt. sogar kauft

Semantic-Web-Technologien

Eine ganze Reihe von Technologien wurde in den letzten Jahren für das Semantic Web entwickelt und angewandt:

- Sprachen für die Definition von Fakten und Regeln (z.B. RDF, OWL, SWRL, RuleML, SBVR, Prolog)
- Anfragesprachen (z.B. SPARQL)
- “Reasoner” und Regelsysteme für automatisches Schlussfolgern und Ausführen von Regeln (z.B. FaCT++, RacerPro, OntoBroker, Pellet)
- Ontologie-Editoren (z.B. Protégé, NeOn)

Diese Technologien können auch ausserhalb des Semantic Webs sehr nützlich sein.

Ontologien

Die Semantic-Web-Technologien ermöglichen die Definition, Verwaltung und Anwendung von **Ontologien**.

Was sind Ontologien und wozu sind sie nützlich?

Eine Ontologie ...

- ... ist eine Weiterentwicklung eines kontrollierten Vokabulars
- ... definiert in formaler Weise die Objekte **und deren Relationen** eines gegebenen Anwendungsgebiets
- ... soll einen höheren Grad an Automation ermöglichen (z.B. die automatische Ausführung von Regeln)

Formale Sprachen versus natürliche Sprachen

Problem:

- **Computer verwenden formale Sprachen** (Programmiersprachen, Anfragesprachen, Logiksprachen etc.), die schwierig anzuwenden sind für Menschen ohne fundierte Ausbildung
- **Menschen verwenden natürliche Sprachen** (Deutsch, Englisch etc.), die von Computern nur sehr unzuverlässig interpretiert werden können

Lösung: Verwendung von **kontrollierter natürlicher Sprache** (CNL)

- sieht wie natürliche Sprache aus
- enthält Einschränkungen, die die automatische Interpretation durch Computer erlauben

Beispiel

Beispiel aus dem letzten KnowTalk zu Produktpaketen:

Alle Produktangebote, die im Duo Pack enthalten sind, sind auch im Trio Pack enthalten.

In einer Semantic-Web-Sprache (OWL in Manchester-Syntax):

```
product_offer and some (inverse contain) Duo_Pack  
SubClassOf some (inverse contain) Trio_Pack
```

In einer kontrollierten natürlichen Sprache (ACE):

The Trio Pack contains every product offer that the Duo Pack contains.



- AceWiki ist ein Wiki (wie z.B. Wikipedia)
- Artikel sind in ACE geschrieben
- Der Inhalt wird automatisch in eine Semantic-Web-Sprache (OWL) übersetzt
- Ein Semantic-Web-Reasoner (FaCT++) ist integriert:
 - Konsistenzprüfung
 - Beantwortung von Fragen (formuliert in ACE)
 - Extraktion von Wort-Hierarchien
- Spezieller Editor um syntaktisch korrekte ACE Sätze zu erstellen

AceWiki-Demo

The screenshot shows a web browser window with the URL `attempto.ifi.uzh.ch/webapps/acewikigeo/`. The page title is "AceWiki" and the user is logged in as "Anonymous". The main content area displays the article "planet" with a navigation bar above it containing "Article", "References", "Individuals", and "Hierarchy".

AceWiki
Geography Wiki

Navigation:
- [Main Page](#)
- [Index](#)
- [Search](#)
- [About](#)
- [Random Article](#)

Actions:
- [New Word...](#)
- [Export...](#)

[Article](#) | [References](#) | [Individuals](#) | [Hierarchy](#)

planet

- ▷ *From Wikipedia: A planet, as defined by the International Astronomical Union (IAU), is a celestial body orbiting a star or stellar remnant that is massive enough to be rounded by its own gravity, is not massive enough to cause thermonuclear fusion, and has cleared its neighbouring region of planetesimals.*
- ▷ Every planet is a celestial-body.
- ▷ No planet is a star.
- ▷ No planet is a dwarf-planet.
- ▷ No planet is a moon.
- ▷ Every planet orbits a star.
- ▷ Every planet that does not orbit the Sun is an extrasolar planet.
- ▷ Which planets orbit the Sun?
 - Earth
 - Jupiter
 - Mars
 - Mercury
 - Neptune
 - Saturn
 - Uranus
 - Venus
- ▷ Every planet is a terrestrial planet or is a gas giant.

Evaluation

Verschiedene Experimente zeigen:

- Kontrollierte natürliche Sprachen (CNL) sind **einfacher zu verstehen** als andere formale Sprachen
- Zudem benötigen sie **weniger Lernzeit** und werden **von Benutzern bevorzugt**
- CNL-Systeme sind **einfach zu bedienen**, selbst für Benutzer ohne Informatik-Hintergrund und ohne Einarbeitungszeit
- Auch bei Benutzern mit Erfahrung in formalen Sprachen zeigen sich die Vorteile von CNL-Systemen

Kontrollierte natürliche Sprache im Unternehmen?

Bis jetzt haben kontrollierte natürliche Sprachen hauptsächlich im akademischen Bereich Verwendung gefunden.

Die Zeit ist reif für den operativen Einsatz!

Das Potential ist gross:

- Immer höher werdender Grad an Automation
- Die Kommunikation zwischen Mensch und Computer wird immer komplexer
- Die meisten Leute haben keine Erfahrung mit formalen Sprachen

Anwendung kontrollierter natürlicher Sprache in Unternehmen

Erfolgsversprechendes Anwendungsgebiet:

- Verschiedene Arten von Objekten, die in verschiedenen (nicht allzu komplexen) Relationen zueinander stehen
- Häufige Änderungen durch Nicht-Informatiker
- Nutzen durch automatische Ausführung von einfachen Regeln
- **Beispiele:** Produktpakete, Verträge, Service Level Agreements, Preispläne, Logistik, Projektmanagement, Marketing

Erfolgsversprechende Anwendung:

- Basierend auf erprobten Technologien (z.B. Semantic-Web-Technologien)
- Eigene kontrollierte natürliche Sprache, massgeschneidert auf das Anwendungsgebiet (z.B. inspiriert von ACE)
- Benutzerfreundlicher Editor

Fazit

Das Semantic Web hat bis jetzt noch nicht stattgefunden

- Es sind aber einige nützliche Technologien daraus hervorgegangen
- Die Benutzerinteraktion bleibt ein grosses Problem

Kontrollierte natürliche Sprachen können die Lösung sein

- Solche Sprachen sind einfacher zu verstehen und zu benutzen als andere formale Sprachen
- Der Einsatz in Unternehmen birgt grosses Potential

Weitere Informationen

ACE:

`http://attempto.ifi.uzh.ch/site/`

AceWiki:

`http://attempto.ifi.uzh.ch/acewiki/`