

# Das Semantic Web

---

## Was ist das Semantic Web?

### Definition

Mit dem Begriff "Semantic Web" werden unterschiedliche **Ziele** und **Methoden** beschrieben. Man kann unterscheiden zwischen:

- Das Semantic Web als eine **universelle Bibliothek** für menschlichen Zugriff. Das Semantic Web als ein einzelner, **globaler Datenraum**, der aus verknüpften RDF-Daten besteht - Das Ziel der Entwicklung des Semantic Web ist ein **integriertes Bedeutungssystem**, das eine Sammlung weit verstreuter Information organisiert. Stichwort: „Semantik“.
- Das Semantic Web als ein **Habitat** für automatisierte **Agenten** und Web Services - Im Semantic Web sollen nicht nur Menschen, sondern auch **Maschinen Daten auch auf Grund ihrer Bedeutung** bearbeiten können. Stichwort: „Maschinen“.
- Das Semantic Web als eine Methode, eine große Vielfalt von **Datenbanken und Knowledge Bases** zu fördern - Mit dem Semantic Web sollen Daten in so einer Weise verlinkt werden, dass sie eine **große, globale, vereinte Datenbank** bilden, die **dezentral** auf Millionen von Rechnern liegen können. Der **Wert** der Information liegt nicht nur in den Informationen an sich, sondern auch in großem Maße in der Weise, wie sie mit anderen Informationen **verknüpft** ist. „Stichwort: „Web“.

### Ziele

Im Augenblick unterscheidet man im Zusammenhang mit dem Semantic Web zwei Ziele:

#### Ziel 1: Das Semantic Web als Web von Daten

Das Hauptziel des Semantic Web ist die **Integration** aller strukturierten und semi-strukturierten Daten im WWW. Datensätze sollen im RDF Format im Web **veröffentlicht** werden und Ontologien die gemeinte **Semantik** dieser Datensätze erklären. Über **URIs** können sie miteinander verknüpft und in neuen, unerwarteten Möglichkeiten wiederverwendet werden.

#### Ziel 2: Das Semantic Web als eine Bereicherung des heutigen WWW

Das Semantic Web ist kein zusätzliches, separates Web, sondern eine **Erweiterung des WWW**.

**Automatische Prozesse** werden die benötigten semantischen **Metadaten** erzeugen. Durch diese Automatisierung werden die Daten in so einer Weise angereichert, dass dies wiederum semantische Schlussfolgerungen für Mensch und Maschine erleichtert.

### Entwicklung

Im Zusammenhang mit dem Semantic Web können **zwei Entwicklungsstränge** beobachtet werden. Der erste hat eher mit dem „Semantic“ zu tun, während der zweite vom „Web“ handelt.

#### Entwicklungsstrang 1: Künstliche Intelligenz

Das Semantic Web ist ein dritter Versuch nach **ELIZA** in den 1960er und **Künstliche Intelligenz** in den 1970er und 1980er Jahren, Computer das Verstehen beizubringen.

Es geht darum Information im Netz **formal und regelbasiert so aufzubereiten**, dass die Regeln von Maschinen verstanden und ausgeführt werden können. Berners-Lee sagte dann auch über das Semantic Web, dass Computer die Information nicht „verstehen“ würden, so wie wir Menschen es tun, sondern dass sie in der Lage sind, Begriffe effizient **so zu verarbeiten, dass es nützlich und bedeutsam für den menschlichen Benutzer** sei.

## Entwicklungsstrang 2: Net – Web – Semantic Web

Tim Berners-Lee sprach von **drei historischen Phasen** in der Entwicklung hin zum Semantic Web: Netz, Web und Semantic Web. Nachdem mit dem Netz **Rechner** und im Web **Dokumente** miteinander verknüpft wurden, kommt man nun in eine dritte Phase, in der erkannt wird, dass es nicht die Dokumente sind, die interessieren, sondern die **Daten**, die in den Dokumenten enthalten sind.

Auch die **Verfügbarmachung** von Daten durchläuft Phasen. Zuerst konnten Daten, wenn auch eher unbeholfen, durch z.B. „**Screen Scraping**“ aus den Dokumenten herausgezogen werden. Im Augenblick werden Daten hauptsächlich über **Web Services und APIs** strukturiert zur Verfügung gestellt. In einer nächsten Phase, im **Semantic Web** in einem maschinenlesbaren Format ohne irgendwelche Barrieren und Bedingungen im Netz zur Verfügung gestellt.

## Technologien des Semantic Web

### Begriffserklärungen

In Essenz geht es beim Semantic Web um **Ressourcen** und deren **Eigenschaften**.

#### *Ressourcen*

Eine Ressource kann alles sein, das eine **Identität** hat, was benannt oder adressierbar ist oder mit dem umgegangen werden muss. Eine Ressource ist ein „**Ding**“ worüber man sich im Semantic Web austauschen will. Ressourcen können z.B. (auch bereits verstorbene) **Autoren, Bücher, Herausgeber, Orte, Proteine** usw. sein. Auch **Dienstleistungen** (der heutige Wetterbericht von Köln), **abstrakte oder mythische Begriffe** oder **Glaubensinhalte** können Ressourcen sein. Weiterhin ist es möglich, dass Sammlungen von Ressourcen eine Ressource ausmachen.

Im Zusammenhang mit dem Semantic Web wird zwischen **zwei Ressourcen-Typen** unterschieden:

- **Informations-Ressourcen:** Dateien, Dokumente oder anderen Ressourcen auf die **über eine URL** zugegriffen werden kann (z.B. auch Bilder und andere Medien)
- **Nicht-Informations-Ressourcen:** es handelt sich hier um die Dinge, die nicht Informations-Ressourcen im obigen Sinne sind, Dinge, auf die man **nicht über eine URL** zugreifen kann. Alle „**real-world Objects**“ die außerhalb des WWW existieren, sind Nicht-Informations-Ressourcen.

Im Denken über und Verstehen von Ressourcen gehen wir nicht mit realen Objekten, sondern mit **Konzepten** von Objekten um. Alle Ressourcen werden über Konzepte begriffen.

In der Objektmodellierung werden die **abstrakten Konzepte** „**Klassen**“ genannt. Dinge im wirklichen Leben sind „Objekte“ oder „**Instanzen** von Klassen“.

Beispiel:

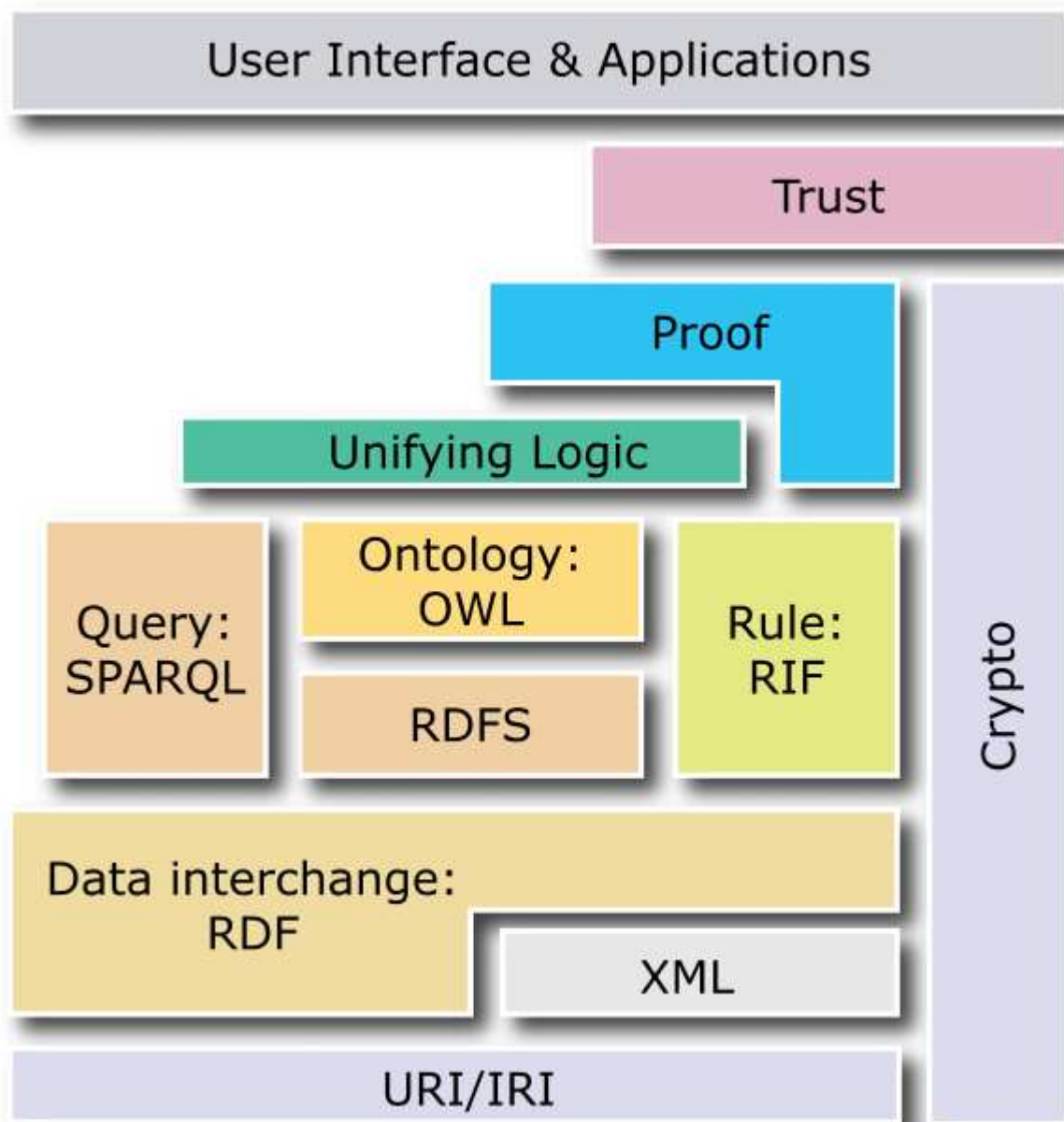
- Die Klasse „Buch“ kann eine Instanz/Objekt namens „The dispossessed“ haben.
- Die Klasse „Autor“ kann eine Instanz/Objekt namens „Ursula le Guin“ haben.

### Eigenschaften / Properties

Eigenschaften sind eine **besondere Art von Ressourcen**. Sie **beschreiben** Ressourcen und die **Beziehungen**, zwischen den Ressourcen. Beispiele sind „geschrieben von“, „hat den Titel“

### Der Semantic Web Layer Cake

Das W3C veröffentlicht im Internet eine grafische Darstellung der Semantic Web-Technologien und Designprinzipien, die als „Semantic Web Layer Cake“ bekannt geworden ist.<sup>1</sup>



<sup>1</sup> <http://www.w3.org/2007/03/layerCake.png>

## URI

Eine grundlegende Semantic Web-Technologie ist der URI, der **Uniform Resource Identifier**.

Ressourcen inkl. Eigenschaften werden im Semantic Web durch URIs **identifiziert**.

Im Semantic Web geht es darum, durch den Einsatz von **URIs** und **HTTP** strukturierte Daten im Web zu verknüpfen und so **Ressourcen mit anderen Ressourcen zu verbinden**.

### *Dereferenzierung*

Mit der Dereferenzierung von URIs wird der Prozess gemeint, in dem URIs über z.B. einem Browser aufgerufen werden mit dem Ziel, **auf die mit der URI referenzierten Ressource zuzugreifen**.

Da alle Ressourcen URIs haben sollen, aber nicht alle Ressourcen auf dem WWW lokalisiert werden können, sollte **bei Nicht-Informationen-Ressourcen** wenigstens **Information über** die Ressource gegeben werden.

## XML Technologie

XML ist die Abkürzung für eXtensible Markup Language und ist eine **Metasprache zur Modellierung von Auszeichnungssprachen**.

Ein XML-Dokument besteht aus einem **Prolog** mit mindestens einer **XML-Deklaration** und einer Reihe von **Elementen**. Die Elemente sind die **Entitäten**, die in XML-Dokumenten referenziert werden. Ein Element besteht aus einem **Start-Tag** und einem **End-Tag**, die den **Inhalt** des Elements umschließen. Der Inhalt von XML-Elementen kann aus **Text, anderen Elementen oder nichts** bestehen. Außerdem ist es möglich, **Attribute** zu einem Element zu vergeben. Attribute werden im Start-Tag eines Elements verzeichnet. Weitere Sprachelemente sind **Kommentare** und **Verarbeitungsinstruktionen**.

### *DTD und XML Schema*

Es gibt zwei XML-Mechanismen, mit denen die **Syntax einer XML-Datei definiert** werden können: DTD (eher veraltet) und XML Schema. Es handelt sich um **Modellierungssprachen**, die XML um **Datentypen, Strukturierungs- und Wiederverwendungsmöglichkeiten** erweitert. XML Schema beinhaltet im Vergleich zu DTDs stärker ausgereifte Mengen von **Datentypen** und ist selbst auch auf **XML basiert**.

### *XML Namensräume*

Ein **wichtiges Konzept**, das in XML benutzt wird (aber auch z.B. nach Turtle übernommen wurde) ist die Idee der Namensräume. Ziel der Namensräume ist es **Namenskonflikte bei der Verknüpfung von Daten aus verschiedenen Quellen** zu vermeiden.

### *XML Datentypen*

XML Schema bietet einen Mechanismus, Datentypen zu deklarieren. So kann z.B. festgelegt werden, dass die Zeichenfolge, die im Element gegeben wird kein String, sondern ein Integer ist. Einige Grunddatentypen sind bereits in XML Schema definiert. Es ist außerdem möglich, weitere Datentypen zu bestimmen.

## RDF

RDF, das „**Resource Description Framework**“ ist eine **universelle Sprache**, mit der **Ressourcen** und deren **Eigenschaften** und **Beziehungen** zu anderen Ressourcen auf einfache Art und Weise modelliert werden.

Wie mit XML werden mit RDF **keine semantischen Festlegungen** gemacht, aber dieses Datenmodell bietet die Möglichkeit semantische Festlegungen auf solcher Art und Weise zu **kodieren**, dass sie **maschinenlesbar** sind.

### *Das grundlegende Datenmodell von RDF: Tripel*

Die **Beziehungen** zwischen und **Eigenschaften** von **Ressourcen** werden in dem Datenmodell von RDF als sogenannte „**Tripel**“ ausgedrückt. Diese Tripel haben die Form von **Subjekt-Prädikat-Objekt**.

Beispiel einer typischen RDF-Aussage:

Das Buch „The Dispossessed“ wurde geschrieben von Ursula Le Guin.

**Subjekt:** Das Buch „The Dispossessed“

**Prädikat:** wurde geschrieben von

**Objekt:** Ursula Le Guin

„Ursula le Guin“ hat in diesem Beispiel eine Beziehung zu dem Buch „The Dispossessed“, weil sie die Autorin des Buches ist. Sie hat aber gleichzeitig auch **andere Beziehungen**. Sie ist z.B. auch Tochter von Alfred L. Kroeber und Einwohnerin von Seattle.

### *Graphen*

**Subjekt und Objekt** können auch als „**Knoten**“ und die Beziehungen zwischen ihnen als „**gerichteten, benannten Graphen**“ bezeichnet werden, z.B.



### *Verknüpfungen*

Für das **Subjekt** und **Prädikat** müssen **immer** und für das **Objekt** sollte, **wo möglich**, mit **URIs** auf Ressourcen verlinkt werden.



Es kann auch der **Namensraum-Mechanismus** von XML eingesetzt werden:

```
xmlns:dbpedia="http://dbpedia.org/pages/"
```

```
xmlns:dc="http://purl.org/dc/elements/1.1/"
```



Wenn es keine verlinkbare Ressource gibt, können **Objekte** auch eine **Zeichenfolge**, beinhalten oder leer sein.

### *Komplexere Aussagen*

Komplexere Aussagen können mit **leeren Knoten, Reifikation oder Container-Elementen** gemacht werden.

### *Serialisierung*

Damit die RDF-Aussagen **maschinenlesbar** sind, müssen sie z.B. mit **XML serialisiert** werden. Durch die Nutzung von XML kann RDF die **Vorteile**, die mit XML verbunden sind, **erben**.

Andere Formate der Serialisierung sind **Turtle** oder **N3**, die für Menschen einfacher zu lesen sind

### *Ontologien*

Bedeutung wird im Semantic Web durch Ontologien definiert.

In der Informatik ist „Ontologie“ ein **Artefakt**, das **Wissen über eine bestimmte Domäne** definiert.

Eine Ontologie ist ein **Begriffssystem**, das die Bedeutung von **Klassen** und **Konzepten**, deren **Eigenschaften, Beziehungen** und **Funktionen** definiert. Außerdem definieren Ontologien **Inferenz-Regeln**, die es erlauben, Ableitungen aus der Information machen zu können.

Die Welt ist voller „Ontologien“: Firmen haben **Produktkataloge**, Organisationen haben **Glossarien** und **Terminologien**, Bibliotheken haben **Metadaten Schemata, Normdaten, Schlagwortlisten, Klassifikationen, Thesauri**, etc.

Damit das Semantic Web gesponnen werden kann, ist es wichtig, dass es **Verknüpfungen zwischen Ontologien**, z.B. zwischen Einträgen der dbpedia und der PND gibt.

### *Maschinenlesbare Ontologien*

Folgendes sind **Hauptanforderungen** an Modellierungssprachen für Ontologien:

- klar **definierte Syntax**
- **formalisierte Semantik**
- **zweckmäßige Ausdrucksweise**
- **ausreichende Ausdruckskraft**
- Unterstützung für **logisches Schließen**

Zwischen den letzten beiden Anforderungen gibt es einen Konflikt: Je **ausdrucksstärker** eine Sprache ist, desto **ineffizienter** werden logische Operationen unterstützt.

Eine maschinenlesbare Modellierungssprache muss folgende **logischen Operationen** unterstützen:

- **Klassenzugehörigkeit**
- **Äquivalenz von Klassen**
- **Konsistenz**
- **Klassifikation**

### *RDF Schema*

RDFS (RDF Schema) ist eine **Schematasprache**, mit der **nicht-komplexe Ontologien** definiert werden können. Mit RDFS kann ein Vokabular definiert und Eigenschaften der **Konzepte, Werte** und **Einschränkungen**, die durch die Konzepte angenommen werden dürfen und die **Beziehungen** zwischen den Konzepten dargestellt werden.

RDFS erwartet **nicht**, dass die Klassen eine strikte **Hierarchie** bilden. Eine Klasse darf mehrere Superklassen haben.

Genauso wie es **Unterklassen** gibt, gibt es auch **Untereigenschaften**. So könnte „istAutorVon“ eine Untereigenschaft von „istSchöpferVon“ sein.

**Kernklassen** von RDFS sind:

- **rdfs:Resource**
- **rdfs:Class**
- **rdfs:Literal**
- **rdfs:Property**
- **rdfs:Statement**

Die Kerneigenschaften, mit denen Beziehungen definiert werden, sind:

- **rdfs:type**
- **rdfs:subClassOf**
- **rdfs:subPropertyOf**

Eine RDFS-Dokument ist ein normales RDF-Dokument und die XML-basierte Syntax von RDF kann eingesetzt werden.

### *OWL*

Die Web Ontology Language (OWL) ist eine **mächtigere Sprache** als RDFS um Ontologien zu beschreiben. Wie RDFS definiert OWL die **Bedeutung von Konzepten** in Ontologien, sowie die **Beziehungen** zwischen ihnen.

Folgendes sind Beispiele von **Operationen**, die nicht mit RDFS, wohl aber **mit OWL** möglich sind:

- in OWL können **Einschränkungen** für ausgewählte Klassen deklariert werden.
- In OWL ist ausdrückbar, dass Klassen **disjunkt** sind (z.B. „männlich“ und „weiblich“)
- OWL bietet die Möglichkeit, **neue Klassen aus Kombinationen existierender Klassen** zu bilden. (z.B. eine Klasse „Person“, die aus den Klassen „männlich“ und „weiblich“ bestehen.)
- OWL bietet **Kardinalitätseinschränkungen** (eine Person hat genau zwei Eltern)

OWL gibt es mehreren **Ausprägungen** in unterschiedlichen **Mächtigkeiten / Komplexitäten**:

- **OWL-Lite** wurde ursprünglich definiert um die Nutzer zu unterstützen, die in erster Linie eine **Klassifikationshierarchie** und **einfache Einschränkungen** brauchten.
- **OWL-DL** wurde definiert um ein **Maximum an Ausdruckskraft** zu ermöglichen, aber **rechnerische Vollständigkeit** und **Entscheidbarkeit** zu erhalten und praktische **Inferenz**-Algorithmen zur Verfügung zu stellen. OWL-DL enthält **alle Sprachkonstrukte** von OWL, aber sie können nur **eingeschränkt** benutzt werden.
- **OWL-Full** wurde entwickelt um eine **Kompatibilität mit RDFS** zu ermöglichen. Es unterstützt **alle Sprachkonstrukte** von OWL **ohne** Einschränkungen. Es ist aber wahrscheinlich **unmöglich rechnerisch** die **volle Inferenz** von OWL Full zu unterstützen.

Wie bei RDFS besteht eine OWL-Ontologie aus einer **Sammlung von Aussagen**, die **Klassen** und die **Beziehungen** zwischen den Klassen **einschränken**. Diese Aussagen ermöglichen es Systemen **zusätzliche Information** auf den bereits gegebenen Daten zu **schließen**.

OWL bietet wie RDFS **Sprachkonstrukte** für Klassen und Eigenschaften. Klassen können **disjunkt** sein und mit **booleschen** Operatoren kombiniert werden. Eigenschaften können **Beziehungen** darstellen oder Typenwerte beinhalten, **invertiert** werden oder **äquivalent** zu anderen Eigenschaften erklärt werden.

Alle OWL-Ontologien verwenden eine **RDF-Syntax** und werden meistens in **XML** serialisiert.

## SKOS

Das Simple Knowledge Organisation System (SKOS) dient der **Repräsentation** und dem **Austausch kontrollierter Vokabulare** (Klassifikationen, Thesauri, usw.) in einer Semantic Web-Umgebung.

SKOS definiert die Klassen und Eigenschaften, die die **allgemeinsten Funktionen** in kontrollierten Vokabularen ausdrücken. Jedes SKOS-Konzept ist eine **RDF-Ressource** und jedes Konzept kann **RDF-Eigenschaften** haben.

Eigenschaften können Konzepte entweder **dokumentieren** (z.B. skos:prefLabel) oder **Beziehungen** zwischen Konzepten anzeigen (z.B. skos:broader, skos:narrower)

Konzepte können in **Hierarchien** organisiert werden indem man weitere bzw. engere Beziehungen angibt. Außerdem können sie durch nicht-hierarchische (**assoziative**) Beziehungen darstellen. Konzepte können in **Konzept-Gruppen** gesammelt werden.

Das **Ziel** von SKOS ist es, das **Erbe der bestehenden Vokabularen** im Semantic Web zu veröffentlichen und zwar leichter als über das komplexere OWL. Mit SKOS wurden bereits einige Vokabularen wie **LCSH** oder der **STW Thesaurus for Economics** der ZBW in Kiel als Linked Data veröffentlicht.

## SPARQL

Für das Semantic Web wird eine neue **Abfragesprache** gebraucht. **XQuery** oder **XPath** können den Ansprüchen nicht genügen, da sie auf einer **niedrigeren Ebene der Abstraktion** als RDF gelegen sind. Darum wurde SPARQL entwickelt, welches es erlaubt, Daten in **RDF Tripel** zu prozessieren.

Es gibt **keine Beschränkung** auf die **Komplexität** einer SPARQL-Abfrage, außer die **Zeit**, die es braucht, sie zu prozessieren.

SPARQL-Abfragen **ähneln SQL-Abfragen** aber vergleichen die **Muster von RDF-Graphen** miteinander. Die Abfrage wird auch als **RDF-Tripel** gestellt, aber es ist möglich, eine **Variable** anstelle eines URI in den Positionen des Subjekt, Prädikats und Objekt zu stellen. Variablen sind z.B. die aus SQL bekannten **SELECT, FROM, WHERE**.

SPARQL ist **unverbindlich**, was die Unterstützung von **RDFS-Semantik** betrifft. Das Ergebnis der Abfrage hängt davon ab, ob das System, das die Abfrage beantwortet RDFS-Semantik unterstützt oder nicht.

## Logik

Logik wird im Semantic Web eingesetzt werden um **Schlüsse** aus den bereitgestellten Daten zu ziehen.

Die **Aufgaben** der Logik sind:

- **Verknüpfungen** von Aussagen
- **Schlussfolgerungen** aus den verknüpften Aussagen
- Formale **Beweisführung** durch Schlussregeln
- **Prioritätensetzung** für logische Patt-Situationen.

Die eingesetzte Logik muss stark genug sein, **komplexe Eigenschaften** von Objekten zu bearbeiten, aber Vorsicht muss geboten sein, da das Abfragesystem mit einem Paradox konfrontiert werden könnten.

Die logischen **Ableitungen** erfolgen über spezielle **Reasoner (Inferenzmaschinen)**.

Folgende Logikparadigmen werden für Ontologien eingesetzt.

### *Beschreibungslogik (bei RDFS, OWL-Lite und OWL-DL)*

Beschreibungslogik ist in den meisten Fällen **entscheidbar** und basiert auf die **Open World Assumption**. Da sie sehr komplex werden kann, kann Schließen **aufwändig** werden und deshalb ist es fraglich, ob Beschreibungslogik **effizient skaliert** werden kann.

### *Logic Programming (F-Logic) (bei OWL-Full)*

Logic Programming hat eine sehr **hohe Ausdrucksstärke** und kann **sehr effizient** inferieren. Sie basiert auf der **Closed World Assumption** und kann zu **unentscheidbaren** Situationen kommen. Sie ist **keine klassische Logik** und, im Gegensatz zur Beschreibungslogik, **kein W3C-Standard**.

### *DLP und Semantic WebRL als Brückenschlag zwischen Description Logic und Logic Programming*

- **Description Logic Programs (DLP)**, die bis zu **99%** der OWL-DL-Aussagen verarbeiten kann und **entscheidbar** bleibt
- **Semantic Web Rules Language (Semantic WebRL)**, die eine Vereinigung von OWL-DL und Logic Programming bildet, erweitert um RuleML. Sie kann **100%** der OWL-DL-Aussagen verarbeiten, kann aber **unentscheidbar** sein.

## Trust

Je mehr RDF-Daten veröffentlicht werden, desto öfter werden Aussagen der **Daten miteinander im Konflikt** stehen. In solchen Fällen wird **Vertrauen** in die Datenquelle eine wichtige Rolle spielen.

Vertrauen (Trust) wird durch eine **Sicherheitsarchitektur** erreicht. Die Ziele von Trust sind:

- **Verfügbarkeit**
- **Datenintegrität**
- **Vertraulichkeit**
- **Authentifizierung**
- **Autorisierung**

Trust wird durch folgende **Instrumente** erreicht:

- Digitale **Signaturen**, z.B. XMLSignature
- **Datenverschlüsselung**, z.B. XMLEncryption

Vertrauen im eigentlichen Sinne wird aber nicht so sehr durch Technologie sondern eher in Verbindung mit **Voting, Rating- und Reputationssystemen** aufgebaut. Dafür gibt es drei Strategien:

- Die **zentrale Strategie**, bei der eine zentrale Autorität Softwareagenten und Datenlieferanten **zertifiziert**
- Die **Recherche-Strategie**, bei der die Softwareagenten einen **Vertrag** aushandeln
- **Transitive Strategie**, die auf **Bewertungen** des Vertrauens durch Benutzer/Agenten untereinander basiert.

## Die Entwicklung des Semantic Web: Wo stehen wir heute?

Heute befinden wir uns auf der Ebene, in der **Datenmodelle** und **Schemata** die verschiedene Perspektiven auf Konzepte und Eigenschaften haben, entwickelt werden.

Folgende **Forschungsprobleme** befinden sich gerade in Bearbeitung:

- **Skalierbarkeit**
- **Entwicklung und Bedeutungswandel von Ontologieinhalten**
- **Ontologie-Mapping**

Außerdem gibt es noch **Probleme**, die zwar bekannt, aber im Augenblick noch **nicht** so sehr bearbeitet werden:

- **Interaktion** zwischen maschinengängigen Repräsentationen und **dynamischen sozialen Benutzer-Netzwerken**
- Verfahren für den Umgang mit **Vertrauen, Reputation, Integrität und Herkunft**
- Inferenz- und Suchinstrumente, die **hinreichend robust mit begrenzten Ressourcen** angewandt werden können und die zu einer intelligenten **Priorisierung** bei Konflikten zwischen Ressourcenbeschränkungen und Ergebnisqualität im Stande sind.

Die Semantic Web-Bewegung befand sich lange Zeit in einer **Zwickmühle**: solange es keine **Applikationen** gab, die mit den **Daten** umgehen können, wurden keine Daten zur Verfügung gestellt. Die Applikationen können aber nur entwickelt werden, wenn es Daten gibt, die sie verarbeiten können. Die **Linked Data Bewegung** hat es zum Ziel so schnell und so viel wie möglich Daten im Semantic Web zu veröffentlichen, so dass Applikationen entwickelt werden können.

## Verstehen / Semantik / Bedeutung: welche Rolle spielen sie wirklich im Semantic Web?

Inwiefern befähigt das Semantic Web **Computer** Dinge zu **verstehen**?

Nach der Meinung von Tim Berners-Lee können Maschinen Information **nicht wirklich „verstehen“**, aber sie können Informationen sehr **effizient manipulieren**, so dass sie für den Menschen **brauchbar** und **sinnvoll** werden. „Semantik“ sollte eher als „**semantische Integration**“ verstanden werden und als ein Prozess, in dem **Informationen** aus verschiedenen Quellen **miteinander in Beziehung** gesetzt werden.

Wenn überhaupt, werden Maschinen höchstens **auf der Ebene der Logik** intelligente Schlüsse ziehen können. **Verstehen**, auch im kognitiven Sinne wird kaum möglich sein, wegen anderen, eher **nicht-rationalen** Aspekte, die auch eine Rolle im menschlichen „Verstehen“ spielen: Emotionen, Vorlieben und Abneigungen, Träume und Visionen. Selbst die Science Fiction, die oft eine sehr viel stärkere Maschinenintelligenz als „nur“ das Semantic Web zur Vision hat, stellt sich die Frage: „Do androids dream of electric sheep?“ (Philip K. Dick)

## Beurteilung des Semantic Web

Das Semantic Web ist ein so **großes und kompliziertes** Unterfangen, dass sich die Frage gestellt werden muss, ob es sich **lohnt**, es zu verwirklichen. Im Folgenden soll erklärt werden, welche Vorteile das Semantic Web mit sich bringt, und warum dessen Entwicklung vorangetrieben werden sollte.

- **Maschinelle Verarbeitung** - Die wichtigste Eigenschaft des Semantic Web ist, dass **Maschinen Information verstehen** und verarbeiten sollen. Dies hat eine **Arbeitserleichterung** zur Folge. Ein Beispiel ist die Beherrschung der **Informationsflut**. Durch Semantic Web Technologie können Computer Information erschließen.
- **Verbesserte Suche** - Die Suche wird sich verbessern, da Systeme Ontologien benutzen können um genau die Information zu finden, die der Nutzer braucht. Nicht nur kann eine Suche um **verwandte Begriffe** ausgebreitet oder eingeschränkt werden, sondern auch, wie bei **Homonymen**, auf genau den richtigen Begriff beschränkt werden. Außerdem kann durch den Einsatz von Semantic Web Technologien auch Fragen beantwortet werden, deren vollständigen Antworten komplexer sind und **bruchstückhaft verstreut** auf mehreren Websites liegen können.
- **Dezentrale Datenhaltung** - Da URIs zur Identifikation von Ressourcen benutzt werden, können Systemen **dezentralisiert wachsen**.
- **Integration von Daten** - Durch Semantic Web-Technologie ist es möglich Daten aus **verschiedenen, heterogenen Quellen** zu integrieren.
- **Koexistenz verschiedener Datenformate**
- **Vereinfachte Verfügbarmachung von Daten** - Daten können zu **minimalen Kosten** veröffentlicht, und einem **globalen Publikum** zur Verfügung gestellt werden.
- **Erhöhter Wert der Daten** - Durch die Bereitstellung von Daten an andere Domänen werden die **Daten wertvoller** als sie es im Augenblick sind.

## Vorteile gegenüber APIs oder Web Services

**APIs und Web Services** ermöglichen auch ohne Semantic Web-Technologien einen Zugriff auf Daten, z.B. bei Amazon, EBay, Yahoo und Google oder auch bibliothekarische APIs wie z39.50 oder SRW/SRU. Allerdings bietet das Semantic Web entscheidende **Vorteile** gegenüber APIs:

- Die Daten werden leichter durch **Suchmaschinen gecrawlt**
- **Generische Datenbrowser** können auf die Daten zugreifen
- **Verlinkung zwischen Daten** aus verschiedenen Quellen ist möglich
- Es ist einfacher neue Datenquellen in **Mashups** einzufügen
- Verschiedene APIs haben unterschiedliche **Zugangsmodalitäten**. Mit Semantic Web-Technologie ist Zugang vereinheitlicht.

RDF kann also als ein **universelles API** gesehen werden.

## Hürden, die noch genommen werden müssen

- **Wichtige Spieler sind noch nicht vom Semantic Web überzeugt** - Einige wichtige Datenprovider können mit dem Semantic Web nicht viel anfangen. Es **fehlt noch ein Verständnis** für den Nutzen dieser Technologie.
- **Semantic Web-Technologie ist nicht einfach** - Es werden aber voraussichtlich, so **wie mit dem Dokumentenweb** sein: Es werden Werkzeuge entwickelt werden, mit denen die Daten semantisch aufbereitet und veröffentlicht werden können, so dass jeder, ohne auch nur ein bisschen Hintergrundwissen übers Semantic Web zu haben, dennoch seine Daten dafür bereitstellen kann.
- **Wichtige Aspekte des Semantic Web müssen noch entwickelt werden** - Während die grundlegende Technologie steht, so dass Daten im RDF-Format bereitgestellt werden können, sind die **oberen Schichten des Semantic Web-Layer Cakes** in einigen Aspekten noch Theorie.
- **Daten müssen veröffentlicht und verknüpft werden** - Damit die oben genannten **Herausforderungen** angegangen werden können, braucht es vor allem **Daten, die veröffentlicht werden müssen**.

## Bibliographie

Folgende Quellen wurden für diesen Aufsatz konsultiert:

**Addis A.**, 2003. SCULPTEUR: Towards a New Paradigm for Multimedia Museum Information Handling - ECS EPrints Repository. In Semantic Web ISWC. Available at: <http://eprints.ecs.soton.ac.uk/8593/>.

**Altenhöfner, R.**, 2007. Normdaten als Wissensknoten im Netz: Ein neues kooperatives Servicemodell? In 3. Leipziger Kongress für Information und Bibliothek AG Verbundsysteme. Leipzig. Available at: <http://www.opus-bayern.de/bib-info/volltexte/2007/379/>.

**Anonymous**, 2007a. Digital collections, museums and the semantic web - discussion. Available at: <http://culturalsemanticweb.wordpress.com/workshop-reports/workshop-four-19-february-2007-london-victoria-and-albert-museum-agenda/digital-collections-museums-and-the-semantic-web-discussion/>.

**Anonymous**, 2007b. Seamus Ross and his view of the semantic web. Available at: <http://culturalsemanticweb.wordpress.com/workshop-reports/workshop-two-glasgow-september-2006/1-seamus-ross-and-his-view-of-the-semantic-web/>.

**Antoniou, G. & van Harmelen, F., 2008.** A Semantic Web Primer 2undefined Aufl., Cambridge, Mass.: The MIT Press.

**Bergman, M., 2008.** Linked Data Comes of Age - LinkedData Planet and LDOW Set the Pace for 2008. Available at: <http://www.mkbergman.com/?p=416> [Zugegriffen März 14, 2009].

**Bergman, M.,** More Structure, More Terminology and (hopefully) More Clarity. Available at: <http://www.mkbergman.com/?p=391> [Zugegriffen März 14, 2009].

**Bergman, M.,** What is the Structured Web? Available at: <http://www.mkbergman.com/?p=390> [Zugegriffen März 22, 2009].

**Berners-Lee, T.,** Giant Global Graph. Available at: <http://dig.csail.mit.edu/breadcrumbs/node/215> [Zugegriffen März 14, 2009].

**Berners-Lee, T.,** Linked Data. Available at: <http://www.w3.org/DesignIssues/LinkedData.html> [Zugegriffen März 14, 2009].

**Berners-Lee, T., Hendler, J. & Lassila, O., 2001.** The semantic web : a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American. Available at: <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>.

**Bizer, C., Cyganiak, R. & Heath, T.,** How to Publish Linked Data on the Web. Available at: <http://sites.wiwiss.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/20070727/> [Zugegriffen Januar 14, 2009].

**Bizer, C., Gauß, T. u. a.,** Semantic Web Client Library. Available at: <http://www4.wiwiss.fu-berlin.de/bizer/ng4j/semwebclient/> [Zugegriffen März 14, 2009].

**Bizer, C., Heath, T. u. a.,** Linked Data on the Web (LDOW2008). In 17th International World Wide Web Conference. Beijing, China. Available at: <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-369/paper00.pdf> [Zugegriffen März 14, 2009].

**Borst, T. & Neubert, Joachim, 2009.** Case Study: Publishing STW Thesaurus for Economics as Linked Open Data. Available at: <http://www.w3.org/2001/sw/sweo/public/UseCases/ZBW/> [Zugegriffen Juni 15, 2009].

**Brooks, T.A., 2002.** The Semantic Web, Universalist Ambition and Some Lessons from Librarianship. Information Research, 7(4). Available at: <http://informationr.net/ir/7-4/paper136.html> [Zugegriffen Juni 15, 2009].

**Buzinkay, M., 2009.** Semantic Web, eine potentielle Spielwiese für BibliothekarInnen. MBI Blog. Available at: <http://www.buzinkay.net/blog-de/2009/05/semantic-web-eine-potentielle-spielwiese-fur-bibliothekarinnen/> [Zugegriffen Juni 15, 2009].

**Cho, A., 2009.** How RDF Can Use MARC in the Semantic Web World: Using Existing Library Cataloguing Methods in Organizing the Web. Available at: [http://internet.suite101.com/article.cfm/how\\_marc\\_realizes\\_the\\_rdf\\_in\\_the\\_semantic\\_web](http://internet.suite101.com/article.cfm/how_marc_realizes_the_rdf_in_the_semantic_web) [Zugegriffen Juni 15, 2009].

**D'Arcus, B. & Giasson, F.**, 2008. Bibliographic Ontology Specification. Available at: <http://bibliontology.com/> [Zugegriffen Juni 16, 2009].

**Dekkers, M., Gradmann, S. & Meghini, C.**, 2009. Europeana Outline Functional Specification : for development of an operational European Digital Library, Available at: [http://dev.europeana.eu/public\\_documents/EDLnet%20D2.5\\_Outline\\_Functional\\_Specifications20090301\\_version%201.7\\_consWithoutHistory\\_lossless.pdf](http://dev.europeana.eu/public_documents/EDLnet%20D2.5_Outline_Functional_Specifications20090301_version%201.7_consWithoutHistory_lossless.pdf) [Zugegriffen März 18, 2009].

**Du Charme, B.**, 2008. Scraping and linked data. Available at: <http://www.snee.com/bobdc.blog/2008/01/scraping-and-linked-data.html> [Zugegriffen März 14, 2009].

**Dunsire, G.**, 2008. The Semantic Web and expert metadata: pull apart then bring together. In Archives, Libraries, Museums 12 (AKM12). Porec, Croatia: Centre for Digital Library Research. Available at: <http://eprints.rclis.org/15880/1/akm2008semanticweb.pdf>.

**Généreux, M.**, 2007. Cultural Heritage Digital Resources: from Extraction to Querying. In ACL 2007 Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2007). Prague, Czech Republic. Available at: <http://eprints.brighton.ac.uk/3211/>.

**Giustini, D. & Cho, A.**, The Semantic Web as a Large, Searchable Catalogue: A Librarian's Perspective. Semantic Universe. Available at: <http://www.semanticuniverse.com/articles-semantic-web-large-searchable-catalogue-librarian%E2%80%99s-perspective.html> [Zugegriffen Juni 16, 2009].

**Goodall, S. u. a.**, 2004. Knowledge-based exploration of multimedia museum collections. In European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT). London, U.K. Available at: <http://eprints.ecs.soton.ac.uk/10899/>.

**Gradmann, S.**, 2008. Signal. Information. Zeichen. Zu den Bedingungen des Verstehens in semantischen Netzen. In Antrittsvorlesung. Berlin: Humboldt-Universität zu Berlin, Institut für Bibliotheks- und Informationswissenschaft. Available at: <http://edoc.hu-berlin.de/humboldt-vl/157/gradmann-stefan-3/PDF/gradmann.pdf> [Zugegriffen April 2, 2009].

**Gruber, T.**, 2008. Ontology. In L. Liu & M. T. Özsu, hrsg. Encyclopedia of Database Systems. Springer. Available at: <http://tomgruber.org/writing/ontology-definition-2007.htm> [Zugegriffen März 14, 2009].

**van Harmelen, F.**, 2006. Semantic Web Research anno 2006: main streams, popular fallacies, current status and future challenges. In Payne, hrsg. 10th International Workshop on Cooperative Information Agents. Lecture Notes in Artificial Intelligence. Springer, S. 1-7. Available at: <http://www.cs.vu.nl/~frankh/abstracts/CIA06.html> [Zugegriffen April 22, 2009].

**Haslhofer, B.**, 2008. Metadaten im Semantic Web. Available at: <http://www.opus-bayern.de/bib-info/volltexte/2008/581/> [Zugegriffen Juni 16, 2009].

**Haslhofer, B. & Schandl, B.**, 2009. The OAI2LOD Server: Exposing OAI-PMH Metadata as Linked Data. In Linked Data on the Web Workshop. CEUR Workshop Proceedings. Beijing, China. Available at: <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-369/paper03.pdf>.

**Heery, R. & Wagner, H.**, 2002. A Metadata Registry for the Semantic Web. D-Lib Magazine, 8(5). Available at: <http://www.dlib.org/dlib/may02/wagner/05wagner.html> [Zugegriffen Juni 16, 2009].

**Hyvönen, E., Holi, M. & Viljanen, K.**, 2004. Designing and creating a web site based on RDF content. In WWW2004 Workshop on Application Design, Development, and Implementation Issues in the Semantic Web. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.1.1.2561>.

**Hyvönen, E., Junnila, M. u. a.**, 2004. Publishing Museum Collections on the Semantic Web - The MuseumFinland Portal. In WWW2004. New York. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.2.454> [Zugegriffen März 14, 2009].

**Idehen, K. & Blakeley, C.**, 2008. Deploying Linked Data. Available at: [http://virtuoso.openlinksw.com/Whitepapers/pdf/Deploying\\_Linked\\_Data\\_Final1.pdf](http://virtuoso.openlinksw.com/Whitepapers/pdf/Deploying_Linked_Data_Final1.pdf) [Zugegriffen März 14, 2009].

**Keene, S.**, 2007. Towards the semantic web in museums. Available at: <http://culturalsemanticweb.wordpress.com/workshop-reports/workshop-four-19-february-2007-london-victoria-and-albert-museum-agenda/suzanne-keene-towards-the-semantic-web-in-museums/>.

**Le Boeuf, P. u. a.**, 2005. Using an ontology for interoperability and browsing of museum, library and archive information. In International Council of Museums 14th Triennial Meeting. The Hague. Available at: <http://eprints.ecs.soton.ac.uk/14790/> [Zugegriffen März 18, 2009].

**Malmsten, M.**, Making a Library Catalogue Part of the Semantic Web. In Proceedings of the International Conference on Dublin Core and Metadata Applications. Available at: <http://dcpapers.dublincore.org/ojs/pubs/article/view/927/923> [Zugegriffen Juni 14, 2009].

**Mazzocchi, S., Garland, S. & Lee, R.**, 2005. SIMILE: Practical Metadata for the Semantic Web. xml.com. Available at: <http://www.xml.com/pub/a/2005/01/26/simile.html> [Zugegriffen Juni 15, 2009].

**Miles, A. u. a.**, 2005. SKOS Core: Simple Knowledge Organisation for the Web. Available at: <http://isegserv.itd.rl.ac.uk/public/skos/press/dc2005/dc2005skospaper.pdf> [Zugegriffen Juni 27, 2009].

**Miller, E.**, 1998. An Introduction to the Resource Description Framework. D-Lib Magazine, 4(5). Available at: <http://www.dlib.org/dlib/may98/miller/05miller.html> [Zugegriffen März 22, 2009].

**Mulholland, P., Zdrahal, Z. & Collins, T.**, 2008. Investigating the Effects of Explanatory Semantic Search on the Use of a Museum Archive. In IEEE 2008 International Conference on Distributed Human-Machine Systems. Athens, Greece. Available at: <http://oro.open.ac.uk/10718/1/mulholland-DHMS08.pdf>.

**Nilsson, M. u. a.**, 2008. Expressing Dublin Core metadata using the Resource Description Framework (RDF). Available at: <http://dublincore.org/documents/documents/2008/01/14/dc-rdf> [Zugegriffen April 5, 2009].

**van Ossenbruggen, J. u. a.**, 2007. Searching and Annotating Virtual Heritage Collections with Semantic-Web Techniques. In San Francisco. Available at:

<http://www.archimuse.com/mw2007/papers/ossenbruggen/ossenbruggen.html> [Zugegriffen Juni 16, 2009].

**Palmer, S.B.**, 2001. The Semantic Web, Taking Form. infomesh.net. Available at: <http://infomesh.net/2001/06/swform/> [Zugegriffen Juni 17, 2009].

**Panzer, M.**, 2008. Cool URIs for the DDC: Towards Web-Scale Accessibility of a Large Classification System. In Proceedings of the International Conference on Dublin Core and Metadata Applications. Available at: <http://dcpapers.dublincore.org/ojs/pubs/article/view/932/928> [Zugegriffen Juni 15, 2009].

**Powell, A.**, 2009. Repositories and linked data. eFoundations. Available at: <http://efoundations.typepad.com/efoundations/2009/06/repositories-and-linked-data.html> [Zugegriffen Juni 18, 2009].

**Sauermann, L., Cyganiak, R. & Völkel, M.**, 2007. Cool URIs for the Semantic Web. Available at: <http://www.dfki.uni-kl.de/~sauermann/2006/11/cooluris/> [Zugegriffen Februar 26, 2009].

**Semenova, E.**, 2007. Ontologie als Begriffssystem : theoretische Überlegungen und ihre praktische Umsetzung bei der Entwicklung einer Ontologie der Wissenschaftsdisziplinen. In 11. Tagung der Deutschen Sektion der ISCO. Archäologisches Landesmuseum Konstanz.

**Stevenson, A. u. a.**, 2003. Semantic Web Techniques for Multimedia Museum Information Handling. In CIDOC/ADIT 2003. State Russian Museum, St Petersburg, Russia. Available at: [http://eprints.ecs.soton.ac.uk/8920/1/SCULPTEUR\\_CIDOC\\_2003.pdf](http://eprints.ecs.soton.ac.uk/8920/1/SCULPTEUR_CIDOC_2003.pdf) [Zugegriffen März 22, 2009].

**Styles, R., Ayers, D. & Shabir, N.**, 2009. Semantic MARC, MARC21 and the Semantic Web. In Linked Data on the Web Workshop. CEUR Workshop Proceedings. Beijing, China. Available at: <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-369/paper02.pdf> [Zugegriffen März 17, 2009].

**Summers, E. u. a.**, 2008. LCSH, SKOS and Linked Data. In Proceedings of the International Conference on Dublin Core and Metadata Applications. Available at: <http://dcpapers.dublincore.org/ojs/pubs/article/view/916/912> [Zugegriffen Juni 15, 2009].

**Sure, Y.**, 2005. Semantic Web Tutorial. semanticweb.org. Available at: <http://rease.semanticweb.org/ubp> [Zugegriffen Juni 17, 2009].

**Szász, B. u. a.**, 2006. Cultural Heritage on the Semantic Web - the Museum24 project. In Symposium on Digital Semantic Content across Cultures. Paris, the Louvre. Available at: <http://www.seco.tkk.fi/events/2006/2006-05-04-websemantique/presentations/articles/Szasz-museum24Paris.pdf> [Zugegriffen Februar 18, 2009].

**Tennant, R.**, The Semantic Web, Part 1 : Promises, Promises. Library Journal. Available at: <http://www.libraryjournal.com/blog/1090000309/post/1910038791.html> [Zugegriffen März 15, 2009].

**Tennant, R.**, The Semantic Web, Part II : Linked Data. Library Journal. Available at: <http://www.libraryjournal.com/blog/1090000309/post/1930038793.html> [Zugegriffen März 15, 2009].

**Tolksdorf, R.**, 2005. REASE - Semantic Web - Überblick und Einleitung. In semanticweb.org. Available at: <http://rease.semanticweb.org/ubp> [Zugegriffen Juni 17, 2009].

**Tonkin, E. & Strelnikov, A.**, 2009. Spinning a Semantic Web for Metadata: Developments in the IEMSR. Ariadne, 59(April 2009). Available at: <http://www.ariadne.ac.uk/issue59/tonkin-strelnikov/> [Zugegriffen Juni 15, 2009].

**W3C**, Best Practice Recipes for Publishing RDF Vocabularies. Available at: <http://www.w3.org/TR/swbp-vocab-pub/> [Zugegriffen März 14, 2009].

**W3C**, 2004. RDF Primer. Available at: <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/> [Zugegriffen April 7, 2009].

**W3C**, 2004b. RDF/XML Syntax Specification (Revised). Available at: <http://www.w3.org/TR/rdf-syntax-grammar/>.

**W3C**, 2004c. Resource Description Framework (RDF): Concepts and Abstract Syntax. Available at: <http://www.w3.org/TR/rdf-concepts/>.

**W3C**, SKOS Simple Knowledge Organization System Primer : W3C Working Draft 29 August 2008. Available at: <http://www.w3.org/TR/2008/WD-skos-primer-20080829/> [Zugegriffen Januar 5, 2009].

**Wikipedia**, 2009a. Linked Data. Available at: [http://en.wikipedia.org/wiki/Linked\\_Data](http://en.wikipedia.org/wiki/Linked_Data) [Zugegriffen März 14, 2009].

**Wikipedia**, 2009b. Simple Knowledge Organization System - Wikipedia, the free encyclopedia. <http://en.wikipedia.org>. Available at: [http://en.wikipedia.org/wiki/Simple\\_Knowledge\\_Organization\\_System](http://en.wikipedia.org/wiki/Simple_Knowledge_Organization_System) [Zugegriffen Juni 28, 2009].

**Wikipedia**, 2009c. Web Ontology Language. [en.wikipedia.org](http://en.wikipedia.org). Available at: [http://en.wikipedia.org/wiki/Web\\_Ontology\\_Language](http://en.wikipedia.org/wiki/Web_Ontology_Language) [Zugegriffen Juni 28, 2009].