

RDF basierte semantische Annotationsansätze

Art der Arbeit

Betreuer: ...

vorgelegt von: ...
Straße
PLZ Ort
Telefonnummer
E-Mail

Matrikelnummer: ...

Bearbeitungszeitraum: ... -

Abgabetermin: ...

Inhaltsverzeichnis

Formelverzeichnis	V
Abkürzungsverzeichnis	VI
1 Einleitung	1
1.1 Motivation	1
1.2 Methodik	1
1.3 Aufbau	2
2 RDF Grundlagen	3
2.1 Entwicklung	3
2.2 Beschreibung	3
2.2.1 Syntax	4
2.2.2 Semantik	5
3 RDF und Linked-Data	6
Eidesstattliche Erklärung	IX
Literaturverzeichnis	IX

Abbildungsverzeichnis

Abb. 1 Kernkonzepte des RDFs	4
---	---

Tabellenverzeichnis

Tab. 1 Die Entwicklung von RDF (vgl. [Abs. 2]Klyne:04:RDF, [Abs. 2]Schreiber:14:RP, Wood:14:WNR)	3
---	---

Formelverzeichnis

Abkürzungsverzeichnis

1. Einleitung

1.1. Motivation

Im Zuge der Globalisierung sind viele Unternehmen vom hohen Wettbewerbsdruck betroffen (cmt: Handbuch Interorganisationssysteme). Um ihren Ertrag langfristig zu sichern und innovativ gegenüber der Konkurrenz zu bleiben, bieten viele Industrieunternehmen, zum Beispiel „Value Added Services“, um die steigende Dienstleistungsnachfrage ihrer Kunden zu befriedigen [vgl.][S. 4]Meffert2015. Das Erbringen dieser Dienstleistungen heutzutage erfordert Geschäftsprozesse die zwischenbetrieblich ausgerichtet sind [vgl.][S.19]fleisch2001netzwerkunternehmen. Im Kontext einer vernetzten Welt [vgl.]bmwi2013 ist es von besonderer Bedeutung, dass Datenintegration während der Unterstützung zwischenbetrieblicher Tätigkeiten auch erfolgen kann, wenn diese Daten über das World Wide Web ausgetauscht werden. Dies kann insbesondere Hilfreich sein um Daten- und Funktionsintegration zwischen kleinen und mittelständischen Unternehmen zu unterstützen, im Fall das die Implementierung von Electronic Data Interchange (EDI) Alternativen zu teuer und unrentabel geschätzt würden. (cmt: empty cite)

1.2. Methodik

Das Ziel der vorliegenden Arbeit ist ein Literatur-Review durchzuführen, um die relevante Literatur über das Resource Description Framework auszuwählen, zu analysieren und zusammenzufassen im Hinblick auf den möglichen Einsatz von RDF für semantische Annotation von Daten innerhalb kleine und mittelständische Unternehmen. Das Review wird anhand der Methodik von fettke2006state durchgeführt. Eine strukturierte Literatursuche wurde mithilfe der Leitlinien von [S. XVI]webster2002analyzing umgesetzt um eine Liste von relevanten Literatur zu erstellen.

1. Im ersten Schritt wurden *EBSCOHost*, *Google Scholar* und *W3C Recommendations* verwendet, um eine Literaturliste zu erstellen.
2. Danach wurden alle Quellen der Literaturliste untersucht, um weitere relevante Literatur zu finden (sogenannte „Go Backward“ Ansatz).
3. Schließlich wurden Google-Scholar und Semantic-Scholar benutzt, um weitere Publikationen zu identifizieren, die auf der Literaturliste von Schritt eins und zwei, verweisen (sogenannte „Go Forward“ Ansatz).

Die folgenden Suchschlüssel wurden während der Literatursuche benutzt:

- Primärschlüssel: `Resource Description Framework`, `RDF`, `JSON-LD`, `RDFa`
- Sekundärschlüssel: `semantic annotation`, `semantic web`, `linked data`

Zum Schluss wurde die Liste der relevanten Literatur mit Blick auf der Zielstellung verarbeitet und nach wichtige Konzepte systematisiert [vgl.][S. XVI]webster2002analyzing.

1.3. Aufbau

Nach der Einleitung wird die Entwicklung von RDF im zweiten Abschnitt kurz beschrieben, und die benötigte RDF-Grundlagen und zusammenhängende Begrifflichkeiten werden herausgearbeitet. Die Syntax von RDF wird auch anhand einiger Beispiele demonstriert. Der dritte Abschnitt erläutert die Rolle von RDF in Linked-Data-Anwendungen. Das Konzept des Semantic-Webs wird vorgestellt und das Resource Description Framework in Attributes (RDFa) wird als exemplarischer Ansatz beschrieben. Ferner wird die Anwendung von Linked-Data erweitert, indem neue Ansätze gezeigt werden. Die Rolle von JSON-Linked-Data (JSON-LD) als alternative RDF-Serialisierungssyntax, um REST APIs erweiterbar zu gestalten, wird erkundet. Im letzten Abschnitt werden die Kernkonzepte der Arbeit zusammengefasst und es wird beschrieben, welche Weiterentwicklungen künftig von Interesse sein könnten.

2. RDF Grundlagen

2.1. Entwicklung

Das Resource Description Framework (RDF) wurde ursprünglich 1999 vom World Wide Web Consortium (W3C) als Empfehlung verabschiedet Lassila:99:RMS. 2004 wurde diese Version aktualisiert und in mehrere Dokumente aufgeteilt Beckett:04:RSS. Die aktuelle, erweiterte Version (RDF 1.1) wurde 2014 veröffentlicht Schreiber:14:RP. RDF 1.1, im Weiteren nur als RDF bezeichnet sofern nicht anders festgelegt, hat das Ziel die Unterstützung neuer Anwendungsfelder für RDF zu stärken. Tabelle 1 zeigt die Erweiterung von RDF 1.0 auf RDF 1.1. Davon kann man ableiten, dass RDF sich in die Richtung entwickelt, immer mehr domänenspezifische Anwendung zu unterstützen.

Unterstützt	RDF 1.0	RDF 1.1
Web-Metadaten	RDF/XML	HTML5+RDFa 1.1, JSON-LD
Datenaustausch zw. Datenbanken	RDF/XML ¹	JSON-LD, TriG, N-Quads
API-Feeds Verbinden	RDF/XML ¹	JSON-LD, RDF/XML

Tab. 1 Die Entwicklung von RDF
(vgl. [Abs. 2]Klyne:04:RDF, [Abs. 2]Schreiber:14:RP, Wood:14:WNR)

2.2. Beschreibung

Laut **Schreiber:14:RP** in **Schreiber:14:RP**

RDF is intended for situations in which information on the Web needs to be processed by applications, rather than being only displayed to people. RDF provides a common framework for expressing this information so it can be exchanged between applications without loss of meaning.

Wie die Name schon vermuten lässt, bietet das Resource Description Framework

¹Diese Anwendungsfelder wurden in [Abs. 2]Klyne:04:RDF nicht expliziert gegeben, aber RDF/XML in diesem Kontext zu nutzen ist auch möglich.

ein Gerüst (Modell, Sprachen und Syntaxen) für die Beschreibung von Attributen, Funktionen und Beziehungen der Ressourcen, wobei Ressourcen können alles was einen einzigartigen Identifier (URI oder IRI) haben kann sein [vgl.][Folie 6]Dekeyzer2013. Im Hinblick auf die Semiotische Ebenen der Integration [vgl.]Schissler2004, können ein Sender und Empfänger mittels RDF eine gemeinsame Sprache haben, um sich auf der syntaktische, und mithilfe von Ontologien und Vokabularen (vgl. Abschnitt 2.2.2), semantische Ebene verständigen zu können.

RDF stellt Zusammenhänge zwischen Ressourcen als eine Menge gerichteter Graphen dar. Ein gerichteter Graph (genannt „Tripel“ in RDF) hat eine Knote (Subjekt), die über eine gerichtete Kante (Prädikat) mit einer anderen Knote (Objekt) in Verbindung steht. Abbildung 1 veranschaulicht diesen Aspekt. Das Beispiel² drückt die Beziehung zwischen einem Kartograph und seinen produzierten Werken in der RDF-abstrakt Syntax aus.

Abb. 1 Kernkonzepte des RDFs anhand eines Beispielgraphs

Abbildung 1 schildert auch, als Orientierung, die Ähnlichkeiten zwischen den Subjekt-Prädikat-Objekt Denkweise von RDF und die typische Objekt-Orientierte Denkweise. Die konkrete RDF-Serialisierungssyntaxen werden im Abschnitt 2.2.1 erörtert. Jeder Komponent eines RDF-Tripels kann ein von drei Ausprägungen haben [vgl.][Abs. 3,1]Wood:14:RCA.

- Ein Subjekt ist ein *IRI* oder ein *Blank Node*.
- Ein Prädikat ist ein *IRI*.
- Ein Objekt ist entweder ein *IRI*, ein *Literale* oder ein *Blank Node*.

ein International Resource Identifier (IRI) ist eine generalisierung von URI (als definiert in)... ein Literal ist Blank node leiht die Fähigkeit Subjekt nicht definieren zu müssen was anwendung in RDFa Ansätze findet (siehe ??)

2.2.1. Syntax

Mit der Freigabe von RDF 1.1 wurden, unter anderem, zwei nicht XML-basierten Serialisierungssyntaxen eingearbeitet [vgl.][Abs. 3]Wood:14:WNR. Die Resource-Description-Framework-in-Attributes (RDFa) Syntax und JSON-Linked-Data (JSON-LD) Syn-

²Das Datenmodell basiert auf einer Web-Applikation ([Globmaplab](#)), die für die Arbeit mit den historischen Beständen der Sammlung Perthes konzipiert wurde.

taxen ermöglichen semantische Annotation (vgl. Abschnitt 3 für Definition) in Anwendungsfelder wo es vorher mit klassischen RDF/XML nicht geeignet war.

RDFa macht es möglich Maschine-lesbare Metadaten in Web-Seiten einzubinden indem es neue HTML-Attributen für diesen Zweck festlegt. Diese „Anreicherung“ der Metadaten einer Website führt dazu, dass Fremdsoftware (english: Third Party Applications) in der Lage sind die Metadaten der Webseite automatisch verarbeiten zu können, und dass Suchmaschinen eine gezielter Darstellung des Websiteinhalts für Suchergebnisse anbieten können [vgl.][Abs. 2]Schreiber:14:RP, wobei der letztere Punkt insbesondere dann der Fall ist, wenn die Metadaten an weit verbreitete Ontologien und Vokabulare angeglichen werden (vgl. Abschnitt 2.2.2). Eine RDFa Serialisierungsmöglichkeit für das Datenmodell in Abbildung 1 wurde in Listing 1 veranschaulicht.

Listing 1 Datenmodell in RDFa

JSON-LD kann auch benutzt werden um Webseite semantisch zu annotieren [vgl.]Vincent2015, aber es wurde ursprünglich entwickelt mit der Absicht Interoperable Web-Services für Linked Data (vgl. Definition in Abschnitt 3) auszubauen und Linked Data in JSON-basierte Datenbanksysteme abzuspeichern [vgl.][Abs. 1]Lanthaler:14:J. Angenommen Ein Client schickt eine HTTP GET Anfrage an eine API, die JSON-LD unterstützen würde, würde die Antwort (in Bezug wieder auf das Datenmodel von Abbildung 1) Listing 1 entsprechen.

2.2.2. Semantik

3. RDF und Linked-Data

Eidesstattliche Erklärung

Ich versichere an Eides statt durch meine Unterschrift, dass ich die vorstehende Arbeit selbstständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich oder annähernd wörtlich aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe. Die Arbeit hat in dieser oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

Ilmenau, Tag. Monat Jahr _____
Unterschrift