

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	4
2.2.1 Serialisierung	5
2.2.2 Semantik	6
3 RDF und Linked-Data	7
Literaturverzeichnis	IX
Eidesstattliche Erklärung	XI

Abbildungsverzeichnis

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

Tabellenverzeichnis

Tab. 1 Die Entwicklung von RDF (vgl. Klyne und Carroll, Abs. 2, Schreiber und Raimond, Abs. 2, Wood)	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. Meffert u. a. 2015, S. 4). Das Erbringen dieser Dienstleistungen heutzutage erfordert Geschäftsprozesse die zwischenbetrieblich ausgerichtet sind (vgl. Fleisch 2001, S.19). Im Kontext einer vernetzten Welt (2013, vgl.) 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 Fettke durchgeführt. Eine strukturierte Literatursuche wurde mithilfe der Leitlinien von Webster und Watson, S. XVI 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. Webster und Watson 2002, S. XVI).

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 und Swick 1999). 2004 wurde diese Version aktualisiert und in mehrere Dokumente aufgeteilt (Beckett 2004). Die aktuelle, erweiterte Version (RDF 1.1) wurde 2014 veröffentlicht (Schreiber und Raimond 2014). 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. Das bedeutet, dass kleinere Unternehmen mit vereinfachte JSON basierte Web-APIs Datenintegration auf die syntaktischen und semantischen¹ Ebene RDF wirksam einsetzen betreiben zu können (und mithilfe von Vokabularen teilweise auch auf der semantischen) Ebene von der Im Hinblick auf die semiotische Ebenen der Integration (Schissler u. a. 2004, vgl.) haben Sender und Empfänger mittels RDF eine gemeinsame Sprache, um sich auf der syntaktischen (und mithilfe von Ontologien und Vokabularen auch auf der semantischen) Ebene verständigen zu können.

Anwendungsfall	RDF 1.0	RDF 1.1
Web-Metadaten	RDF/XML	HTML5+RDFa 1.1, JSON-LD
Datenaustausch zw. Datenbanken ²	RDF/XML	JSON-LD, TriG, N-Quads
API-Feeds Verbinden ²	RDF/XML	JSON-LD, RDF/XML

Tab. 1 Die Entwicklung von RDF
(vgl. Klyne und Carroll, Abs. 2, Schreiber und Raimond, Abs. 2, Wood)

¹mithilfe von Vokabularen als definiert in Abschnitt 2.2.2. Datenintegration im Kontext von Ebenen der Integration (Schissler u. a. 2004, vgl.).

²Diese Anwendungsfelder für RDF wurden in Klyne und Carroll, Abs. 2 nicht expliziert gegeben, aber RDF/XML dort zu nutzen ist auch möglich.

2.2. Beschreibung

Laut Schreiber und Raimond in *RDF 1.1 Primer*:

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 der Name vermuten lässt, bietet das Resource Description Framework ein Gerüst (Modell, Sprachen und Syntaxen) für die Beschreibung von Attributen, Funktionen und Beziehungen der Ressourcen. Ressourcen können alles sein, was einen einzigartigen Identifier (URI oder IRI) hat (vgl. De Keyzer u. a. 2013, Folie 6). RDF legt eine abstrakte Syntax fest um Zusammenhänge zwischen Ressourcen als eine Menge gerichteter Graphen darzustellen. 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 Beispiel³ drückt die Beziehung zwischen einem Kartograph und seinen produzierten Werken aus.

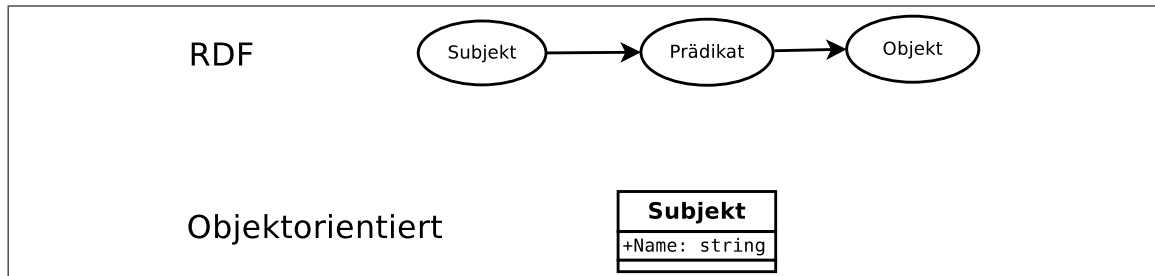


Abb. 1 Kernkonzepte des RDFs anhand eines Beispielgraphs

Die konkrete RDF-Serialisierungssyntaxen werden im Abschnitt 2.2.1 erörtert. Jeder Komponent eines RDF-Tripels kann ein von drei Ausprägungen haben (vgl. Wood u. a. 2014, Abs. 3,1).

- 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 IRI (International Resource Identifier) ist eine generalisierung von URI (als de-

³Das Datenmodell basiert auf einer Web-Applikation ([Globmaplab](#)), die für die Arbeit mit den historischen Beständen der Sammlung Perthes konzipiert wurde.

finiert in)... ein Literal ist Blank node leiht die Fähigkeit Subjekt nicht definieren zu müssen was anwendung in RDFa Ansätze findet (siehe ??)

2.2.1. Serialisierung

Mit der Freigabe von RDF 1.1 wurden, unter anderem, zwei nicht XML-basierten Serialisierungssyntaxen eingearbeitet (vgl. Wood 2014, Abs. 3). Die Resource Description Framework in Attributes (RDFa) Syntax und JSON Linked Data (JSON-LD) Syntaxen 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 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. Schreiber und Raimond 2014, Abs. 2), 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.

```
1 <html>
2   <body>
3     <p>test</p>
4   </body>
5 </html>
```

Listing 1 Datenmodell in RDFa

JSON-LD kann benutzt werden um Webseiten semantisch zu annotieren (Vincent 2015, vgl.), aber es wurde ursprünglich entwickelt mit der Absicht die Interoperabilität von Webservices für Linked Data (vgl. Definition in Abschnitt 3) auszubauen und Linked Data in JSON-basierte Datenbanksysteme abzuspeichern (vgl. Lanthaler u. a. 2014, Abs. 1). 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

Um die .. ist die

3. RDF und Linked-Data

Literaturverzeichnis

Beckett D (2004), *RDF/XML Syntax Specification (Revised)*, W3C Recommendation, <http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/>, W3C.

De Keyzer M, Loutas N, Goedertier S (2013), *Einführung in RDF & SPARQL*.

Fettke DWIP (2006) State-of-the-Art des State-of-the-Art. *Wirtschaftsinformatik* 48(4): 257–266.

Fleisch E (2001) Das Netzwerkunternehmen: Strategien und Prozesse zur Steigerung der Wettbewerbsfähigkeit in der „Networked economy“ Springer-Verlag.

Klyne G, Carroll J (2004), *Resource Description Framework (RDF): Concepts and Abstract Syntax*, W3C Recommendation, <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>, W3C.

Lanthaler M, Sporny M, Kellogg G (2014), *JSON-LD 1.0*, W3C Recommendation, <http://www.w3.org/TR/2014/REC-json-ld-20140116/>, W3C.

Lassila O, Swick RR (1999), *Resource Description Framework (RDF) Model and Syntax Specification*, W3C Recommendation, <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>, W3C.

Meffert H, Bruhn M, Hadwich K (2015) *Dienstleistungsmarketing* Springer Science + Business Media.

Monitoring-Report Digitale Wirtschaft 2013: Digitalisierung und neue Arbeitswelten (2013).

Schissler M, Zeller T, Mantel S (2004), „Überbetriebliche Integration von Anwendungssystemen: Klassifikation von Integrationsproblemen und -lösungen“, In: *Überbetriebliche Integration von Anwendungssystemen* Bartmann D, Mertens P, Elmar J. S (Hrsg), FORWIN-Tagung, Aachen, S. 1–20.

Schreiber G, Raimond Y (2014), *RDF 1.1 Primer*, W3C Note, <http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/>, W3C.

Vincent J (2015), *Video SEO Just Got Better: Meet Wistia's New Standard Embed*.

Webster J, Watson RT (2002) Analyzing the past to prepare for the future: Writing a literature review. *Management Information Systems Quarterly* 26(2): 3.

Wood D (2014), *What's New in RDF 1.1*, W3C Note, <http://www.w3.org/TR/2014/NOTE-rdf11-new-20140225/>, W3C.

Wood D, Lanthaler M, Cyganiak R (2014), *RDF 1.1 Concepts and Abstract Syntax*, W3C Recommendation, <http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>, W3C.

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