

Abschnitt 6

Wissensrepräsentation

Resource Description Framework (RDF)

- Sprache zur Beschreibung von “Ressourcen” insbesondere im WWW
- ursprünglich zur Beschreibung von Webseiten gedacht (Autor, Erstellungsdatum, Lizenz,...)
- Zielt auf **Maschinenlesbarkeit statt Menschenlesbarkeit** ab
- Ressource: beliebige “Dinge”, die eindeutig identifiziert werden können
 - häufig über das Web abrufbar (z.B. Webseiten)
 - aber nicht zwangsläufig (Personen, Unternehmen, Bücher, Kunstwerke,...)

Resource Description Framework (RDF) (2)

- Beschreibung von Ressourcen in der Form einfacher **Subjekt-Prädikat-Objekt-Aussagen**
- Identifizierung von Ressourcen durch **Uniform Resource Identifiers (URIs)**; z.B. eine URL
- Aussagen über Ressourcen nehmen die Form eines Graphen (**RDF-Graph**) an

Beispiel RDF-Graph

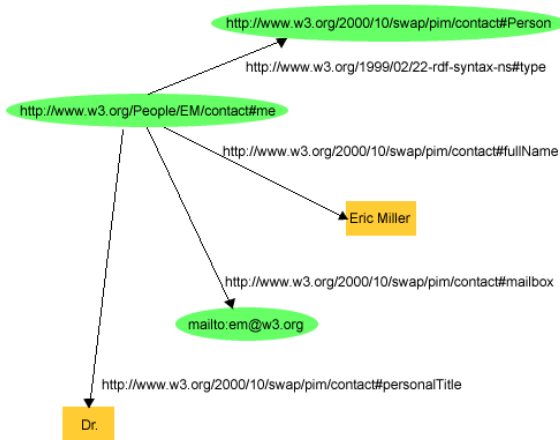


Abbildung: Beschreibung der Person “Eric Miller” als RDF-Graph; Ressourcen sind als Ellipsen dargestellt, Literale als Rechtecke. Bildquelle:

<https://www.w3.org/TR/rdf-primer/>

Uniform Resource Identifiers (URIs)

- Zeichenkette zur konsistenten und eindeutigen Bezeichnung von “Ressourcen”
- Beispiele: Internetseiten, Email-Adressen, Bücher (ISBN-Nummern), Geo-Positionen, aber auch Personen, Kunstwerke, Konzepte
- Zwei Unterkategorien von URIs:
 - URLs (Uniform Resource Locators) bezeichnen eine Ressource über die Zugriffsweg (z.B. Webadresse
`http://www.julielab.de/`)
 - URNs (Uniform Resource Names) bezeichnen eine Ressource durch einen spezifischen Namen. Insbesondere für nicht online verfügbare Ressourcen (z.B. `urn:isbn:0-486-27557-4`)

URI References

- RDF verwendet eine besondere Variante von URIs, sog. **URI References (URIsrefs)**
- Unicode-kompatibel (für viele Sprachen einsetzbar)
- bestehen aus URI und (optional) einem **fragment identifier**.
Diese sind durch ein “#” voneinander abgegrenzt
- Bsp.: `http://www.example.org/index.html#section2`

RDF-Aussagen

- Aussagen in RDF (auch **RDF-Tripel** genannt) bestehen aus Subjekt, Prädikat und Objekt (SPO)
- Subjekt und Prädikat müssen durch URIref bezeichnet sein. Objekte können URIrefs oder alternativ einfache Zeichenketten (**Literale**) sein.
- Ein einfaches Beispiel:

“The website <http://www.example.org/index.html> was created by John Smith.”

Darstellung von RDF-Aussagen (Graph)

“The website **http://www.example.org/index.html** (*Subjekt*) **has a creator** (*Prädikat*) whose value is **John Smith** (*Objekt*).”

http://www.example.org/index.html

http://purl.org/dc/elements/1.1/creator

http://www.example.org/staffid/85740

Darstellung von RDF-Aussagen (Tripel-Notation)

“The website **http://www.example.org/index.html** (*Subjekt*) **has a creator** (*Prädikat*) whose value is **John Smith** (*Objekt*).”

```
<http://www.example.org/index.html>  
<http://purl.org/dc/elements/1.1/creator>  
<http://www.example.org/staffid/85740> .
```

- (eigentlich in eine Zeile zu schreiben)
- SPO werden direkt hintereinander geschrieben, durch Leerzeichen getrennt
- Volle URIs werden durch “< >” umschlossen
- Aussagen enden mit Leerzeichen und Punkt
- Problem: **Aussagen werden schnell sehr lang**
- **QNames** als Lösung (s. nächste Folien)

Nachtrag: XML-Namensräume

- Per Design kann jeder XML-Entwickler seine eigenen XML-Tags definieren.
- Dadurch können Namenskonflikte entstehen, wenn XML-Dokumente verschiedener Quellen kombiniert werden
- Menschen können ggf. sofort erkennen, dass solche Tags unterschiedliche “Dinge” auszeichnen. Für den Computer ist dies zunächst nicht erkennbar. Beispiel:

```
<table >
  <tr >
    <td>Apples </td >
    <td>Bananas </td >
  </tr >
</table >
```

```
<table >
  <name >
    African Coffee Table
  </name >
  <width >80</width >
  <length >120</length >
</table >
```

Nachtrag: XML-Namensräume

Solche Namenskonflikte können durch das Einführen von spezifischen **Präfixen** gelöst werden:

```
<h:table >
```

```
  <h:tr >
```

```
    <h:td>Apples </h:td >
```

```
    <h:td>Bananas </h:td >
```

```
  </h:tr >
```

```
</h:table >
```

```
<f:table >
```

```
  <f:name>African Coffee Table </f:name>
```

```
  <f:width >80</f:width >
```

```
  <f:length >120</f:length >
```

```
</f:table >
```

Nachtrag: XML-Namensräume

- Um Präfixe benutzen zu können, müssen diese einem spezifischen **Namensraum (namespace)** zugeordnet werden.
- Namensräume dienen zur eindeutigen Identifizierung von XML-Elementen und werden durch eine für sie spezifische URI referenziert
- In einem XML-Dokument werden Namensräume durch das Attribut **xmlns** im Start-Tag eines XML-Elements definiert
- Die zu verwendende Syntax ist dabei `xmlns:prefix="URI"`
- Das Präfix eines Namensraums bildet dann zusammen mit einem **lokalen Namen** (das eigentliche XML-Element) einen sog. **qualified name (QName)**
- Wenn ein Namensraum/Präfix in einem XML-Element definiert ist, gilt dieses ebenso für alle Tochter-Elemente.
- Es können mehrere Namensräume/Präfixe in einem Element definiert werden. Dies ist z.B. bei Wurzel-Elementen hilfreich.

Nachtrag: XML-Namensräume

Beispiel mit zwei Namensraumdefinitionen für zwei **unterschiedlichen** XML-Elemente:

```
<root>
  <h:table xmlns:h="http://www.w3.org/TR/html4/">
    <h:tr>
      <h:td>Apples </h:td>
      <h:td>Bananas </h:td>
    </h:tr>
  </h:table>
  <f:table xmlns:f="https://www.w3schools.com/furniture">
    <f:name>African Coffee Table </f:name>
    <f:width>80</f:width>
    <f:length>120</f:length>
  </f:table>
</root>
```

Nachtrag: XML-Namensräume

Beispiel mit zwei Namensraumdefinitionen für **ein** XML-Elemente:

```
<root xmlns:h="http://www.w3.org/TR/html4/"
      xmlns:f="https://www.w3schools.com/furniture">
  <h:table>
    <h:tr>
      <h:td>Apples</h:td>
      <h:td>Bananas</h:td>
    </h:tr>
  </h:table>
  <f:table>
    <f:name>African Coffee Table</f:name>
    <f:width>80</f:width>
    <f:length>120</f:length>
  </f:table>
</root>
```

Darstellung von RDF-Aussagen (Verkürzte Tripel-Notation)

In der Tripel-Notation können vollständige URIs durch QNames mit vorher definierten Präfixen ersetzt werden, um die Zeilenlänge massiv zu reduzieren:

- **ex** für `http://www.example.org/`
- **dc** für `http://purl.org/dc/elements/1.1/creator/`
- **exstaff** für `http://www.example.org/staffid/`

Vorher:

```
<http://www.example.org/index.html>  
<http://purl.org/dc/elements/1.1/creator>  
<http://www.example.org/staffid/85740> .
```

Nacher:

```
ex:index.html dc:creator exstaff:85740 .
```

Darstellung von RDF-Aussagen (RDF/XML)

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">
  <contact:Person rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
    <contact:personalTitle>Dr.</contact:personalTitle>
  </contact:Person>
</rdf:RDF>
```


Darstellung von RDF-Aussagen (Vermischtes)

- In der Tripel-Notation werden vollständige URIs mit “< >” umschlossen, nicht jedoch QNames
- Literale werden in Tripel-Notation in Doppelte Anführungszeichen gesetzt.
- Alle drei Darstellungsarten sind äquivalent und können die gleiche Information darstellen. Die Darstellungsart ist gegenüber dem beschriebenen (formalen) Graphen nachrangig!

Übung

Stellen Sie den unten abgebildeten RDF-Graphen in verkürzter Tripel-Notation dar. Definieren Sie sich zunächst die dafür geeigneten Präfixe.



Übung (Beispiellösung)

```
ericmiller http://www.w3.org/People/EM/contact#  
contact http://www.w3.org/2000/10/swap/pim/contact#  
rdf http://www.w3.org/1990/02/22-rdf-syntax-ns#
```

```
ericmiller:me rdf:type contact:Person .  
ericmiller:me contact:fullName "Eric Miller" .  
ericmiller:me contact:mailbox mailto:em@w3.org .  
ericmiller:me contact:personalTile "'Dr'" .
```

Ausblick

- RDF
 - RDF/XML
 - blank nodes
 - Typen von Literalen
- Suchsprachen für RDF-Tripel
- Terminologien und SKOS
- Beschreibungslogik
- Ontologien und OWL
- “Semantisches Rechnen”
- Semantic Web

Literatur zur Wissenrepräsentation

- <https://www.w3.org/TR/rdf-primer/>
- <https://www.w3.org/TR/webarch/#identification>
- <https://t3n.de/news/url-uri-unterschiede-516483/>
- <https://www.w3.org/TR/rdf-primer/#documents>
- https://www.w3schools.com/xml/xml_namespaces.asp