



schema.org und Bibliotheken

Struktur

2

- Schema.org Basics
- Reaktionen auf Schema.org
- Pflege und Weiterentwicklung
- Schema Bib Extend Community Group

Schema.org

Was ist Schema.org?

4

- Start: Juni 2011
- Gemeinsame Initiative von Google, Microsoft, Yahoo! und (seit November 2011) Yandex
- Ziel: Anreicherung von Suchergebnissen mit strukturierten Daten (Rich Snippets)
- Mittel: Entwicklung eines geteilten Vokabulars für die Anreicherung von Webseiten mit strukturierten Daten

Rich Snippets

5

[Vegetarian Chilli Recipe - Taste.com.au](#)



www.taste.com.au/recipes/5923/vegetarian+chilli ▼

★★★★★ Rating: 4.5 - 52 reviews

Vegetarian chilli **recipe** - Place 2 tablespoons of oil in a large saucepan over medium ... This is a staple in our home which we have renamed **chilli sin carne**.

[Fat Freddy's Drop Tour Dates 2013 — Fat Freddy's Drop Concert ...](#)

www.songkick.com/artists/282353-fat-freddys-drop ▼

Find **Fat Freddy's Drop** live concert **tour dates**, tickets, reviews, and more on Songkick. Be the first to know when **Fat Freddy's Drop** is playing live in your town!

Fri, Dec 6 [The Bedford Bigtop](#)
Sun, Dec 29 [Coroglen Tavern](#)
Mon, Dec 30 [Brewers Field](#)

[RESTful Web APIs - O'Reilly Media](#)

shop.oreilly.com/product/0636920028468.do ▼

★★★★★ Rating: 4 - 2 reviews

A proof of this approach are the first three chapters that talk more about HTTP and how it works more than **RESTful API**. This makes the **book** very good for ...

Schema.org - Verbreitung I

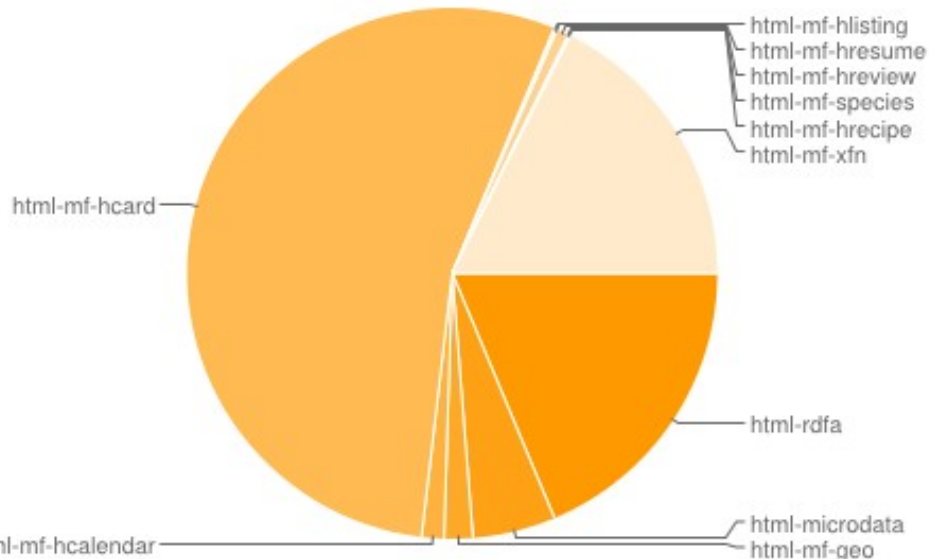
6

- Untersuchung von Bizer et al. 2013
- Basis: Common-Crawl-Daten von August 2012
- Strukturierte Daten (Microformats, Microdata, RDFa) in Common Crawl
- 369 M. von 3 Mrd. URLs (**12,3 %**)
- 2 286 277 von 40,600,000 Domains (**5,64%**)

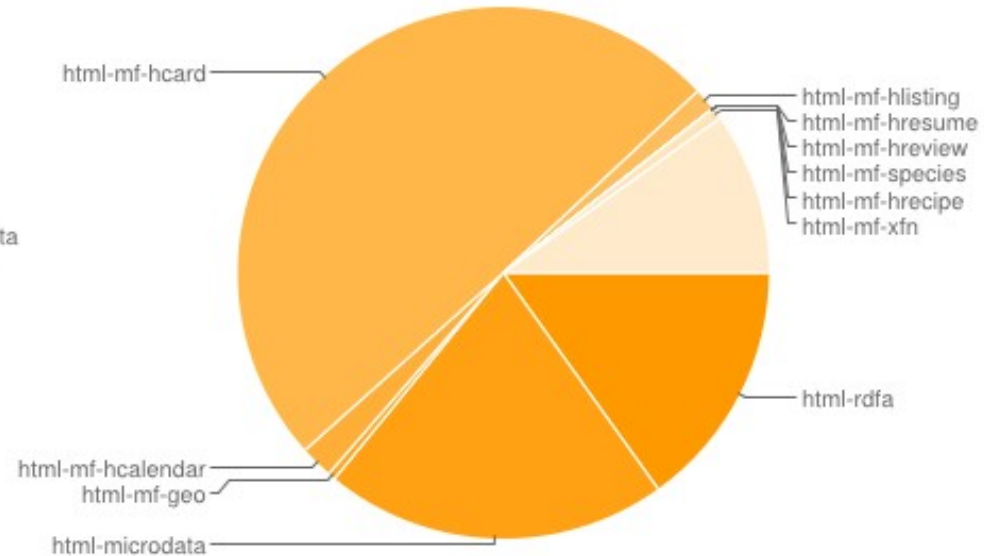
Schema.org - Verbreitung II

7

Domains with Triples



Triples



Quelle: <http://webdatacommons.org/2012-08/stats/stats.html>

**Achtung: Nicht alles
Microdata = schema.org
aber der größte Teil!**

(siehe <http://webdatacommons.org/2012-08/stats/stats.html> für Details)

Schema.org - Vokabular

9

- Zielgruppe: Webmaster, Webanwendungsentwickler
- Umfasst mehr als 300 **Klassen/Typen** (Person, Product, Book, Organization etc.) und ca. 270 **Properties** (openingHours, creator, ...)
- Klassenhierarchie wie bei einem RDF-Vokabular
 - Oberste Ebene: schema:Thing
 - Properties werden auf untergeordnete Typen vererbt

Ein pragmatischer Ansatz

10

„Not asking an unconstrained 'so, how do we describe cars?', but 'how can we improve markup on existing pages that describe cars?' (or Comics, SoftwareApps, Sports, ...)”

Dan Brickley (2012)

Thing

The most generic type of item.

Property	Expected Type	Description
Properties from Thing		
additionalType	URL	An additional type for the item, typically used for adding more specific types from external vocabularies in microdata syntax. This is a relationship between something and a class that the thing is in. In RDFa syntax, it is better to use the native RDFa syntax - the 'typeof' attribute - for multiple types. Schema.org tools may have only weaker understanding of extra types, in particular those defined externally.
description	Text	A short description of the item.
image	URL	URL of an image of the item.
name	Text	The name of the item.
sameAs	URL	URL of a reference Web page that unambiguously indicates the item's identity. E.g. the URL of the item's Wikipedia page, Freebase page, or official website.
url	URL	URL of the item.

More specific types

- [Action](#)
- [Class](#)
- [CreativeWork](#)
- [Event](#)
- [Intangible](#)
- [MedicalEntity](#)
- [Organization](#)
- [Person](#)
- [Place](#)
- [Product](#)
- [Property](#)

Thing > CreativeWork > Book

A book.

Property	Expected Type	Description
Properties from Thing		
additionalType	URL	An additional type for the item, typically used for adding more specific types from external vocabularies in microdata syntax. This is a relationship between something and a class that the thing is in. In RDFa syntax, it is better to use the native RDFa syntax - the 'typeof' attribute - for multiple types. Schema.org tools may have only weaker understanding of extra types, in particular those defined externally.
description	Text	A short description of the item.
image	URL	URL of an image of the item.
name	Text	The name of the item.
sameAs	URL	URL of a reference Web page that unambiguously indicates the item's identity. E.g. the URL of the item's Wikipedia page, Freebase page, or official website.
url	URL	URL of the item.
Properties from CreativeWork		
about	Thing	The subject matter of the content.
accountablePerson	Person	Specifies the Person that is legally accountable for the CreativeWork.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
alternativeHeadline	Text	A secondary title of the CreativeWork.
associatedMedia	MediaObject	The media objects that encode this creative work. This property is a synonym for encodings.
audience	Audience	The intended audience of the item, i.e. the group for whom the item was created.
audio	AudioObject	An embedded audio object.
author	Organization or Person	The author of this content. Please note that author is special in that HTML 5 provides a special mechanism for indicating authorship via the rel tag. That is equivalent to this and may be used interchangeably.
award	Text	An award won by this person or for this creative work.
awards	Text	Awards won by this person or for this creative work. (legacy spelling; see singular form, award)
citation	CreativeWork or Text	A citation or reference to another creative work, such as another publication, web page, scholarly article, etc. NOTE: Candidate for promotion to ScholarlyArticle.
comment	UserComments	Comments, typically from users, on this CreativeWork.
contentLocation	Place	The location of the content.
contentRating	Text	Official rating of a piece of content—for example, 'MPAA PG-13'.
contributor	Organization or Person	A secondary contributor to the CreativeWork.
copyrightHolder	Organization or Person	The party holding the legal copyright to the CreativeWork.
copyrightYear	Number	The year during which the claimed copyright for the CreativeWork was first asserted.
creator	Organization or Person	The creator/author of this CreativeWork or UserComments. This is the same as the Author property for CreativeWork.

Unterstützte Standards

13

- Microdata
- RDFa (seit Ende 2011)
- JSON-LD (seit Juni 2013)

The Catcher in the Rye - Mass Market Paperback

by J.D. Salinger

4 stars - 3077 reviews

Product details

224 pages

Publisher: Little, Brown, and Company - May 1, 1991

Language: English

ISBN-10: 0316769487

...Microdata-Anreicherung

```
<div itemscope itemtype="http://schema.org/Book">

<span itemprop="name">The Catcher in the Rye</span> -
  <link itemprop="bookFormat" href="http://schema.org/Paperback">Mass
Market Paperback
by <a itemprop="author" href="/author/jd_salinger.html">J.D.
Salinger</a>
```

```
<div itemprop="aggregateRating" itemscope
itemtype="http://schema.org/AggregateRating">
  <span itemprop="ratingValue">4</span> stars -
  <span itemprop="reviewCount">3077</span> reviews
</div>
```

Product details

```
<span itemprop="numberOfPages">224</span> pages
Publisher: <span itemprop="publisher">Little, Brown, and Company</span>
-
  <meta itemprop="datePublished" content="1991-05-01">May 1, 1991
Language: <span itemprop="inLanguage">English</span>
ISBN-10: <span itemprop="isbn">0316769487</span>
```



```
<div itemscope itemtype="http://schema.org/Book">

<span itemprop="name">The Catcher in the Rye</span> -
  <link itemprop="bookFormat" href="http://schema.org/Paperback">Mass
Market Paperback
by <a itemprop="author" href="/author/jd_salinger.html">J.D.
Salinger</a>
```

```
<div itemprop="aggregateRating" itemscope
itemtype="http://schema.org/AggregateRating">
  <span itemprop="ratingValue">4</span> stars -
  <span itemprop="reviewCount">3077</span> reviews
</div>
```

Product details

```
<span itemprop="numberOfPages">224</span> pages
Publisher: <span itemprop="publisher">Little, Brown, and Company</span>
-
  <meta itemprop="datePublished" content="1991-05-01">May 1, 1991
Language: <span itemprop="inLanguage">English</span>
ISBN-10: <span itemprop="isbn">0316769487</span>
```

...RDFa-Anreicherung

```
<div vocab="http://schema.org/" typeof="http://schema.org/Book">
```

```

```

```
<span property="name">The Catcher in the Rye</span> -
```

```
<link property="bookFormat" href="http://schema.org/Paperback">Mass  
Market Paperback
```

```
by <a property="author" href="/author/jd_salinger.html">J.D.  
Salinger</a>
```

```
<div property="aggregateRating"
```

```
typeof="http://schema.org/AggregateRating">
```

```
<span property="ratingValue">4</span> stars -
```

```
<span property="reviewCount">3077</span> reviews
```

```
</div>
```

Product details

```
<span property="numberOfPages">224</span> pages
```

```
Publisher: <span property="publisher">Little, Brown, and Company</span>
```

-

```
<meta property="datePublished" content="1991-05-01">May 1, 1991
```

```
Language: <span property="inLanguage">English</span>
```

```
ISBN-10: <span property="isbn">0316769487</span>
```

```
</div>
```

**...der Vollständigkeit
halber:**

JSON-LD

```
{
  "@context":
    {
      "@vocab": "http://schema.org/"
    },
  "@type": "Book",
  "publisher": "Little, Brown, and Company",
  "bookFormat": "Paperback",
  "numberOfPages": "224",
  "inLanguage": "English",
  "aggregateRating":
    {
      "@type": "AggregateRating",
      "ratingValue": "4",
      "reviewCount": "3077"
    },
  "isbn": "0316769487",
  "author": "/author/jd_salinger.html",
  "name": "The Catcher in the Rye",
  "datePublished": "1991-05-01",
  "image": "catcher-in-the-rye-book-cover.jpg"
}
```

Reaktionen

Kritik...

23

- Kritik an Entstehungsprozess: nicht offen und community-driven
- Zentrales, allumfassendes Vokabular als Konkurrenz zu verteiltem LOD-Ansatz und existierenden Vokabularen
- Kontrolle liege bei kommerziellen Unternehmen (Stewart: „Land Grab“)
- Gradmann: „Ontologischer Darwinismus“ & „Zensur“, „oligopolistisches Eigentum“ an den Ausdrucksmitteln im Web

...und Hoffnung

24

- schema.org als „Linked Data Lite“ für die Massen:
- schema.org als One-Stop-Shop für Webmaster, die mit LOD bisher nichts zu tun hatten
- Keine steile Lernkurve wie bei LOD
- „Google and the search engine triumvirate understand well — much better than many of the researchers and academics that dominate mailing list discussions — that use and adoption trump elegance and sophistication.“ (Bergman 2011)
- Ergebnis: schema.org bringe dem Web mehr strukturierte Daten als Semweb LOD in Jahren, wovon auch die LOD-Community profitiere

Pflege und Weiterentwicklung

Web Schemas Task Force

26

- Web Schemas Task Force der W3C Semantic Web Interest Group (SWIG)
 - Chairs: R.V.Guha (Google) and Dan Brickley (Google)
 - Wiki: <http://www.w3.org/wiki/WebSchemas>
 - Mailingliste: <http://lists.w3.org/Archives/Public/public-vocabs/>
 - Issue Tracker: <http://www.w3.org/2011/webschema/track/>
- Allgemeine Diskussion zu Metadatenschemata mit Fokus auf schema.org:

„The Web Schemas Task Force is devoted to practical issues around data schemas for large-scale use in the public Web.“

schema.org-Erweiterungen

27

- Prozess:
 - Ziel: grober Konsens und inkrementelle Verbesserungen
 - Realistische Beispiele, Einfachheit für Publisher und möglichst Wiederverwendung des aktuellen Vokabulars
- Liste der (vorgeschlagenen) Erweiterungen:
<http://www.w3.org/wiki/WebSchemas/SchemaDotOrgProposals>
- Beispiele: LRMI, Schema für Stellenangebote, Medical/Health Proposal, ...

Schema Bib Extend Community Group

Allgemeines

29

- Gegründet: September 2012
- Chair: Richard Wallis (OCLC)
- Ziel: Erweiterung von schema.org zum Teilen von -
im weitesten Sinne - bibliographischen
Informationen
- Mailingliste: <http://lists.w3.org/Archives/Public/public-schemabibex/>
- Wiki: http://www.w3.org/community/schemabibex/wiki/Main_Page

Vorschläge/Entwürfe

30

Proposals

- ~~Work-Instance~~ - Renamed to CreativeWork Relationships
- [CreativeWork Relationships](#)
- [Identifier](#)
- [Identifier-2](#)
- [commonEndeavor](#)
- [Content-Carrier](#)
- [citation](#) - **Accepted**
- [Audiobook](#) - Proposed
- [Collection](#) - Proposed
- [Parts](#)
- [Holdings](#) - alternatively, rather than extend, [Holdings via Offer](#)

Content-Carrier

31

- Verwendung der Product Types Ontology
 - URIs für Produkttypen basierend auf Wikipedia
- Beispiel:

```
@prefix schema: <http://schema.org/> .
```

```
@prefix schemap: <http://proposed-schema.org/> .
```

```
@prefix pto: <http://www.productontology.org/id/> .
```

```
[ ] a schema:Book, pto:Audiobook, pto:Compact_Disk;  
    schema:name "Dune";  
    schema:author <http://viaf.org/viaf/59083797>;  
    schema:isbn "1427201439";  
    schemap:sameThingAs <http://www.worldcat.org/oclc/153312079> .
```

Beispieltitelbeschreibung

32

```
@prefix schema: <http://schema.org/> .
@prefix pto: <http://www.productontology.org/id/> .

<http://lobid.org/resource/HT014576567>
  a schema:Book, pto:Hardcover ;
  schema:name "System und Performanz : symboltheoretische Grundlagen
von Medientheorie und Sprachwissenschaft" ;
  schema:publisher "Velbrück Wiss." ;
  schema:numberOfPages "337 S." ;
  schema:creator <http://d-nb.info/gnd/115875549> ;
  schema:datePublished "2005" ;
  schema:inLanguage <http://id.loc.gov/vocabulary/iso639-2/deu> ;
  schema:about <http://d-nb.info/gnd/4054498-9>,
<http://d-nb.info/gnd/4074250-7>, <http://d-nb.info/gnd/4128325-9> ;
  schema:bookEdition "1. Aufl." ;
  schema:isbn "3938808004" ;
  schema:gtin13 "9783938808009" .
```


“Holdings as Offer”-Ansatz

33

```
@prefix schema: <http://schema.org/> .
@prefix gr: <http://purl.org/goodrelations/v1#> .

<http://lobid.org/item/HT014576567%3AHWB25011>
  a schema:Offer ;
  schema:businessFunction gr:LeaseOut ;
  schema:sku "HWB25011" ;
  schema:itemOffered <http://lobid.org/resource/HT014576567> ;
  schema:seller <http://lobid.org/organisation/DE-294> .
```

Ausstehende Arbeiten

34

- Verlinkung von Werken, Auflagen, Exemplaren: CreativeWork relationships proposal und commonEndeavour
- Viele kleine Ergänzungen

Schema.org in the wild

35

- <http://www.worldcat.org>
- Open-Source-Bibliothekssysteme Koha & Evergreen:
siehe Dan Scott (2013)
-?....

Danke.

Fragen?

**Gerne jetzt oder auch später an
pohl@hbz-nrw.de oder @acka47.**

Referenzen

37

Phil Archer (2012): Danbri has moved on – should we follow?

Mike Bergman (2011): Structured Web Gets Massive Boost

Christian Bizer, Kai Eckert, Robert Meusel, Hannes Mühleisen, Michael Schuhmacher and Johanna Völker (2013):

Deployment of RDFa, Microdata, and Microformats on the Web – A Quantitative Analysis. The 12th International Semantic Web Conference (ISWC2013), Oct. 2013

Dan Brickley (2012): Schema.org and One Hundred Years of Search

Stefan Gradmann (2013): Semantische Nachhaltigkeit und Kontrolle: Gedanken zu schema.org und Linked Open Data. Vortrag am 31.5.2013 bei der Abschlussveranstaltung der 8. Initiative "Nachhaltigkeit in der digitalen Welt". (Folien, Aufzeichnung)

Dan Scott (2013): RDFa and schema.org all the library things

Manu Sporny (2011): The False Choice of Schema.org

Darin Stewart (2011): Schema.org: Webmaster One-Stop or Linked Data Land Grab?

Lizenz

38



Diese Folien stehen unter einer
Creative-Commons-Lizenz:

<http://creativecommons.org/licenses/by/3.0/de/>