

WEB FOR HUMANS

Chuck Norris is born March 10, 1940 in Ryan, Oklahoma, US.

Chuck Norris knows Bruce Lee.

Chuck Norris knows karate.

Bruce Lee knows kung-fu.

TRIPLETS

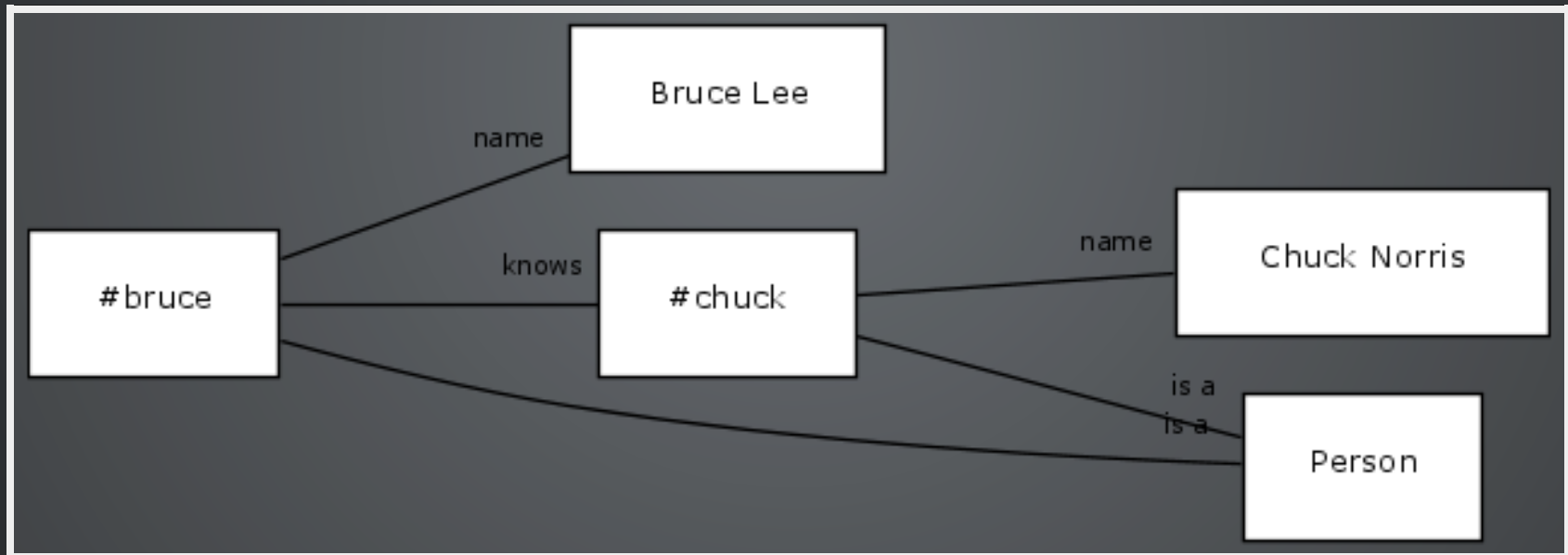
subject predicate object

Chuck is a Person
Chuck name Chuck Norris
Chuck knows Bruce

WEB FOR MACHINES

graph

Facebook: Open Graph
Google: Knowledge Graph



LINKED DATA

<http://linkeddata.org/>

Linked Data is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods. More specifically, Wikipedia defines Linked Data as "a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF."

URIS

GENERIC MEANS TO IDENTIFY ENTITIES OR CONCEPTS IN THE WORLD

SUBJECT PREDICATE OBJECT

[HTTP://EXAMPLE.COM/CHUCK](http://example.com/chuck)

[HTTP://VOCABULARY.ORG/TYPE](http://vocabulary.org/type)

[HTTP://SCHEMA.COM/PERSON](http://schema.com/person)

LINKED DATA - RDF

The Resource Description Framework is a family of W3C specifications originally designed as a metadata data model. It has come to be used as a general method for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats.

1. abstract format
2. serializations (RDF/XML, JSON-LD, N-Triples, Turtle...)

Metadata delivered separately from web page
RDFa - metadata embedded into markup

Not widely used, slow adoption because of complexity and low incentive

SCHEMA.ORG

Joint effort by Bing, Google, Yahoo! and Yandex

The principles are the same:

- vocabulary / ontology
- triplets
- microdata / RDFa lite (experimental)

SCHEMA.ORG - MICRODATA

```
<div itemscope itemtype = "http://schema.org/Movie">
  <h1 itemprop="name">Avatar</h1>
  <div itemprop="director" itemscope itemtype="http://schema.o
rg/Person">
    Director: <span itemprop="name">James Cameron</span> (born <
span itemprop="birthDate">August 16, 1954)</span>
  </div>
  <span itemprop="genre">Science fiction</span>
  <a href=" ../movies/avatar-theatrical-trailer.html" itemprop=
"trailer">Trailer</a>
</div>
```

Metadata is embedded into markup of web page

SCHEMA.ORG - VOCABULARY

- Creative works: [CreativeWork](#) [Book](#) [Movie](#) [MusicRecording](#) [Recipe](#) [TVSeries](#) ...
- Embedded non-text objects: [AudioObject](#) [ImageObject](#) [VideoObject](#)
- [Event](#)
- [Organization](#)
- [Person](#)
- [Place](#) [LocalBusiness](#) [Restaurant](#) ...
- [Product](#) [Offer](#) [AggregateOffer](#)
- [Review](#) [AggregateRating](#)

Banana Banana Bread Recipe - Allrecipes.com



[allrecipes.com/recipe/banana-banana-bread/](https://www.allrecipes.com/recipe/banana-banana-bread/)

★★★★★ Rating: 4.6 - 7,836 reviews - 1 hr 20 mins - 229 cal

The excerpt from the page will show up here. The reason we can't show text from your webpage is because the text depends on the query the user types.

A voc

Quick adoption because of clear incentive.

SCHEMA.ORG - VOCABULARY

Thing > Person

A person (alive, dead, undead, or fictional).

Property	Expected Type	Description
Properties from Thing		
<u>additionalType</u>	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.
<u>alternateName</u>	Text	An alias for the item.
<u>description</u>	Text	A short description of the item.
<u>image</u>	URL	URL of an image of the item.
<u>name</u>	Text	The name of the item.
<u>sameAs</u>	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.
<u>url</u>	URL	URL of the item.
Properties from Person		
<u>additionalName</u>	Text	An additional name for a Person, can be used for a middle name.
<u>address</u>	PostalAddress	Physical address of the item.
<u>affiliation</u>	Organization	An organization that this person is affiliated with. For example, a school/university, a club, or a team.
<u>alumniOf</u>	EducationalOrganization	An educational organizations that the person is an alumni of.
<u>award</u>	Text	An award won by this person or for this creative work.
<u>awards</u>	Text	Awards won by this person or for this creative work. (legacy spelling; see singular form, award)

SCHEMA.ORG AND EZ PUBLISH

Microdata should correctly reflect the
primary content of the page

Mapping

For some part it is simple:

type -> ContentType

type property -> FieldDefinition (values and relations)

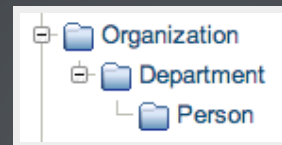
SCHEMA.ORG AND EZ PUBLISH

Mapping that is not so simple

- relationships defined by hierarchy
- enumeration properties
- web page specific stuff

SCHEMA.ORG AND EZ PUBLISH

Relationships defined by hierarchy



1. organisation has a department
2. person works in a department
3. person is an employee of an organisation

SCHEMA.ORG AND EZ PUBLISH

Enumeration properties

```
<title>Kenmore White 17" Microwave</title>

<div itemscope itemtype="http://schema.org/Product">
  <span itemprop="name">Kenmore White 17" Microwave</span>

  <div itemprop="offers" itemscope itemtype="http://schema.org/Offer">
    <span itemprop="price">$55.00</span>
    <meta itemprop="priceCurrency" content="USD" />
    <link itemprop="availability" href="http://schema.org/InStock"/>In stock
  </div>

  <div itemprop="aggregateRating" itemscope itemtype="http://schema.org/AggregateRating">
    Rated <span itemprop="ratingValue">2.5</span>
    /<span itemprop="bestRating">5</span>
    based on <span itemprop="reviewCount">2</span> customer reviews
  </div>
</div>
```

SCHEMA.ORG AND EZ PUBLISH

Web page specific stuff

```
<title>Breadcrumb Books</title>

<body itemscope itemtype="http://schema.org/WebPage">
...
<div itemprop="breadcrumb">
  <a href="category/books.html">Books</a>
  <a href="category/books-literature.html">Literature & Fiction</a>
  <a href="category/books-classics">Classics</a>
</div>
```

[Breadcrumb Books](#)

[Books](#) › [Literature & Fiction](#) › [Classics](#) - [Cached](#)

The excerpt from the page will show up here. The reason we can't show text from your webpage is because the text depends on the query the user types.

SCHEMA.ORG AND EZ PUBLISH

RichText

- schema needs to be extended to allow microdata markup

SCHEMA.ORG AND EZ PUBLISH

RichText embeds

- embedded Content renders its own microdata markup
- there needs to be a way to connect embedded Content with primary Content of the page

SCHEMA.ORG AND EZ PUBLISH

RichText embeds

```
<!-- PRIMARY CONTENT -->
<div itemscope itemtype="http://schema.org/Person">
  <span itemprop="name">Jane Doe</span>
  
  <span itemprop="jobTitle">Professor</span>
<br>
  <!-- START EMBEDDED CONTENT -->
  <div <span style="background-color: rgb(255, 0, 0);">itemprop="address"</span>
    <span itemprop="streetAddress">
      20341 Whitworth Institute
      405 N. Whitworth
    </span>
    <span itemprop="addressLocality">Seattle</span>,
    <span itemprop="addressRegion">WA</span>
    <span itemprop="postalCode">98052</span>
  </div>
  <span style="font-family: monospace; font-size: 17px; font-style: normal; font-weight: normal;">
    <span itemprop="telephone">(425) 123-4567</span>
    <a href="mailto:jane-doe@xyz.edu" itemprop="email">jane-doe@xyz.edu</a>
  </div><br>
```

Schema.org and eZ Publish

RichText custom/template tags (widgets)

- should be similar to embeds

Standard tags

- schema should allow metadata on these
- too much freedom - possibility to create nonsense

SCHEMA.ORG AND EZ PUBLISH

Implementation

- possible to model vocabulary and relationships separately from our core structures
- not in core but isolated in a bundle - keep APIs simple
- open for custom vocabularies
- possible to implement custom resources (RDF/XML, JSON-LD) to more easily connect to existing semantic tools (Apache Stanbol)
- template helpers to handle rendering of microdata

SCHEMA.ORG AND EZ PUBLISH

Implementation questions

- how to model hierarchy relationships (if at all)
- how to model webpage specific stuff