



# Semantic Technologies

Juan David Ocampo, Juan Saavedra, Ziad Elharairi and Sahitya Chaudhary



# Books RDFa

```
<div typeof="Book" resource="2">
```

```
<h1>
```

```
  <span property="name">Head First Data Analysis: A learner's guide to big numbers, statistics,  
  and good decisions</span>
```

```
</h1>
```

```
  <div>Author: <span property="author"></span></div>
```

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

```
  <div>ISBN <span property="isbn">978-0596153939</span></div>
```

```
  <div>Publisher <span property="publisher">O'Reilly Media</span></div>
```

```
  <div>Date Published <span property="datePublished">August 18, 2009</span></div>
```

```
  <div>Number of Pages <span property="numberOfPages">484</span></div>
```

```
  <div>Average rating on Amazon <span property="contentRating">4.3</span></div>
```

```
  <div>Book Edition <span property="bookEdition"> 1st edition </span></div>
```

```
</div>
```

# RDFa validator

Book		0 ERRORI 0 AVVISI ^
ID: https://example-test.site/2		
@type	Book	
@id	https://example-test.site/2	
name	Head First Data Analysis: A learner's guide to big numbers, statistics, and good decisions	
inLanguage	English	
isbn	978-0596153939	
datePublished	2009-08-18	
numberOfPages	484	
contentRating	4.3	
bookEdition	1st edition	
publisher		
@type	Thing	
name	O'Reilly Media	

# Books Turtle

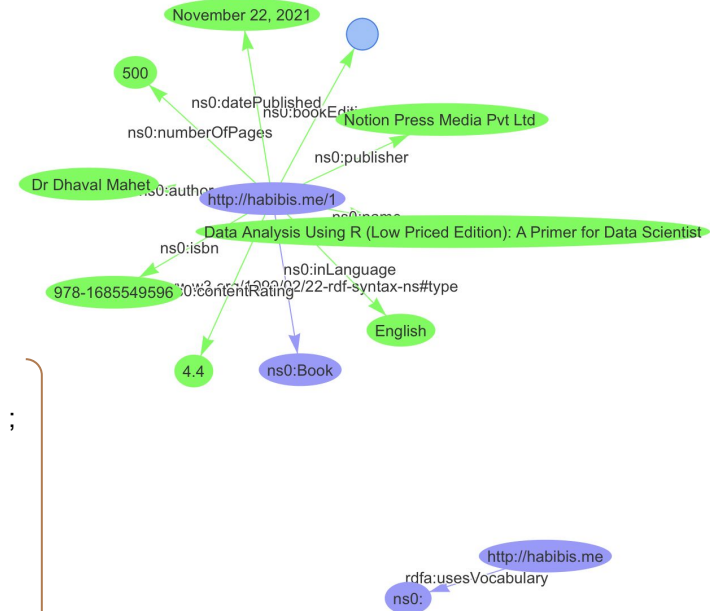
## Structure

<[https://www.amazon.com/Data-Analysis-Using-Low-Priced/dp/1685549594/ref=sr\\_1\\_16?keywords=data+analysis&qid=1671164640&sr=8-16](https://www.amazon.com/Data-Analysis-Using-Low-Priced/dp/1685549594/ref=sr_1_16?keywords=data+analysis&qid=1671164640&sr=8-16)>

ns0

### Book

**name** "Data Analysis Using R (Low Priced Edition): A Primer for Data Scientist" ;  
**author** " Dr Dhaval Mahet" ;  
**inLanguage** "English" ;  
**isbn** "978-1685549596" ;  
**publisher** "Notion Press Media Pvt Ltd " ;  
**datePublished** "November 22, 2021" ;  
**numberOfPages** "500" ;  
**contentRating** "4.4" ;  
**bookEdition** "" .



# Query: Get authors and book names

**PREFIX** rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

**PREFIX** rdfs: <http://www.w3.org/2000/01/rdf-schema#>

**PREFIX** ns0: <https://schema.org/>

**SELECT** ?author ?name

**WHERE** {

?x a ns0:Book.

?x ns0:name ?name; ns0:author ?author

} **LIMIT** 2



Author	Name
""	"Data Analytics, Data Visualization & Communicating Data: 3 books in 1"
"Benjamin Bengfort, Rebecca Bilbro"	"Applied Text Analysis with Python"

# Query: Get the Book names and IsBn which have more than 500 pages

**PREFIX** : <http://example.org/ontology/>  
**PREFIX** rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
**PREFIX** rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
**PREFIX** ns0: <https://schema.org/>  
**PREFIX** xsd: <http://www.w3.org/2001/XMLSchema#>

**SELECT** ?bookName ?isbn  
**WHERE** {  
 {  
 ?x ns0:name ?bookName; ns0:numberOfPages ?stringValue.  
 **OPTIONAL** {?x ns0:isbn ?isbn.}  
 **BIND**(xsd:integer(?stringValue) AS ?integerValue)  
 }  
 **FILTER** (?integerValue > 500)  
} **LIMIT** 2



BookName	IsBn
Time Series Analysis with Python Cookbook	978-1801075541
Python Data Science Handbook	978-1098121228

# Query: Get the books with a content rating greater than 4 and the ones published by Pearson

**PREFIX** rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

**PREFIX** rdfs: <http://www.w3.org/2000/01/rdf-schema#>

**PREFIX** ns0: <https://schema.org/>

**PREFIX** xsd: <http://www.w3.org/2001/XMLSchema#>

**SELECT** ?bookName ?integerValue

**WHERE** {

{  
 {  
 ?book a ns0:Book; ns0:name ?bookName; ns0:publisher ?P.  
 ?book ns0:contentRating ?stringValue.  
 **BIND**(xsd:double(?stringValue) AS ?integerValue)}  
 **FILTER**(?integerValue > 4)  
 }  
}



**UNION**

{  
 ?book a ns0:Book; ns0:name ?bookName; ns0:publisher ?P.  
 ?book ns0:contentRating ?stringValue.  
 ?book ns0:publisher "Pearson".  
}

**LIMIT** 2

BookName	IntegerValue
Data Analytics, Data Visualization & Communicating Data	4.6
Data Science for Supply Chain Forecasting	4.4

# Query: Retrieve the book names and number of pages that have 'SQL' in their name

**PREFIX** rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

**PREFIX** rdfs: <http://www.w3.org/2000/01/rdf-schema#>

**PREFIX** ns0: <https://schema.org/>

**PREFIX** xsd: <http://www.w3.org/2001/XMLSchema#>

**SELECT** ?BookName ?NumPages

**WHERE** {

  ?s ns0:name ?BookName; ns0:numberOfPages ?NumPages .

**FILTER**( **REGEX**( xsd:string(?BookName ), "SQL" ) )

}

**LIMIT** 2



BookName	NumPages
SQL QuickStart Guide	249
SQL for Data Analysis	360



# Query: Get the average rating by each publisher

**PREFIX** rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

**PREFIX** rdfs: <http://www.w3.org/2000/01/rdf-schema#>

**PREFIX** ns0: <https://schema.org/>

**PREFIX** xsd: <http://www.w3.org/2001/XMLSchema#>

**SELECT** ?publisher (AVG(?integerValue) AS ?AVGINTValue)

**WHERE** {

?book a ns0:Book; ns0:publisher ?publisher; ns0:contentRating ?stringValue.

**FILTER NOT EXISTS** { ?book ns0:contentRating ""; ns0:publisher "" . }

**BIND**( xsd:double( ?stringValue ) AS ?integerValue )

}

**GROUP BY** ?publisher

**ORDER BY DESC**(?AVGINTValue)

**LIMIT** 2



<b>publisher</b>	<b>AVGINTValue</b>
"Cambridge University Press"	5.0e0
"Holy Macro! Books "	5.0e0