

Embedding Microdata in Library Online Services for Better Discovery of Content

Vinit Kumar

Assistant Professor
Bundelkhand University, Jhansi, U.P., India
mailvinitkumar@gmail.com

Abstract. HTML5 has seen the day of light and with this the much awaited microdata mechanism has also become a standard. Browsers conventionally use the HTML tags for better readability of webpages for humans. There are cases when it is felt to provide structures to the HTML pages and the inability of HTML tags defined in the HTML specifications to handle such structure. The new specifications of HTML, HTML5 provide a very good mechanism to define structure in webpages. The mechanism popularly known as microdata is easier to implement than the other formats such as, microformats and RDFa. No doubt these structures are machine-readable providing the machines/software programs to search for specific types of information. Recently, some of the popular search engines such as Google, Yahoo and Bing joined hands to support microdata. This collaboration resulted in introduction of Schema.org, a collection of terms that webmasters can use to markup their pages to improve the display of search results. This paves a way for libraries to also augment and implement their web pages with microdata. The present article focuses on the techniques and elements used to create a semantic structure for a webpage using microdata. The paper starts by discussing various efforts going on in the arena of Microdata and microformats etc. then it puts forth the ways by which libraries can use it. The various available tools, and the know how for publishing microdata is also discussed in detail keeping in mind the libraries will be benefited at large.

Keywords: Metadata, Microdata, Schema.org, Rich Snippets, microformats

1. Background

Libraries and librarians in their day-to-day activities involve in creation and maintenance of content. With the arrival of Web and its growing popularity libraries started providing online access to the content available in libraries. The Web made up of HTML pages was not enough to provide relevant content to the users. Libraries also started

providing metadata with the information objects. The non-availability of a mechanism to provide structure to webpages was a main factor for the arrival of metadata. The various markup languages ranging from SGML¹(structured General Markup Language), HTML² (Hypertext Markup Language) and the XML³ (eXtensible Markup Language) were developed to provide structure to the documents on the web. The XML is used to define the metadata standards. Metadata is data about data. This definition of metadata defines the basic concept of metadata but not very meaningful, it refers to everything and nothing. NISO (NISO, 2004) defines metadata as "Structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource" similarly ALA Committee on Cataloguing: Description and Access (CC: DA) task force on metadata defines it as "Structured, encoded data that describe characteristics of information-bearing entities to aid in the identification, discovery, assessment, and management of the described entities"(CC:DA Task Force on Metadata, 2000). The above definitions underline the importance to have structure, description and management as an aspect of metadata. The World Wide Consortiums Tim Berners-Lee defines the metadata in very crude sense as "Machine-understandable information about Web resources or other things"(Berners-Lee, 1997).

The provision of information about the web resources helps in resource discovery, organizing, interoperability and digital identification (NISO, 2004). Resource discovery serves the same functions as a good catalogue does. It allows resources to be found by relevant criteria such as subject, author, title, keyword etc. It helps in identifying the resources and giving the location information. The metadata helps in organizing the resources in a database or on the Internet i.e., by bringing the similar resources together and distinguishing dissimilar resources. As almost every metadata scheme contains an element for standard number it ensures the permanent digital identification of the information object, further this leads to the interoperability between multiple systems with different hardware and software platforms, data structure, and interfaces to exchange data with minimal loss of content

¹http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=16387

² <http://www.w3.org/html/wg/>

³ <http://www.w3.org/XML/>

and functionality. Although metadata model is quite popular, it involves some issues discussed later in the paper.

2. Approaches to Add or Extend the Semantics of Web page

There have been several approaches to add semantics to the Web pages.

2.1 Surrogate Approach

To communicate the structured data about an information object surrogate approach is used. In this approach the HTML representation is linked to an alternate representation of the content. The example for this is to have a RIS formatted citation data available from the page where an information object is available. RIS is a standardized tag format developed by Research Information Systems, Incorporated (the format name refers to the company) to enable citation programs to exchange data (RefMan, 2011). This approach is helpful in many cases such as in the area of Web Syndication with RSS and Atom feed files available along with the HTML page. This alternate representation is sent to the subscribers via HTTP that helps the consumer to decide upon whether to visit the html page or not. Also the automatic capturing of citation data by tools such as bibliographic management tools etc. has been very successful. Example:

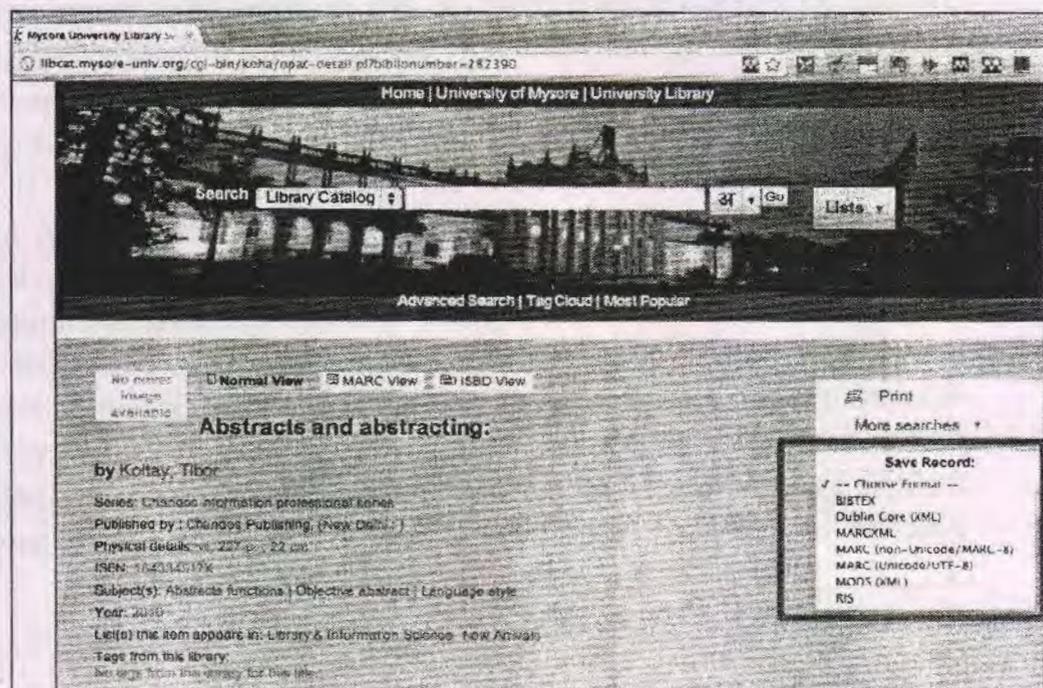


Figure 1. Screenshot from University of Mysore Library OPAC displaying different formats available to user for download

Although, libraries heavily use this approach it has some problems too. Techniques like this require consumers of the data to know where to look for this invisible content. It also adds a layer of complexity by requiring a site to have APIs or metadata gateways that can be burdensome to setup and expensive for organizations to maintain (Ronallo, 2012).

2.2 Metadata Approach

Another approach is to embed the data directly into the webpage. This is done by tagging the pages with meta tags these meta tags attempt to summarize the context of the entire page, adding information such as authorship and revision date. However, they can only describe the page as a whole: as meta tags are written in the <head> section, it is impossible for them to drill down and specify the individual content of the page (Storey, 2011). The meta tags are standardized and released as a standard assuming that every developer will use the same tags. Some of the metadata standards such as Dublin Core⁴, MODS⁵ etc. are very popular.

The screenshot shows a web browser window with the URL <https://drc.isibang.ac.in/handle/1849/506?show=full>. The page title is "Use of web technology in providing information services by south Indian technological universities as displayed on library websites". Below the title, there is a "Show simple item record" link. The metadata is displayed in a table-like format:

dc.contributor.author	Babu, Preedip Balaji	
dc.contributor.author	Kumar, Vinit	
dc.date.accessioned	2012-01-02T15:23:17Z	
dc.date.available	2012-01-02T15:23:17Z	
dc.date.issued	2011-07	
dc.identifier.citation	Preedip Balaji B., Vinit Kumar, (2011) "Use of web technology in providing information services by south Indian technological universities as displayed on library websites", Library Hi Tech, Vol. 29 Iss: 3, pp.470 - 495	en_US
dc.identifier.isbn	0737-8831	
dc.identifier.uri	http://hdl.handle.net/1849/506	
dc.description.abstract	Abstract Purpose - This paper aims to discuss the present status of using new generation web technology, social media and Web 2.0 features among the technological university library websites in south India. It assesses the library websites as a primary platform and one-stop portal for information services and examines as to how much library websites are effective in providing web-based information services. Design/Methodology/Approach - The library websites of the technological universities in south India were evaluated on the basis of a relative weight checklist. The criteria for the checklist was drawn on the basis of availability of websites for library, resource discovery tools, access to scholarly content and Web 2.0 tools. The various issues and challenges in adapting new web technologies in academic environment are discussed. Findings - Using the current web development technologies and deploying for mainstream web information services is not widespread as web information	en_US

Figure 2. Use of Dublin Core Metadata format (metadata displayed)

⁴ <http://dublincore.org/>

⁵ <http://www.loc.gov/standards/mods/>

```

<meta name="DC.creator" content="Basu, Preodip Balaji" />
<meta name="DC.creator" content="Kumar, Vinil" />
<meta name="DCTERMS.dateAccepted" content="2012-01-02T15:23:11Z" scheme="DCTERMS.W3CDTF" />
<meta name="DCTERMS.available" content="2012-01-02T15:23:11Z" scheme="DCTERMS.W3CDTF" />
<meta name="DCTERMS.issued" content="2011-07" scheme="DCTERMS.W3CDTF" />
<meta name="DCTERMS.bibliographicCitation" content="Preodip Balaji B., Vinil Kumar, (2011). <quot>Use of web technology in providing information services by south Indian technological universities as displayed on library websites<quot>; Library & Tech. Vol. 29 Iss. 3, pp.476 - 495" xml:lang="en_US" />
<meta name="DC.identifier" content="0717-8831" />
<meta name="DC.identifier" content="http://hdl.handle.net/1849/504" scheme="DCTERMS.URI" />
<meta name="DCTERMS.abstract" content="Abstract:<math>A</math>Purpose - This paper aims to discuss the present status of using new generation web<math>A</math>technology, social media and Web 2.0 features among the technological university libraries<math>A</math>websites in south India. It assesses the library websites as a primary platform and one-stop<math>A</math>portal for information services and examines as to how such library websites are effective in<math>A</math>providing web-based information services.<math>A</math>Design/Methodology/Approach - The library websites of the technological universities in south<math>A</math>India were evaluated on the basis of a relative weight checklist. The criteria for the checklist was<math>A</math>drawn on the basis of availability of websites for library, resource discovery tools, access to<math>A</math>scholarly content and Web 2.0 tools. The various issues and challenges in adopting new web<math>A</math>technologies in academic environment are discussed.<math>A</math>Findings - Using the current web development technologies and deploying for mainstream web<math>A</math>information services is not widespread as web information services are yet to take off widely in<math>A</math>academic libraries. The majority of university libraries are found to be working in the<math>A</math>conventional library settings and the diffusion rate of web information services is relatively low.<math>A</math>Originality/value - As this is an assessment of the existing online information infrastructure<math>A</math>facilities of the engineering universities in south India, the awareness of web-based information<math>A</math>services, their viability, and service values can be enhanced. More emphasis is underlined to<math>A</math>improve upon the current learning, online educational facilities and benchmarking electronic<math>A</math>information services for sustainability is highlighted.<math>A</math>Paper type - Case study" xml:lang="en_US" />
<meta name="DC.language" content="en" xml:lang="en_US" scheme="DCTERMS.RFC1766" />
<meta name="DC.publisher" content="Emerald" xml:lang="en_US" />
<meta name="DC.relation" content="Vol. 29/No. 3" />
<meta name="DC.subject" content="Web technology, Web information services, Web 2.0, Technological universities, Academic library websites, South India" xml:lang="en_US" />
<meta name="DC.title" content="Use of web technology in providing information services by south Indian technological universities as displayed on library websites" xml:lang="en_US" />
<meta name="DC.type" content="Article" xml:lang="en_US" />

```

Figure 3. Use of Dublin Core Metadata (hidden in a Webpage)

The metadata is also provided through a gateway following OAI-PMH (Open Archive Initiative - Protocol for Metadata Harvesting) for harvesting metadata this also has been supported by major search engines.

2.3 Metadata for content

In the above approaches the services are hidden and are prone to get neglected. So another approach was thought of to embed data in visible HTML, this helps in keeping the representations in sync and now the page authors have to expose only one public version of their data. Embedding data along with the visible HTML also solves the problem of marking up the meaning of the content and not the whole page. In other words, on a single page different items and their metadata can be accommodated.

2.3.1 Microformats

Microformats are one of the pioneers in providing a general approach of using visible HTML markup to publish structured data to the Web. Microformats extend tags via special values for class and div attributes.

Microformats usually use the convention and standard class names to provide meaning on a Web page. Fig. 4 shows the logo of microformats

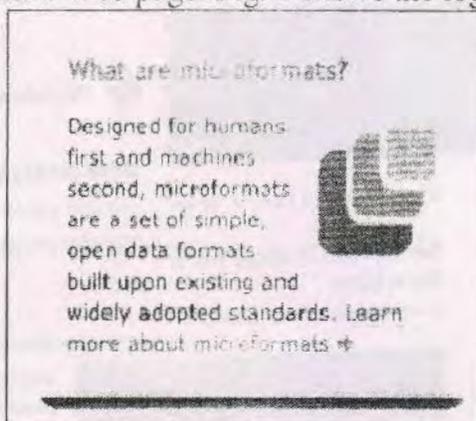


Figure 4. Logo of Microformats

Microformats have strong support by Google, Facebook and YouTube through schemas like

1. hcard (to represent people, companies, organizations, and places),
2. hcalendar (for events),
3. hRecipe (for recipe),
4. XFN (to diagram relationships between people),
5. VoteLinks, hReview (for opinions, ratings and reviews)
6. rel-license (for licenses),
7. rel-tag (for tags, categories and keywords),
8. XOXO (for lists and outlines),

In Feb, 2011 Google Launched its new search feature Recipe View⁶ It enables users to discover recipes that have been marked-up following the hRecipe specification from a variety of sources with a new level of accuracy. On the similar grounds Facebook also supports the hCard and hCalendar for marking up content related to people and events respectively. Support is provided in such as sense that the search engines or Web services can give more weight to the content if they know what kind of content it is. Similarly, Web browsers with the help of plugins can more easily pull out data for other uses than just display. Screen reader software by using the structured elements can more accurately read and provide access to the differently-abled people.

⁶ <http://www.google.com/landing/recipes/>



Figure 5. Microformat captured by a browser plugin.

The fig. [5] shows capture of microformat using a plugin in Google Chrome Web browser. A user can import the automatically captured microformat to Google Calendar or Gmail Contact list. The libraries can also embed the citation data or metadata in the form of microformats for automatic capture by browser plugins.

An Example of Microformat:

```
<div>
  
  <strong>Bob Smith</strong>    Senior editor at ACME
  Reviews    200 Main St    Desertville, AZ 12345
</div>
```

If we markup this as microformat it has to be put it like this:

```

<div class="vcard">
  
  <strong class="fn">Bob Smith</strong>
  <span class="title">Senior editor</span>
  at <span class="org">ACME Reviews</span>
  <span class="adr">
    <span class="street-address">200 Main St</span>
    <span class="locality">Desertville</span>,
    <span class="region">AZ</span>
    <span class="postal-code">12345</span>
  </span> </div>

```

2.3.2 RDFa

Another standard which has developed with the efforts of W3C and the semantic Web community for marking up the content of Web pages is RDFa⁷, Resource Description Framework in Attributes) In this approach too the context to tags are provided through the values added to del and class attributes. It also includes processing rules for embedding RDF and its graph-based framework (RDF Web Applications Working Group, 2012). Google has supported RDFa in some fashion since 2009, over the time found a large set of errors in the application RDFa by webmasters (Google, 2012a). Another version of RDFa, RDFa Lite 1.1⁸ has been developed to remove the complexities in implementing the RDFa. It provides simpler authoring guidelines. For example the following block of HTML code showing contact information of Bob Smith (Google, 2011):

```

<div> My name is Bob Smith but people call me Smithy.
Here is my home page: <a
href="http://www.example.com">www.example.com</a>. I
live in Albuquerque, NM and work as an engineer at ACME
Corp. </div>

```

The same HTML document marked up with RDFa

⁷ <http://www.w3.org/TR/rdfa-primer/>

⁸ <http://www.w3.org/TR/rdfa-lite/>

```

<div xmlns:v="http://rdf.data-vocabulary.org/#"
  typeof="v:Person"> My name is <span
  property="v:name">Bob Smith</span>, but people call
  me <span property="v:nickname">Smithy</span>. Here is
  my homepage: <a href="http://www.example.com"
  rel="v:url">www.example.com</a>. I live in
  Albuquerque, NM and work as an <span
  property="v:title">engineer</span> at <span
  property="v:affiliation">ACME Corp</span>. </div>

```

2.3.3 Microdata

Microdata branched from the approaches to incorporate RDFa in HTML5. Hypertext Markup Language (HTML) the popular language in which the information is represented in the World Wide Web has gone through lot of development in the past 20 years of its origin. Hypertext documents written by following the HTML standard and made accessible over a network by implementing the Hypertext transfer protocol and Transmission control protocol has been the reality for libraries. These standards making the foundation of the Web have gone through many changes and developments. The HTML also has gone through revisions and updates. After being in development for more than 20 years a new revised version of HTML has come known as HTML5⁹. The use of a bunch of div elements with different class names is the popular approach to markup the content of a webpage. This approach provides same semantics for all the content of a page, for handling this problem HTML5 provides new semantic elements such as header, nav, article, aside, section and footer allowing more expressiveness. HTML5 includes a syntax called Microdata¹⁰ that allows Web publishers to layer richly structured metadata directly into their Web pages. The syntax for microdata has been developed in such a way that it is simple for page authors to implement it. In this approach the extension of tag semantics is done via a shared vocabulary. Unlike meta tags, microdata is written in the body of an HTML document; unlike microformats and RDFa, it does not abuse class or del, instead uses new attributes that are built into HTML5.

⁹ <http://www.w3.org/TR/html5/>

¹⁰ <http://dev.w3.org/html5/md/>

2.3.3.1 Syntax of Microdata

The Microdata syntax is completely made up of HTML attributes. These attributes can be used on any valid HTML element. The things being described in an HTML page are recognized as “items”. The data model is made of three new HTML attributes

`itemscope` (marks the presence of a new item inside the element having this attribute)

`itemtype` (marks the type of item inside the element having this attribute)

`itemprop` (marks the properties and values of a item inside the element having this attribute)

An example code

```
<div itemscope itemtype="Institute">  
<span itemprop="Title">Documentation Research and  
Training Centre</span>  
</div>
```

As it is very much clear from the above code, the `itemscope` attribute marks the beginning of an item. The attribute `itemtype` specifies the type of thing we are talking about. Third attribute `itemprop` defines the property “title” and the value as Documentation Research and Training Centre.

3. Importance of Shared Vocabularies

In the above example the microdata snippet is completely valid according to the specifications of Microdata. But as one can see that for the values of the `itemtype` and `itemprop` attributes it uses an arbitrary language and not a standard one. This might be of help in a community or in a small group, but as Web is a public platform one always yearning for reaching the maximum. For reaching other machines and programs like Web crawlers etc., a shared language vocabulary is used so that the page authors and readers interpret and understand the meaning of the content easily. There are two big projects to running at present providing a shared vocabulary. Both are community driven and supported by major search engines.

1. Schema.org¹¹
2. data-vocabulary.org¹²

Schema.org is an attempt to define a broad, Web-scale, shared vocabulary focusing on popular concepts. It is more or less like a middle ontology that does not attempt to have the scope of an "ontology for everything" or go into depth in any one area. Schema.org provides a collection of shared vocabularies webmasters can use to mark up their pages in ways that can be understood by the major search engines: Google, Microsoft, Yandex and Yahoo! At present the vocabulary is in version 0.9.

```
<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>
<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 item , ="reviewCount">3077</span> reviews
</div>
<div itempr "offers" itemscope
itemtype="http://schema.org/Offer">
  Price: <span itemprop="price">$6.99</span>
  <meta itemprop="priceCurrency" content="USD" />
  <link itemprop="availability"
href="http://schema.org/InStock">In Stock
</div>
```

¹¹ <http://schema.org/>

¹² <http://data-vocaulary.org/>

```

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>

```

Reviews:

```

<div itemprop="review" itemscope
itemtype="http://schema.org/Review">
  <span itemprop="reviewRating">5</span> stars -
  <b><span itemprop="name">A masterpiece of
literature</span>" </b>
  by <span itemprop="author">John Doe</span>,
  Written on <meta itemprop="datePublished"
content="2006-05-04">May 4, 2006
  <span itemprop="reviewBody">I really enjoyed this
book. It captures the essential challenge people face
as they try make sense of their lives and grow to
adulthood.</span>
</div>
<div itemprop="review" itemscope
itemtype="http://schema.org/Review">
  <span itemprop="reviewRating">4</span> stars -
  <b><span itemprop="name">A good read.</span>" </b>
  by <span itemprop="author">Bob Smith</span>,
  Written on <meta itemprop="datePublished"
content="2006-06-15">June 15, 2006
  <span itemprop="reviewBody">Catcher in the Rye is a
fun book. It's a good book to read.</span>
</div>
</div>
...
</body>

```

Table 1. Schema.org types and its explanation

Microdata and Schema.org is the latest search engine preferred method of exposing structured data in HTML. The broadest item type is Thing, which has four properties: name, description, URL, and image. More specific types share properties with broader types. For example, in the above example of HTML representation for a book follows a type Book, which is a more specific type of CreativeWork, which is a more specific type of Thing (Schema.org, 2011). More specific types inherit

the properties of their parents. Most commonly used item types from Schema.org:

1. Creative works: CreativeWork, Book, Movie, MusicRecording, Recipe, TVSeries ...
2. Embedded non-text objects: AudioObject, ImageObject, VideoObject
3. Event
4. Organization
5. Person
6. Place, LocalBusiness, Restaurant ...
7. Product, Offer, AggregateOffer
8. Review, AggregateRating

The entire hierarchy is listed at <http://schema.org/docs/full.html>.

4. Steps for Embedding Metadata

4.1 Pick Up a Markup Format

The first step is to decide upon the markup format. Of the above three markup formats discussed above one can use any on these or a combination of these. Google recommends using microdata. Although, all the formats are in use these days and are acceptable it depends on what one is trying to do. If one wants to create an XHTML page, use of Dublin Core and microformats is recommended. If one wants to add a Creative Commons license to a page use of RDFa and Microdata is advised and if one uses the state-of-the-art HTML5 page use Microdata, adding metatags and RDFa as they are supported with namespaces.

4.2 Mark up the content

The next step is to markup the content this can be done by using categories from some shared vocabulary. Google, Bing and Yahoo support Schema.org. But any other shared vocabulary if any arrives will be well acceptable.

4.3 Test the markup

Finally the markup can be tested against rich snippet testing tool¹³ this testing makes sure that the marked up content is well readable by the search engines. It also tells us the possible errors such as, non-

¹³ <http://www.google.com/webmasters/tools/richsnippets>

compliance of usage guidelines for different content types, marked-up content hidden from user, etc. Similarly, incorrect and misleading markup also leads to no extraction by crawlers. Non-representativeness of markup with the content can also be another reason for non-display of snippets in the search results (Google, 2012b).

5. Applications/use of Microdata

There are many applications of Microdata in the Web environment. With the declaration about their support by major search engines it has gained a lot of publicity.

5.1 Exposing Library Pages to Search Engines

In 2005, a major study done by OCLC (Online Computer Library Center) "Perceptions of Libraries and Information Resources" revealed that 72% respondents used Web search engine and only 30% used library catalogue at least once (OCLC, 2005). In the terms of beginning the search for information on a particular topic 84% respondents start their search from web search engines while only 1% of them start from library. Similarly, 80% of respondents said the search engine would be their first choice the next time they need a source for information. This study clearly reveals that the users are more satisfied with the services offered by the Web search engines. On the other hand the holdings and other bibliographic details stored inside the library catalogues are not indexed by the Web search engines. This requires users to separately search in the library OPACs. For example, A resource which exists in the library is visible to the user after querying on a popular Web search engine showing its presence on Amazon but not in library website will imply to the user that the particular resource is not available in his/her library. The reasons behind this obscurity of data or content can be many as the architecture of library OPAC stores the data in such a way that the content becomes obscure to search engine crawlers. There were some approaches to expose the collections by using sitemaps, an XML file having the URLs of all the pages that should be visited by the crawlers of major search engines and then these sitemaps.xml files are registered to the search engines separately. Despite of being too popular this approach is limited to exposing pages and not the content of the page. The search engines are currently using Microdata embedded in pages to display rich snippets in search results.

Last.fm - Android Apps on Google Play



<https://play.google.com/store/apps/details?id=fm.last>

★★★★ 42,471 votes - Free - Android

30 Apr 2012 - Learn about your favorite artists, find events, and share your music tastes. Learn about your favorite artists, find nearby events, and share your ...

image

Bollywood Chutney Chicken Pizz Recipe



food2goodhealth.com/.../Bollywood.../1754.046

38 mins

Easy to Make **Bollywood** Chutney Chicken Pizz, Indian flavors and **Bollywood** movies are quickly gaining popularity in America--and for good reason. Both are ...

Ingredients: chicken, garam masala, ginger, garlic, turmeric, polenta, olive ...

Figure 6. Rich snippets of a product showing its price and rating

This gives a lot more information to the user that helps in deciding whether to click on a particular result or not. Apart from this the rich snippets also reduces clicking websites fooling search engines to increase traffics on their page. The search engines can use this structured data as signals of relevance. Even though the search engines like Google support the rich snippet it is not necessary that the search results will show up in rich snippet view only. It depends on the policy of the search engine, for example Google currently only supports some itemtypes such Applications, Authors, Events, Movie, Music, People, Products, Recipes, Reviews, TV Series etc. The presence of a threshold of properties creates a chance of showing up as rich snippet. Even then a rich snippet may only show up for a page if an item is relevant to the users query.

5.2 Automatic Capture of Metadata

The embedded metadata can be captured by browser plugins and then further processed in different ways for various other purposes. Similarly, by using a common shared vocabulary it becomes quite easy for site scrappers to selectively extract content from websites.

6. Conclusion and Future work

With the increase in popularity of Microdata other consumers for this structured data may appear, such as Web applications searching only a

specific type of items or things from Web or the ongoing faceted search and faceted navigation efforts will also take a leap forward. Although, some content management systems have developed plugins to markup content in Microdata, implementing Schema.org and Microdata is a bit tedious as the popular integrated development environments for Webpages have not updated themselves, but with the increase in popularity in no time it may arrive. Shared vocabulary Schema.org is still in developing stage and at its present format may not fit the needs of the libraries. This unfitness of expressing varied range of content stored in libraries provides an opportunity for libraries and librarians to enumerate the places where the vocabulary could be modified to better fit their data and use cases. Seeing in the work in the direction of mapping Dublin Core Metadata Initiative and Schema.org by the DCMI Schema.org Alignment Task Group (DCMI, 2012) one can see a huge implementation in digital libraries too.

References

- Berners-Lee, T. (1997). Axioms of Web Architecture□: Metadata. Retrieved May 21, 2012, from <http://www.w3.org/DesignIssues/Metadata.html>
- CC:DA Task Force on Metadata. (2000). *CC:DA Task Force on Metadata: Final Report*. Retrieved from <http://www.libraries.psu.edu/tas/jca/ccda/tf-meta6.html>
- DCMI. (2012). Schema.org Alignment. Retrieved May 21, 2012, from http://wiki.dublincore.org/index.php/Schema.org_Alignment
- Google. (2011). About RDFa - Webmaster Tools Help. Retrieved May 21, 2012, from <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=146898>
- Google. (2012a). Rich snippets (microdata, microformats, and RDFa) - Webmaster Tools Help. Retrieved May 21, 2012, a from <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=99170>
- Google. (2012b). Rich snippets not appearing - Webmaster Tools Help. Retrieved May 21, 2012, b from <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=1093493>
- NISO. (2004). *Understanding Metadata*. Retrieved from www.niso.org/standards/resources/UnderstandingMetadata.pdf

- OCLC. (2005). *Perceptions of libraries and information resources* : a report to the OCLC membership (pp. 12-19).
- RDF Web Applications Working Group. (2012). *RDFa 1.1 Primer*. Retrieved from <http://www.w3.org/TR/rdfa-primer/>
- RefMan. (2011). RIS Format Specifications. *RefMan Documentation*. Retrieved May 21, 2012, from http://www.refman.com/support/risformat_intro.asp
- Ronallo, J. (2012). HTML5 Microdata and Schema.org. *The Code4Lib Journal*, (16). Retrieved from <http://journal.code4lib.org/articles/6400>
- Schema.org. (2011). Book - schema.org. Retrieved May 21, 2012, from <http://schema.org/Book>
- Storey, D. (2011). Microdata: Metadata For Content. *demosthenes.info*. Retrieved from <http://demosthenes.info/blog/344/Microdata-Metadata-For-Content>