Building Smarter Books in Browsers with RDFa, Schema.org, & Linked Data: Leveraging Standards & Tools in the Creation of Semantically-Enhanced Reading Systems

Jason A. Clark
@jaclark
Book as Platform:
Linking Data to Build a Better Book
The Swabian.

He was slow in learning how to talk. "My parents were so worried," he later recalled, "that they consulted a doctor." Even after he had begun using words, sometime after the age of 2, he developed a quirk that prompted the family maid to dub him "der Depperte," the dopey one, and others in his family to label him as "almost backwards." Whenever he had something to say, he would try it out on himself, whispering it softly until it sounded good enough to pronounce aloud.

Every sentence he uttered, his worshipful younger sister recalled, "no matter how routine, he repeated to himself softly, moving his lips." It was all very worrying, she said. "He had such difficulty with language that those around him feared he would never learn."

His slow development was combined with a cheeky rebelliousness toward authority, which led
Web Book

Did you mean:

- Ebook
- Always correct to "Ebook"
- Add to personal dictionary
- Ignore all
Every Book its [machine] Reader

Apologies to Ranganathan
What if?

Books were machine indexed for semantic discovery

Journal articles were semantically tagged at the page level
A “Book” in a Browser

Atomizable, Analyzable, Linkable, Reusable
Book as Software

1. Software Development
2. Data Storage and Archiving
3. Metadata
4. Discovery and Analytics
5. API Access
Links and Anchors

A link is the connection between one piece of hypertext and another.

```html
  <title>Hypertext Links</title>
  <h1>Links and Anchors</h1>
  A link is the connection between one piece of
  <a href="WhatIs.html">hypertext</a> and another.
```
The Research

Institute of Museum and Library Services
Sparks Ignition Grant
Web Book Prototyping
Our Initial Project

Digital Book for History Class

MS Word -> Static HTML -> Dynamic HTML

API driven and Linked Data
Book 1
demo: arc.lib.montana.edu/book/home-cooking-history-409
code: github.com/jasonclark/bib-template

Book 2
demo: arc.lib.montana.edu/book/opsis
code: github.com/msulibrary/bib-template-fiction
Recipe and Essay

Gingerbread

Ingredients:

- 1/2 c. shortening
- 2 1/4 c. flour
- 2 T. sugar
- 1 tsp. soda
- 1 egg
- 1/2 tsp. salt
- 1 c. dark New Orleans Molasses
- 1 tsp. ginger
- 1 tsp. boiling water
- 1 tsp. cinnamon

Directions:

- Mix shortening, sugar and egg.
- Blend in molasses and cup of boiling water, cooled first.
- Add flour, soda, salt, ginger and cinnamon.
- Beat until smooth.
- Pour into well-greased 9 inch square pan.
- Bake 45 minutes to 50 minutes at 350 degrees.

Serve with whipped cream and applesauce.

Tradition and Travel

by Jonathan Bergh

Every winter the Bergh family celebrates the season with a special tradition. It is the time of the year when the Bergh family...
young Lyle Bergh gingerbread was the catalyst for exposure to countless foodstuffs that were outside of the farm menu. Once Lyle Bergh took his first bite into gingerbread cake he knew he ate something special. To the Bergh family gingerbread may have represented a season, but for Lyle it represented so much more.

In 1960s northeastern Montana’s food options were limited. For inhabitants of Plentywood, Montana, like Lyle and his family, it meant that you would eat farm food: simple meals like steak, potatoes, milk and coffee. Lyle would come home and have farm food that his mother Bee cooked. When he and his family would go Seattle, Washington, it would represent more than just family gatherings. When the Bergh family headed to Seattle, Washington in the winter, it meant that a family would be reunited, and it meant there would be treats. For Lyle, going to Seattle meant that he going to Seattle meant that he would have a break from farm food and farm work. It meant that he would get to see his grandma Belen. Seattle, Washington is the place that the tradition of gingerbread cake started in the Bergh family. Grandma Belen was in the restaurant business so Lyle knew that he and his family would be eating well-prepared food. It was on these trips that Lyle got his first taste of gingerbread cake. This cake was soft and moist and topped with whipped cream that tingled his taste buds. This sugarless cake also represented that here in Seattle he and the family would have foodstuffs that were not accessible in Montana.

Lyle knew that whenever the family traveled Seattle that he would be exposed to new and foreign foods. Foods like avocados and seafood became some of Lyle’s favorites because they reminded him of family. This is not to say that he disliked farm food, but Seattle food was a delicacy that Lyle could only get once a year. Because he loved gingerbread so much his mother Bee brought the recipe back to Montana. Gingerbread became a family tradition that has continued for generations.

Bringing the recipe for gingerbread back to Montana meant bringing back memories for Lyle. It brought back memories of food being shared with him for the first time along with sharing moments with loved ones. Most importantly, it gave Lyle Bergh a tradition that he continues to this day. Whenever the first snow of winter falls you can always expect that Lyle Bergh will be baking gingerbread. For Lyle baking gingerbread meant taking the little piece of Seattle and bringing it back to northeastern Montana.
Every winter the Bergh family celebrates the season with a special tradition. It is the time of the year when the Bergh family makes gingerbread cake. For a young Lyle Bergh, gingerbread was the catalyst for exposure to countless foodstuffs that were outside of the farm menu. Once Lyle Bergh took his first bite into gingerbread cake he knew he ate something special. To the Bergh family gingerbread may have represented a season, but
The Mysteries of Udolpho, Chapter 1

On the pleasant banks of the Garonne, in the province of Gascony, stood, in the year 1584, the chateau of Monsieur St. Aubert. From its windows were seen the pastoral landscapes of Guienne and Gascony stretching along the river, gay with luxuriant woods and vine, and plantations of olives. To the south, the view was bounded by the majestic Pyrenees, whose summits, veiled in clouds, or exhibiting awful forms, seen, and lost again, as the partial vapours rolled along, were sometimes barren, and gleamed through the blue tinge of air, and sometimes crowned with forests of gloomy pine, that swept downward to their base. These tremendous precipices were contrasted by the soft green of the pastures and woods that hung upon their skirts; among whose flocks, and herds, and simple cottages, the eye, after having scaled the cliffs above, delighted to repose. To the north, and to the east, the plains of Guienne and Languedoc were lost in the mist of distance; on the west, Gascony was bounded by the waters of Biscay.

M. St. Aubert loved to wander, with his wife and daughter, on the margin of the Garonne, and to listen to the music that floated on its waves. He had known life in other forms than those of pastoral simplicity, having mingled in the gay and in the busy scenes of
On the pleasant banks of the Garonne, in the province of Gascony, stood, in the year 1584, the chateau of Monsieur St. Aubert. From its windows were seen the pastoral landscapes of Guienne and Gascony stretching along the river, gay with luxuriant woods and vine, and plantations of olives. To the south, the view was bounded by the majestic Pyrenees, whose summits, veiled in clouds, or exhibiting awful forms, seen, and lost again, as the partial vapours rolled along, were sometimes barren, and gleamed through the blue tinge of air, and sometimes frowned with forests of gloomy pine, that swept downward to their base. These tremendous precipices were contrasted by the soft green of the pastures and woods that hung upon their skirts; among whose flocks, and herds, and simple cottages, the eye, after having scaled the cliffs above, delighted to repose. To the north, and to the east, the plains of Guienne and Languedoc...
Web Book Dev Stack

HTML
CSS
Javascript
MySQL
PHP
schema.org
RDFa
Disruption?

New Roles for Publisher
“But think of the S.E.O. if the title actually named the two cities.”
Old Roles

Validating
Editing
Typesetting
Promotion
Sales
New Roles

Archiving
Metadata [structured]
Discovery
Analytics + Metrics
Sharing + Reuse
Structured Data

Data Model + Machine Readability
Schema.org
controlled vocabulary
CREATE TABLE IF NOT EXISTS `bodymatter` (  
`id` int(11) NOT NULL AUTO_INCREMENT,  
`name` varchar(255) DEFAULT NULL,  
`creator` varchar(255) DEFAULT NULL,  
`url` varchar(255) DEFAULT NULL,  
`description` text,  
`about` text,  
`image` varchar(140) DEFAULT NULL,  
`articleBody` text COMMENT 'http://schema.org/Article',  
`articleSection` text COMMENT 'http://schema.org/Article',  
`publicationType` varchar(255) NOT NULL COMMENT 'http://schema.org/MedicalScholarlyArticle',  
`additionalType` varchar(255) DEFAULT NULL,  
`additionalType2` varchar(255) DEFAULT NULL,  
`isbn` varchar(40) DEFAULT NULL,  
`genre` varchar(140) DEFAULT NULL,  
`keywords` varchar(255) DEFAULT NULL,  
`publisher` varchar(255) DEFAULT NULL,  
`dateCreated` varchar(255) DEFAULT NULL,  
`dateModified` timestamp NULL DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP,  
`datePublished` varchar(255) DEFAULT NULL,  
`version` varchar(140) DEFAULT NULL,  
`learningResourceType` varchar(140) DEFAULT NULL,  
`inLanguage` varchar(5) DEFAULT NULL,  
PRIMARY KEY (`id`),  
FULLTEXT KEY `search` (`name`, `creator`, `description`, `genre`, `keywords`, `dateCreated`)  
) ENGINE=MyISAM DEFAULT CHARSET=utf8;
epub:type

semantic divisions...
<html>
  <section epub:type="bodymatter">
    ...
  </section>
</html>
Gingerbread

Ingredients:
- 1/2 c. soft shortening
- 2 1/4 c. flour
- 2 T. sugar
- 1 tsp. soda
- 1 egg
- 1/2 tsp. salt
- 1 c. dark New Orleans Molasses
- 1 tsp. ginger
- 1 c. boiling water
- 1 tsp. cinnamon

Instructions:
Mix shortening, sugar and egg. Blend in molasses and cup of boiling water, cooled first. And flour, soda, salt, ginger and cinnamon. Beat until smooth. Pour into well-greased 9 inch square...

Cooking Method:
Bake

Cook Time:
PT45M
Discovery

RDFa

JSON-LD

Linked Data Topics
Gingerbread - Jonathan Bergh

45 mins

Table of Contents > Recipe and Essay Gingerbread. Ingredients: 1/2 c. soft shortening; 2 1/4 c. flour; 2 T. sugar; 1 tsp. soda; 1 egg; 1/2 tsp. salt; 1 c. dark New ...
The sand softens shifts ...
CIP Data

Cataloging in Publication

www.loc.gov/publish/cip/techinfo/cipdata.html
<body property="mainEntity" typeof="Book">
<main epub:type="cover" role="main" id="top">
  <a href="#menu" class="control open">menu</a>
</main>
</body>
Analytics + Metrics

Web Analytics, UX
<table>
<thead>
<tr>
<th>Source / Medium</th>
<th>Acquisition</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sessions</td>
<td>% New Sessions</td>
</tr>
<tr>
<td>Cookbook</td>
<td>874 (0.22%)</td>
<td>61.44% (40.89%)</td>
</tr>
<tr>
<td>1. google / organic</td>
<td>334 (38.22%)</td>
<td>73.35% (46.62%)</td>
</tr>
<tr>
<td>2. montana.edu / referral</td>
<td>198 (22.65%)</td>
<td>41.92% (15.46%)</td>
</tr>
<tr>
<td>3. (direct) / (none)</td>
<td>190 (21.74%)</td>
<td>72.63% (25.70%)</td>
</tr>
<tr>
<td>4. t.co / referral</td>
<td>52 (5.95%)</td>
<td>69.23% (6.70%)</td>
</tr>
<tr>
<td>5. lib.montana.edu / referral</td>
<td>16 (1.83%)</td>
<td>0.00% (0.00%)</td>
</tr>
<tr>
<td>6. diffbot.com / referral</td>
<td>11 (1.26%)</td>
<td>0.00% (0.00%)</td>
</tr>
<tr>
<td>7. shelf.lib.montana.edu / referral</td>
<td>11 (1.26%)</td>
<td>0.00% (0.00%)</td>
</tr>
<tr>
<td>8. facebook.com / referral</td>
<td>6 (0.69%)</td>
<td>50.00% (0.58%)</td>
</tr>
<tr>
<td>9. isbn.lib.montana.edu / referral</td>
<td>6 (0.69%)</td>
<td>33.33% (0.37%)</td>
</tr>
<tr>
<td>10. aol / organic</td>
<td>5 (0.57%)</td>
<td>60.00% (0.56%)</td>
</tr>
<tr>
<td></td>
<td>google / organic</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>baklava tradition</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>german chili recipe</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>montana state university</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>angel food cake essay +</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>bison pot roast slow cooker with beer</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>chocolate sheets cakes for mother</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>composition cooking cake with my mother</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>debone redfish fillet</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>essay my mother not at home decided to bake a cake</td>
</tr>
</tbody>
</table>
Kattfulwurst (German Potato Sausage)

Ingredients:
- 16 lbs. potatoes, peeled and diced
- 5 lbs. onions, peeled and chopped
- 4 lbs. ground beef
- 2 lbs. ground pork
- Pepper, Salt, Garlic Powder and Italian seasoning to taste
- 15 pig casings
- 3 cans beets

Directions:
- Using an electric grinder or hand crank, grind potatoes and onions.
- Combine with beef, pork, seasonings, and beer.
- Stuff mixture into casings and use a sewing needle to poke holes into the sides to prevent air bubbles.
- When stuffed tightly, tie off and place onto an oiled sheet pan.
- Poke sausages again, and bake 350 degrees for one 1-1/2 hours.
- If you have leftover mixture, press into oiled casserole dishes, and bake 350 degrees for one hour or until golden brown.

After consulting a native German friend, I discovered the correct spelling of this dish is probably Kattfulwurst und Wurst.
Sharing + Reuse

Web Services + API
Social Sharing
Deep Linking
Your search for adonis resulted in 1 match(es).

Lost In The Garden Of Adonis ...
https://arc.lib.montana.edu/book/opsis/item/4
Content: ... Creator: Carol Clonan... Date: Feb 1 2015 6:43 PM ...
<!-- Social Media Tags -->
<meta name="twitter:title" content="Opsis">
<meta name="twitter:description" content="Literary Arts Journal at Montana State University (MSU)">
<meta name="twitter:image:src" content="https://arc.lib.montana.edu/book/opsis/meta/img/opsis.png">
<meta name="twitter:url" content="https://arc.lib.montana.edu/book/opsis/">
<meta name="twitter:card" content="summary_large_image">
<meta name="twitter:site" content="@msulibrary">
<meta name="twitter:creator" content="@msulibrary">
<meta property="og:title" content="Opsis">
<meta property="og:description" content="Literary Arts Journal at Montana State University (MSU)">
<meta property="og:image" content="https://arc.lib.montana.edu/book/opsis/meta/img/opsis.png">
<meta property="og:url" content="https://arc.lib.montana.edu/book/opsis/">
<meta property="og:type" content="website"/>
<meta property="og:site_name" content="Montana State University (MSU) Library"/>
<!-- End Social Media Tags -->
Check out Opsi (in a browser)
arc.lib.montana.edu/book/opsis/
<table>
<thead>
<tr>
<th>Social Network</th>
<th>Acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookbook</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sessions</td>
</tr>
<tr>
<td></td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>% of Total: 0.02% (392,074)</td>
</tr>
<tr>
<td>1. Twitter</td>
<td>52 (86.67%)</td>
</tr>
<tr>
<td>2. Facebook</td>
<td>8 (13.33%)</td>
</tr>
</tbody>
</table>
Deep-linking the book

HTTP + URL
Ongoing Research

Semantic Classification
Genres - textbook, journal
UX of Reading
Real-time Analytics
Annotations
Offline
getSEMantic [beta]
jason a. clark

Mmmmm... Semantic Yumminess

Ranked Entities

**Engineering**
relevance: 0.944351
concepts & linked data resources:
- http://dbpedia.org/resource/Engineering
- http://rdf.freebase.com/ns/m.02ky346

**Social sciences**
relevance: 0.795123
concepts & linked data resources:
- http://dbpedia.org/resource/Social_sciences
- http://sw.opencyc.org/concept/Mx4rwauQZNc8Qdibm_jCfd9X4g
data=... options in almost every function we use from here forward. Figure 1-4 contains the side-by-side boxplots showing right skew for all the groups, slightly higher median and more variability for the Unattractive group along with some potential outliers indicated in two of the three groups.

```r
> boxplot(Years~Attr, data=MockJury)
```

![Boxplot](image)

**Figure 1-4**: side-by-side boxplot of Years based on picture groups.

The `~` (the tilde symbol, which you can find in the upper left corner of your keyboard) notation will be explained in a later section.
cabin in the woods, seeking immortality. The boy was
and preternaturally ugly. He received his looks from
a woodsman of sorts who engraved likenesses in tree

Context:

Fairy tale

Read Time:

5 min to read

Page:

-30-
The “Book” in a Network

http://arc.lib.montana.edu/book/

Designed for the bot’s experience

Designed for the reader’s experience
Survey of the Field

HTMLBook
https://github.com/oreillymedia/HTMLBook

Leanpub
https://leanpub.com/

GitBook
https://www.gitbook.com/

Penflip
https://www.penflip.com/

gitenberg
https://gitenberg.github.io/

Book in Browser Conference
http://booksinbrowsers.org/

Library Publishing Coalition
http://www.librarypublishing.org/
w3c Groups

Schema.org Community Group
https://www.w3.org/community/schemaorg/

Digital Publishing Interest Group
http://www.w3.org/dpub/

Semantic Web Interest Group
http://www.w3.org/2001/sw/interest/

Web Annotation Working Group
http://www.w3.org/annotation/

Data on the Web Best Practices Working Group
http://www.w3.org/2013/dwbp/wiki/Main_Page
A Research Team Effort

Scott Young
Digital Initiatives Librarian

Jason A. Clark
Head, Library Informatics & Computing

Michelle Gollehon
Digital Production Operations Manager

Jan Zauha
Outreach Librarian
Thanks!

Jason A. Clark
Head, Library Informatics & Computing
Montana State University Library
@jaclark

Scott Young
Digital Initiatives Librarian
Montana State University Library
@hei_scott