

How I fixed my SEO Structured Data problem on SoccerMovieMom.com

Background:

The purpose of this article is to document the steps I took to make my WordPress (WP) site indexed and searchable by Google and Bing search engines. Before I added structured data fields to my web pages, my site failed to show up in search engine result pages (SERPs).

I created SoccerMovieMom.com in May-2014 using the WordPress TwentyFourteen theme. The website is a collection of movie reviews of over 130 soccer films. It was my first WP site and took me 2 weeks of tough work, even though I had a technical background, having retired from software development over 10 years ago.

I was just happy to get the website up and running, and it wasn't until a year later that I realized how impossible it was to actually find my soccer movie reviews in a google search.

Upon researching, I discovered that around the time I created the website, the search engines began requiring that web pages adhere to a Structured Data model in order to be indexed. Before, people loaded up their webpages with keywords for the search engines to find, but now webpages need to be tagged with invisible data fields that identify what is in the page, e.g., date published, author name, type of content. This enables search engines to provide more relevant hits in the result pages.

Moving to structured data though, uncovered error messages about the invisible data that Wordpress generates. The rest of this article explains how I fixed the structured data errors on my website and saw improved search results. I also created JSON-LD structured data per Google's instructions, but I don't believe this has improved my SEO.

Update Jan-1-2016:

After posting this article in Jul-2015, and after I became more familiar with Google search, I found that some of my earlier conclusions were not quite correct, and I have updated those sections. In the five months after this article, I slowly realized a number of issues:

- My improved search results were specious: most of the hits that appeared in my Google Analytics page turned out to be **fake**. They were generated by **referral spam**. I had to make changes in my Google Analytics page to try and screen out the bad data, but referral spam continues to comprise half of my hits in the GA reports.
- The JSON-LD script that I added inside the body of all my posts was sometimes corrupted when WP saved the post. I had to instead use a plugin to put the script in the post's Header.
- My JSON-LD script posed one significant problem for googlebot due to a formatting problem with smart double quotes in text files.
- Lastly, after putting all this work into JSON-LD, I discovered that Bing (and therefore Yahoo) does not support JSON-LD. I was foolish to blindly follow google down a path without checking what other search engines do. Since Bing does not recognize JSON-LD, I will have to add META tags and microdata to my posts as well. I will describe that work in a separate article.

Assess your Current SEO

Making your website searchable is called increasing its SEO (Search Engine Optimization). There is too much to cover step-by-step in detail, so the following are general descriptions.

First of all, assess the current searchability of your website.

1. In Google Chrome, find out how searchable you are by typing in the URL bar "site:soccermoviemom.com", substituting your domain name for mine.
2. At the top of the SERP, google may pop up a box asking you to try [Google Search Console](#).

Try Google Search Console

www.google.com/webmasters/

Do you own **soccermoviemom.com**? Get indexing and ranking data from Google.

Google promotion

This is the former Google Webmaster Tools. You will see all the effects and deficiencies of structured data, so follow that link and set up your website. You will need a google login.

At this time, if google doesn't know about your website and its sitemap, do not submit it yet.

3. The Google Search Console has some web site tracking data, but you will get more data by also setting up [Google Analytics](#) to measure your site's progress. This screenshot shows website traffic picked up after my SEO changes. However, note the Update comment.

Update: Keep in mind that this old data is actually wrong. These 80-90 hits per day and 904 hits in one month? **Most of it is referral spam**



Fixing the “updated” errors in Google Search Console

In Google Search Console, you might see an error saying that you are missing the “updated” field or an “author” field. This is because wordpress creates some data structures that aren’t compatible with the definition that the search engine is using. Here’s what I did to solve this problem.

1. In the WP Admin console, I navigated to the “Appearance > Editor” for the TwentyFourteen theme and fixed the code directly in the theme.
 - Technically, you should never edit the theme but should instead create a Child theme. Otherwise, your edits will be lost when the theme is updated.
 - However, this theme is rarely updated, I was only changing one line, and creating a child theme meant making new sub-dirs etc, I decided to edit the theme directly.
 - I also had to turn off automatic WP updates in my ISP Dreamhost's web panel. Otherwise, any theme update could occur without my knowing, and my code changes would be wiped out. By turning off the automatic update, I can run the update manually and then immediately reinput the code changes into the new files.
 - All the forums I had read suggested checking these files for the line to fix: single.php, content.php, and template-tags.php. I found the **twentyfourteen_posted_on** function within **template-tags.php** and inserted the word “updated” into the highlighted spot below,

```
// Set up and print post meta information.

printf( '<span class="entry-date">
<a href="%1$s" rel="bookmark">
<time class="entry-date updated" datetime="%2$s">%3$s</time>
</a></span> <span class="byline"><span class="author vcard"><a class="url fn n" hr
ef="%4$s" rel="author">%5$s</a></span></span>'
```

2. Update the file in the WP console. This should fix the error for all WP posts. However, if you already had Google Search Console indexing your website, it may take several days for the fix to filter down to all your posts. This is why I earlier suggested not to submit your sitemap to google until you have fixed your structure data problems.

Fixing the missing “author” errors on Wordpress Pages

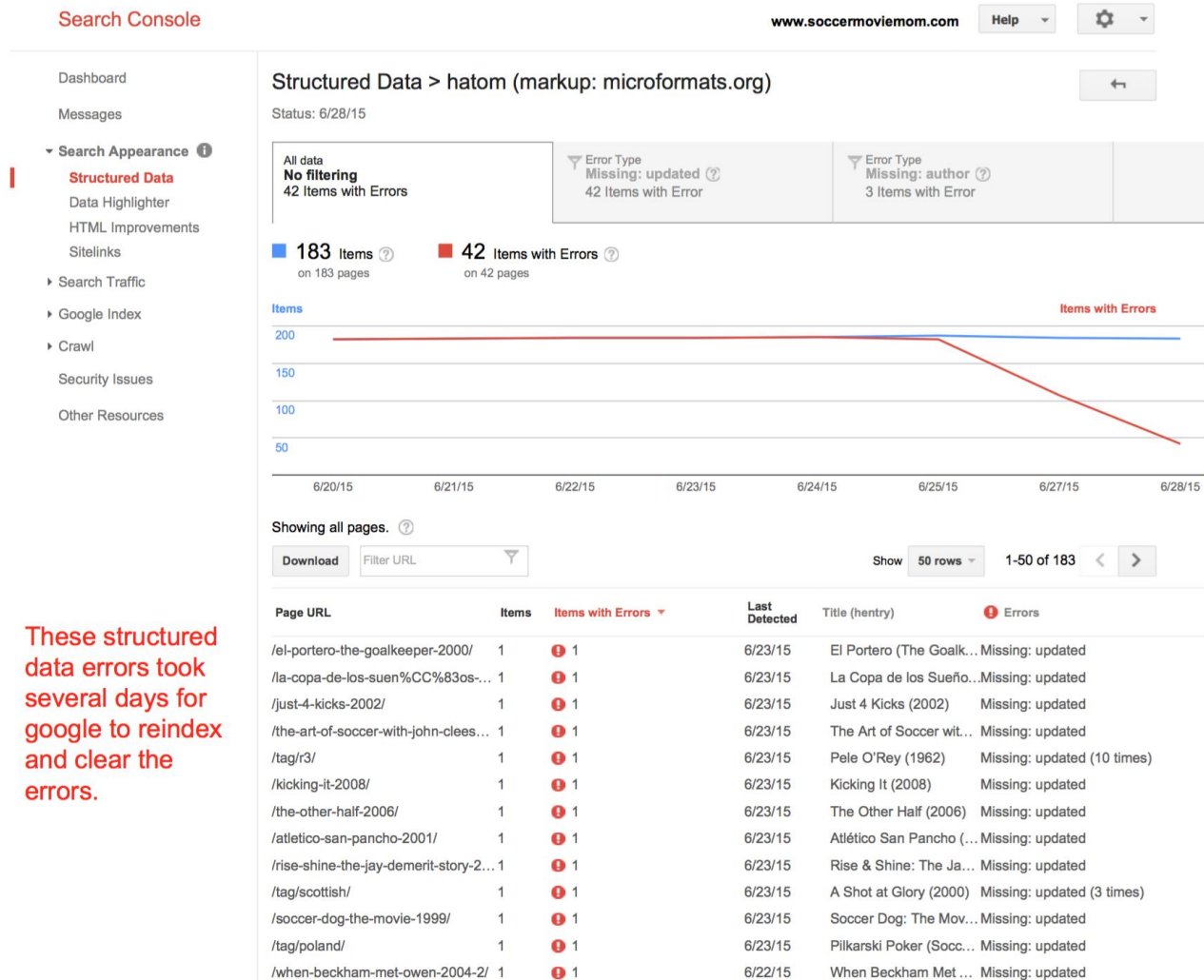
Wordpress has a similar issue for WP pages, such as your home page. In this case, you may see a missing “author” error on the Google Search Console. Again, for me the wordpress.org forum didn’t have any real help, but I found the fix in the [Google Webmaster Central Help Forum](#). To fix this, I had to create a plug-in that effectively tells a browser, “if this is a Page or an Archive, do not load the hentry data structure”.

1. I created a .txt file with the following code for my plug-in.

```
<?php
/*
Plugin Name: MJ Remove hentry
Description: Removes the "hentry" class from pages, archives and the "blog index"
(prevents structured data errors)
*/
function mj_remove_hentry( $classes ) {
if (is_page() || is_archive() || is_home()){ $classes = array_diff(
$classes, array('hentry')); } return $classes; }
add_filter( 'post_class', 'mj_remove_hentry' );
?>
```

2. On the mac, I compressed the plugin .txt file into a zip file.
3. In the WP Admin console, I then uploaded the plugin and activated it.
 - As a precaution, I made sure to only use underscores in the function name, and my file and folder names used hyphens and no spaces.
 - Note that in your Wordpress directory structure, plug-ins are stored in the wp-content folder on your wordpress server.
4. As with the “updated” error, if you have already submitted your sitemap to google, then it may take days for these errors to go away in the Google Search Console (see the following screen shot).
5. You can also see if this worked by using the Google structured data testing tool.
 - Go to the Google Search Console
 - In the left sidebar, click on “Other Resources”.
 - Click on the “Structured Data Testing Tool”.
 - Click on Fetch URL and type in the URL of your blog post.
 - Click on “FETCH & VALIDATE”. The right side of the page will show your structured data fields and if you have any errors.

Since I submitted the sitemap before fixing my structured data problems, the search console screenshot below shows how my code changes fixed 141 errors over several days.



These structured data errors took several days for google to reindex and clear the errors.

Create and Submit your Sitemap

Before proceeding, fix all problems before submitting a sitemap, because once google indexes a page, it seems to be impossible to remove it from google searches. I've also noticed that google only cares about new content, not updated content. Additionally, if you make updates, you can only notify google one page at a time by doing a Fetch as Google. That is very tedious, so get your content right the first time!

To generate a sitemap, I used the [Google XML Sitemaps plugin](#). In the "Sitemap Content" Settings for this plugin, do **NOT** include categories or tags, as shown in the following screenshot. Otherwise, you will be creating duplicate content in your sitemap. Not only does google downgrade sites with duplicate content, but your search results will prioritize the category or tag pages ahead of the individual posts.

Sitemap Content

WordPress standard content:

- ☒ Include homepage
- ☒ Include posts
- ☒ Include static pages
- ☐ Include categories
- ☐ Include archives
- ☐ Include author pages
- ☐ Include tag pages

Further options:

- ☒ Include the last modification time.

This is highly recommended and helps the search engines to know when your content has changed. This option affects *all* sitemap entries.

Once your structured data is okay, submit your sitemap in the Google Search Console. It took 1-2 weeks for google to finish indexing all the posts on soccermoviemom. There is a separate issue about google searching which I will cover in another how-to PDF on my website

Submitting a sitemap made the biggest difference in my SEO. In many searches for various movie titles, my movie reviews show up in the first few SERPs. This is a huge improvement, but SEO and "organic search" by itself can't make your site a success. You need to find other ways to drive people to your site. For example, if I tweet an interesting or useful post on twitter, the number of visitors to my site will increase for 1-3 days. Probably the fastest way to elevate your site would be to get a celebrity or well-known organization to tweet or retweet about you. Good luck!

Other Benefits and Future of Structured Data

Structured data is how your airline reservation now shows up in gmail, because gmail is reading the structured data in the incoming email and displaying it to you in an action-oriented format.

<http://searchengineland.com/next-generation-structured-data-taking-markup-next-level-216134>

Besides helping search rankings, the structured data may be used by search engines to create rich snippets, which are the descriptions in the SERP for each URI that is found. For reviews, Google may put your review rating into the snippet, although I haven't been able to get this to happen.

I also submitted my sitemap to [Bing Webmaster Tools](#), which is for both Bing and Yahoo searches. It took about 10 days for Bing to generate its reports. One advantage of Bing is that it will show you the search phrases that people use as well as where your link shows up in the SERP. However, as stated in the first page of this article, **Bing does not support JSON-LD**. So you still need to use meta tags and microdata on your website. I haven't started doing that yet.

I have noticed that Bing and Google search results can be quite different. Google tries to be more structured and analytical, while Bing results give me a loosey-goosey why-is-that-doc-in-this-SERP feeling. However, Microsoft rewards users for doing searches in Bing, so it is a search engine that we shouldn't ignore.

Creating Structured Data in JSON-LD for each Post

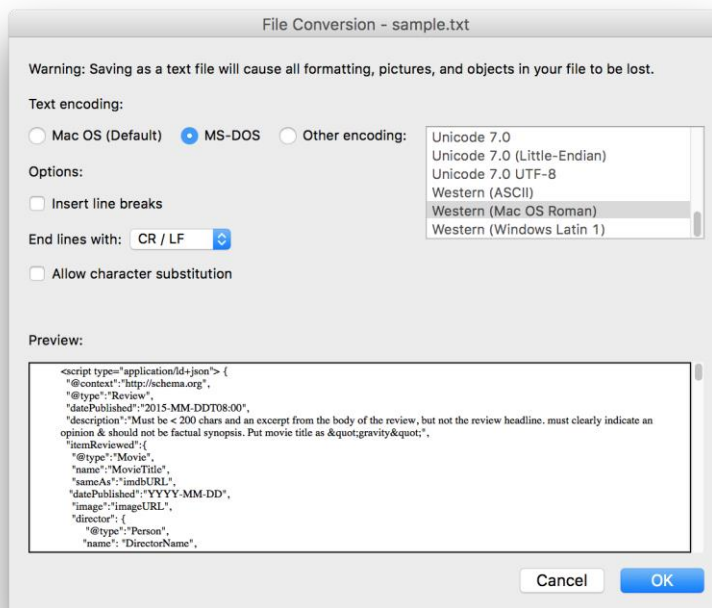
Because SoccerMovieMom is a Movie Review website, I followed the structured data specifications listed on the [Google developers web page for movie reviews](#). I added JSON-LD script to all of my posts and a few pages.

The wordpress.org forum was not very helpful on creating structured data. The google webmaster forum at <https://productforums.google.com/forum/#!forum/webmasters> is helpful in overall issues not specific to wordpress, and I have found many wordpress answers there. From this research, I came up with the following process to add JSON-LD to my posts.

However, as noted before, I have not really seen a benefit from doing all this work. I submitted my site to google's movie review system 3 times, but they never responded, and my search results still don't qualify for a movie review snippet. And as I said, JSON-LD is not supported by Bing/Yahoo. So i still need to add meta data and microdata.

The Process

1. I had to create a text file with the JSON-LD script for each of my wordpress posts. The boilerplate script is shown on the next page, which I saved on the Mac as an MS Word template (.dotx) file with the gray shaded lines as Word fields. The yellow highlighted areas are text that you would customize for your site and organization, but they would not be Word fields because these should stay the same for all your pages. To understand what these fields mean, please read the spec on the google developers page.
2. On my mac, the script for each post is created from the template and then saved as a .txt file. You must save with MS-DOS Text encoding selected:



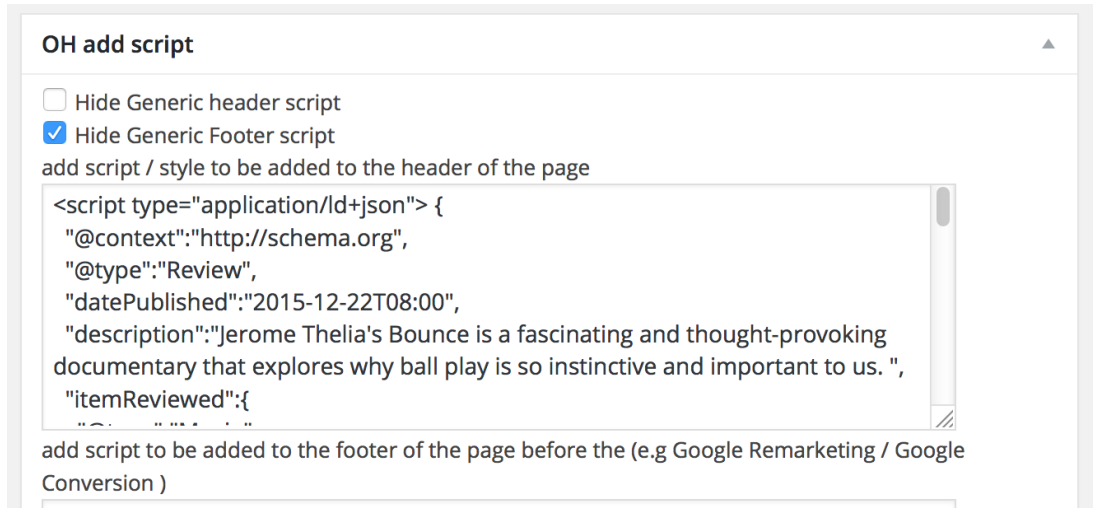
3. My template has two actor fields in it. If you only have one actor, or if you have more than one director, you will need to edit the text file you have generated, in textedit. Pay close attention to the left and right brackets [] when you list multiple values for an @type, as mistakes are easy.

```
<script type="application/ld+json"> {
  "@context":"http://schema.org",
  "@type":"Review",
  "datePublished":"2015-MM-DDT08:00",
  "description":"Must be < 200 chars and an excerpt from the body of the review, but not the review
headline. must clearly indicate an opinion & should not be factual synopsis. Put movie title as
"gravity"",
  "itemReviewed":{
    "@type":"Movie",
    "name":"MovieTitle",
    "sameAs":"imdbURL",
    "datePublished":"YYYY-MM-DD",
    "image":"imageURL",
    "director": {
      "@type":"Person",
      "name": "DirectorName",
      "sameAs":"http://www.imdb.com/name/directorURL"
    },
    "actor":[{
      "@type":"Person",
      "name":"ActorName",
      "sameAs":"http://www.imdb.com/name/actorURL"
    },{
      "@type":"Person",
      "name":"ActorName",
      "sameAs":"http://www.imdb.com/name/actorURL"
    }]
  },
  "reviewRating":{
    "@type":"Rating",
    "worstRating":1,
    "bestRating":10,
    "ratingValue":11
  },
  "url":"http://www.soccermoviemom.com/SMMpostURL",
  "author":{
    "@type":"Person",
    "name":"MJ Lee",
    "sameAs":"https://twitter.com/soccermoviemom" },
  "publisher":{
    "@type":"Organization",
    "name":"Soccer Movie Mom",
    "sameAs":"http://www.soccermoviemom.com" },
  "inLanguage":"en"
} </script>
```

4. You can double-check your script in Google Search Console > Other Resources > Structured Data Testing Tool. Copy and paste your script. The right side of the page will show your structured data fields and if you have any errors.

WARNING: Be careful with the use of the double quote character. In Word and Textedit, if you are copying and pasting text, you might insert a "smart quote" instead of a plain double quote. Google does not recognize the smart quote as a double quote, and it will not read your structured data correctly. I have turned off smart quote in my textedit preferences.

5. In previous versions of this document, I described steps where I compressed the script and inserted it at the end of the post. Grace Langlois on the Google Webmasters forum told me that JSON-LD should not be in the body of the post. It must instead be in the header or footer. I used the header, as in theory, the header is read by googlebot before the rest of the post.
6. I installed the plugin [OH Add Script to Individual Pages Header Footer](#) and placed the JSON-LD script in the Header, as shown below. It took me about 2 weeks to edit 130 posts in this way.



OH add script

☐ Hide Generic header script

☒ Hide Generic Footer script

add script / style to be added to the header of the page

```
<script type="application/ld+json"> {
  "@context":"http://schema.org",
  "@type":"Review",
  "datePublished":"2015-12-22T08:00",
  "description":"Jerome Thelia's Bounce is a fascinating and thought-provoking
documentary that explores why ball play is so instinctive and important to us. ",
  "itemReviewed":{
    "name":"Jerome Thelia's Bounce",
    "url":"http://www.soccermoviemom.com/2015/12/22/jerome-thelia-s-bounce/"
  }
}
```

add script to be added to the footer of the page before the (e.g Google Remarketing / Google Conversion)

7. After you save the post, verify this worked by again going to the Google Search Console > Other Resources > Structured Data Testing Tool. This time, click on Fetch URL and type in the URL of your blog post. Then click on "FETCH & VALIDATE". I suggest you do this whenever you edit the script, because it is very easy to insert an error. This was in fact, how I discovered my problem with the smart quote.
8. BTW, should have mentioned this earlier, but always backup your wordpress site before making big changes.

Tips on Google Snippets and Other Structured Data

On the Google Webmasters forum, Grace Massa Langlois made suggestions for the script content:

- For Movie type, Google recommends the **name** and **SameAs** URL references come from "IMDB, English Wiki page or official website. If the Movie is listed at IMDB then use that link, if not listed at IMDB then link to the English Wiki page, and so on."
- I also include **Director** as well as **Actor(s)** who I mention in the text of the review.
- "the snippet must come from the body of the review; descriptions containing snippets that aren't part of the review will prevent the review from being included in the feature."

Snippets have to be <200 characters. In my JSON-LD script, I specify text from my reviews to define a snippet that, in my opinion, is meaningful. But so far, it appears that google is ignoring my snippets and is instead using its own. Usually these are the first <200 chars from the beginning of the post.

- To easily count your characters, go to <http://www.lettercount.com/>

I recently had a post where google used the last line of the review for the snippet. The line they used exemplified Google's position that snippets should reflect opinions and not synopsis. But it was a very bland line:

- "**Bounce** is a very well-crafted film, and being a soccer fan is not required to enjoy it."

My preferred snippet was

- "Jerome Thelia's **Bounce** is a fascinating and thought-provoking documentary that explores why ball play is so instinctive and important to us."

It makes me wonder if Google's AI is really that good, or if some person at Google has a job where they select snippets all day.

Best wishes to all who read this. If you find it useful, please blog or link about it. Or, drop me a line via the Contact Me page on my website. Or follow me on twitter.

Cheers,
-mj lee
@soccermoviemom