

Understanding Schema.org

A deep dive for healthcare marketers.



GEONETRIC®

Executive Summary

In the early days of the internet, search engines matched users’ keywords with webpages that corresponded to those terms. It’s a quaint solution by today’s standards.

Search engines, especially Google, no longer limit themselves to pointing users to existing, published content on other sites. Instead, they are combining data on the fly to create entirely new information that answers searchers’ questions.

For example, when a prospective patient searches for a nearby provider, Google compiles relevant information into a local knowledge graph panel. In the example of Dr. Scott Coatsworth (see image 1), Google gathered data—photos, maps, address, hours, phone numbers, and reviews—from multiple sources. Unlike search results from the early days, there is no webpage with all of this information published on it. Google has created a dynamic answer that matches the user’s need.

Beyond just presenting information, Google has set their sights on enabling users to take action from within the search results. In fact, limited experiments are already allowing searchers to book appointments with some physicians from the search results page.

This has obvious implications for how health consumers find and evaluate their options for care. As healthcare marketers, how can we maintain control over the presentation of our brands, services, and providers?

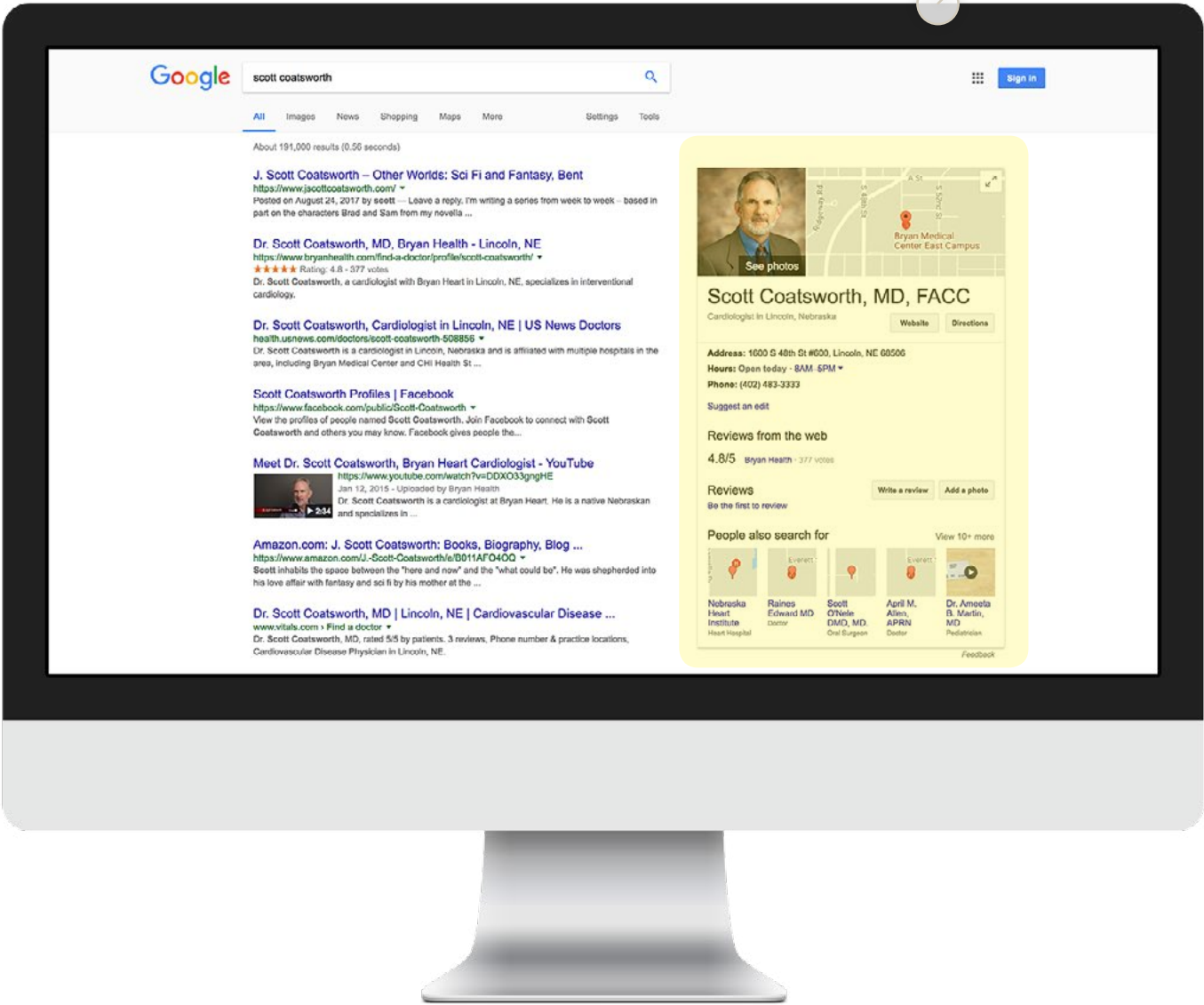
It’s still important to think about gaining visibility on search results pages. But the focus has expanded beyond individual webpages to optimizing the underlying data about those real world entities—the businesses, locations, and services—you’re charged with promoting.

Accomplishing this requires providing search engines with the kind of detailed, structured data that help to build the information and relationships within the knowledge graph. Schema.org is a standardized format for doing just that.

In this white paper, you’ll learn:

- The trends that are reshaping search engines
- The information that you can classify with Schema.org
- The Schema.org entities that apply to health information
- The three metadata formats you can use to implement Schema.org

Image 1: Knowledge Graph Panel



The focus has expanded beyond individual webpages to optimizing the underlying data about those real world entities – the business, locations, and services – you’re charged with promoting.

Principles of Search Engines

To say search engines are popular is an understatement. In fact, there are more than six billion searches a day worldwide.¹ And search engines are constantly improving how they answer all those queries. Great content and a good user experience are important pieces of how search engines evaluate websites and will become more important as search engines get more sophisticated.

Focusing on the User

Although search engines are evolving, one thing remains constant: their focus on users. Google, where more than four billion of those six billion searches start every day,² [lists this as their main goal](#).³

“Our first priority is to help our users find relevant, engaging answers for their search queries.”

To accomplish this, Google needs to identify **great content** that answers users’ questions and ensure a **good experience** for searchers.

Defining “Great Content”

Search engines define “great content” in terms of how they access that content and how they make it available to their users. In practice, this breaks down to how well they can crawl it, index it, and retrieve it:

- **Crawl-able: how well do search engines crawl a website?** The easier a search engine can find and request the content on your site, the better job they’ll do at knowing what content is there. This process involves many technical search engine optimization (SEO) elements: the server setup, configurations, robots.txt files, etc. We even see that things like page response time for the spider can influence how frequently your site gets indexed.

- **Index-able: how well do search engines understand the content on a website?** Search engines used to just break up the words on webpages, count them up, and use those results to determine relevancy for each search query. As we’ll discuss later, search engines are improving the ways they understand your content and index it.
- **Retrievable: how well do search engines match content to users’ search queries?** This is when we begin thinking about page ranking—whether other people vouch for your content by giving you links, sharing your content, or clicking it on the search engines results page.

Defining a “Good User Experience”

When users have a good experience with a search engine, they tend to keep using it. The search engines are aware of this, and that is why providing users with the best experience possible is so important. Notice the subtle language in Google’s statement: “help our users.” As a result, search engines value websites that provide a good user experience, and they devalue websites that don’t. For example, if it takes users a long time to load your website from the search results, they have a poor experience. Many times, users bounce back to that search results page. That’s why factors such as poor page speed and high bounce rates can potentially impact your rankings.

You should use the principles of **great content** and a **good experience** to vet your overall (SEO) strategy. As you review the user experience of your site, be sure to take a search perspective. Too often we think about user experience in terms of someone typing in your site’s URL and landing on the home page. Instead, consider a user typing a query on a search engine and landing on any page within your site. What will his or her experience be? Will it be easy to find relevant information on the page? Can they easily take action? Did the page load quickly and without distraction? Is all of that true on a mobile device?

If you align your goals in this way, you can achieve a synergy: the search engines keep users coming back and your rankings improve.

Use Google’s Search Console to identify common search phrases for your site and see which pages are ranking for those queries. It may not be the page you’d expect!

¹ <http://www.smartinsights.com/search-engine-marketing/search-engine-statistics>

² <https://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustommd=0>

³ <https://developers.google.com/search/docs/guides/intro-structured-data>



Search Engines: Past, Present, and Future

While the principles of search engines will remain consistent, we’ve seen an important evolution in how they implement these principles.

The Past: Searching for Documents

Traditionally, search engines looked for documents. They took the keywords that users entered and tried to find the best match with a webpage—much like going to the library and trying to find the page in the book that mentions a particular topic.

For many years, search engines focused on getting more precise at finding the webpage that matched a user’s query: finding the best, most popular, best-rated page of that book, and bringing that back to the user.

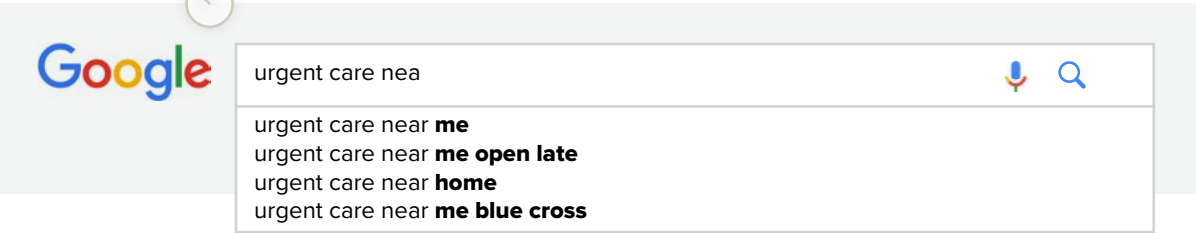
To rank well with this approach, marketers focused on the page title, meta description, meta keywords tag, and URL. While those are still important—minus the keywords tag, which was depreciated a number of years ago—search engines are expanding their focus.

The Present: Searching Entities

Search engines, especially Google, are shifting away from searching for documents that match keywords. Now, they’re increasingly searching for answers to users’ questions. To accomplish that, analyzing documents has given way to understanding entities and their relationships.

An “entity” is any person, place, or thing that exists in the real world. Google is trying to map all entities together, understanding what each entity is and how it relates to other entities. Rather

Image 2: Search Entities



than retrieving documents, search engines want to combine information on the fly from real-world entities and create answers to users’ questions. Thus, the results users see may be a combination of information that doesn’t actually exist in one document and, perhaps, has never existed before.

For example, if you start typing “urgent care,” Google is going to suggest that you might be looking for “urgent care near me,” “urgent care near me open late,” and more (see image 2).

These suggestions are complicated queries for specific information. Search engines need to know businesses that are classified as “urgent care.” They need to know the address for these locations, the user’s location, and what distance counts as “near me.” They need to know the hours of business and what times count as being “open late.” Then they need to connect all that information to provide an accurate answer.

This information doesn’t always exist together in a single document that a search engine can index and retrieve. This is something that is more like a database query—something that’s assembled on the fly based on exactly where users are located. By connecting these pieces of information, search engines are creating a map, or a graph, of entities. It includes people, places, and things—the details for each one and how they interrelate.

In addition, search engines are focusing on actionable information. Users search for “urgent care near me” because they need urgent care soon—or *right now*. As a result, search engines are thinking about not only how to combine the information, but also how to combine it in a way that enables someone to take action.

Google is creating a map (graph) of entities (real world people, places, and things) and actionable information about them.

As a result, search engines aren’t necessarily in the business of sending users to your site, although they definitely do that. If users can get an immediate answer to their question on the search results page that enables them to take their desired action, they’ll prefer that to choosing the right website and waiting for that page to load.

The Future: Evolving the Search Results Page

To see where search engines are going, let’s look at a popular content type: recipes. Lots of users must search for recipes because Google is really focused on enhancing those results. That makes it a great test case for how other content will look in the future.

Here’s an image with the evolution of search results pages for recipes (see image 3): On the left are the standard results. In the middle are rich snippets. These are the same results enhanced with things like images, star ratings, and more information. The next step in the evolution of search results is something Google calls rich cards. In this instance, Google displays them in a carousel toward the top of the page. You can see ratings and reviews, the approximate time to make the recipe, and who published it.

Job listings are another popular content type. Not only do you see them appear on search results, but also you can find and compare job postings on Google (see image 4).

The key to Google being able to present so much relevant data right in the search result? Schema.org. By using Schema.org markup, Google combines information from multiple sources to provide a richer experience. In this example, Google is able to provide the job title, company, location, and a brief description. You also see ratings for the employer, when the job was posted, and whether it’s full-time. Users can sort postings by category, title, city, date posted, type, company type, and employer.

Image 3: Evolution of Search Results

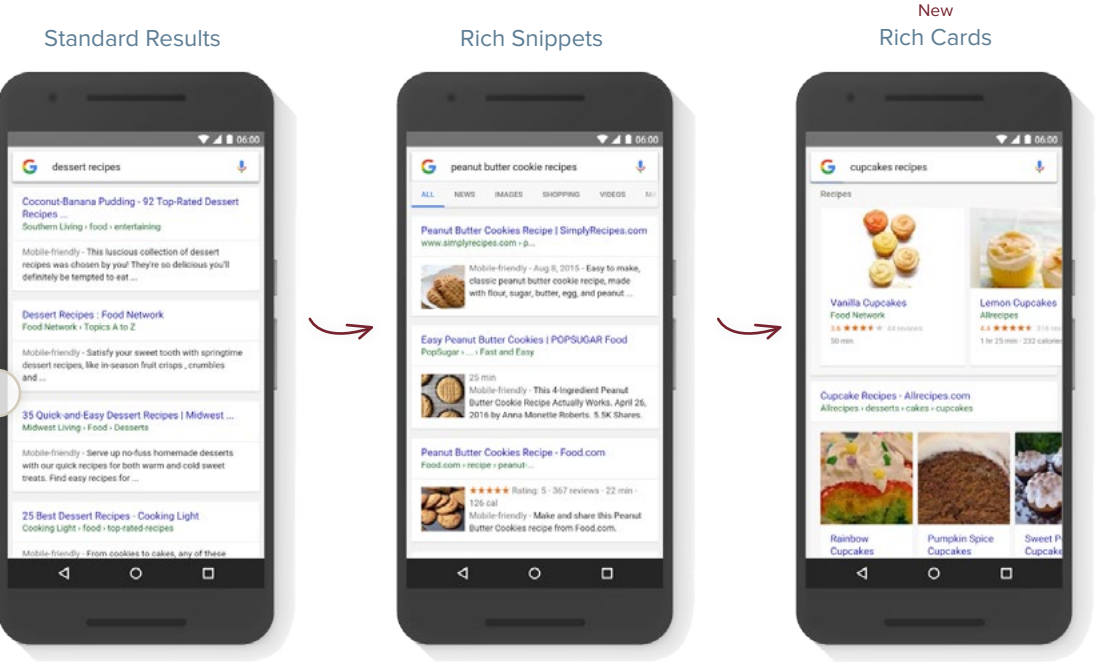


Image source: <https://webmasters.googleblog.com/2016/05/introducing-rich-cards.html>

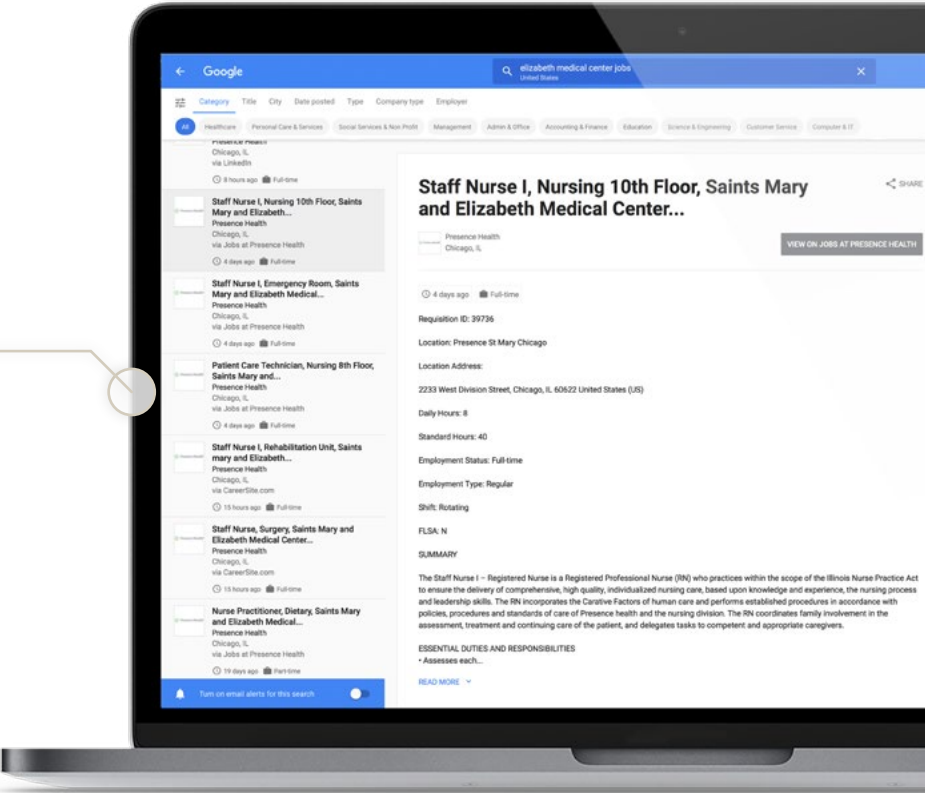


Image 4: Job Listings

What Does This Mean for Healthcare?

Today, users are comparing recipes and job postings before ever visiting a site. In the future, we may see Google try to control the presentation of other content types, such as physicians. While Google's exact plans are unclear, it's evident they will continue to grow their share of the online journey for users seeking answers to their questions. To capture patients' attention and provide a consistent brand experience from initial query to appointment request, healthcare marketers will increasingly rely on Schema.org markup. Read on to learn more about Schema.org entities and how to implement them at your organization.

An Introduction to Schema.org

Schema.org is a standardized vocabulary for providing information about a webpage and classifying its content. Schema.org began back in 2011, when Bing, Google, Yahoo, and Yandex came together with the goal of creating and supporting a common format for schema markup. While “schema” is a general term for a data model, “Schema.org” is the specific term for the markup language supported by the major search engines.

Entities

Schema.org establishes a hierarchy of entities that correspond to people, places, and things *in the real world*. The most general entity in Schema.org is a Thing, which has a few properties for describing it. Nested under the Thing entity, you have more specific entities that inherit the properties of a Thing, such as Action, CreativeWork, and more (see chart 1). Under each of those, you have other entities that get more and more specific. For example, under Event, there are multiple types of events, such as MusicEvent and SocialEvent.

Schema.org & Healthcare

Schema.org allows for the description of a wide range of healthcare and medical entities from drugs to symptoms to clinic locations. However, Google currently recognizes only a small subset. For marketers, this means entities that describe business aspects tend to be most useful. To understand how Schema.org entities relate, let’s look at a common scenario (see chart 2). To start, you have a hospital, which Schema.org considers a Thing, Place, Organization, and LocalBusiness. A hospital

inherits the properties of those entities. For example, as a LocalBusiness, a hospital has hours of operation; as an organization, a hospital has employees. In addition to these inherited properties, hospitals have unique properties of their own. Hospitals have services, which are categorized as MedicalProcedures, MedicalTherapies, and MedicalTests. Following that logic, a MedicalProcedure has a unique property: a MedicalSpecialty. A Physician has a MedicalSpecialty property as well. For example, a hip replacement surgery may be associated with an orthopedic specialty. Similar to a hospital, a Physician is a Thing, a Place, a LocalBusiness, and an Organization. While we usually think of physicians as people, Schema.org treats them as a business—the business of a physician’s individual practice. One property unique to physicians is they can accept new patients. You can define that property for each physician. Physicians can be connected to a MedicalProcedure both directly through AvailableService and indirectly through a shared MedicalSpecialty If you add the remaining connections, then you can start to create a comprehensive picture that pulls together the information in interesting ways (see chart 2). This allows search engines to understand how your physicians, locations, and services fit together, which helps them answer users’ queries. Using this graph, you can figure out possible answers to questions like:

- Which physicians perform hip replacement surgery and are accepting new patients?
- Which orthopedic surgeons are affiliated with the health system?

As you can start to see, using Schema.org entities helps search engines understand your content and how it all interrelates. This helps your content appear in coveted places, such as knowledge graphs, rich snippets, and local business results. But there are opportunities to connect the dots even further.

As you think about how to connect your organization, medical procedures, and medical specialties, you can differentiate among some of the competing entities. This will help you develop a strategy and avoid competing against yourself in search.

Connecting Entities

You can build relationships that connect Schema.org entities not only to each other on your website, but also to information in other systems, such as patient education materials and patient portals. That’s because many of these entity types can be described using medical vocabularies, such as ICD and RxNorm, that are likely used within your EMR and patient portal systems. These relationships can cross websites. For example, marking something as “SameAs,” you can show how an entity on one site is the same as an entity mentioned on another website. This could provide search engines with a better understanding of the importance of content. If you have many physicians and medical procedures referencing one of your key medical specialties, Google could understand that the specialty is important. How search engines will ultimately use Schema.org is yet to be seen. However, applying the relationships provided by Schema.org may help you to solve information architecture challenges common in healthcare organizations. For example, you may have locations within a health system that compete against each other, or service lines that overlap. The clean separation of entities in Schema.org may offer useful clues to tackle these challenges.

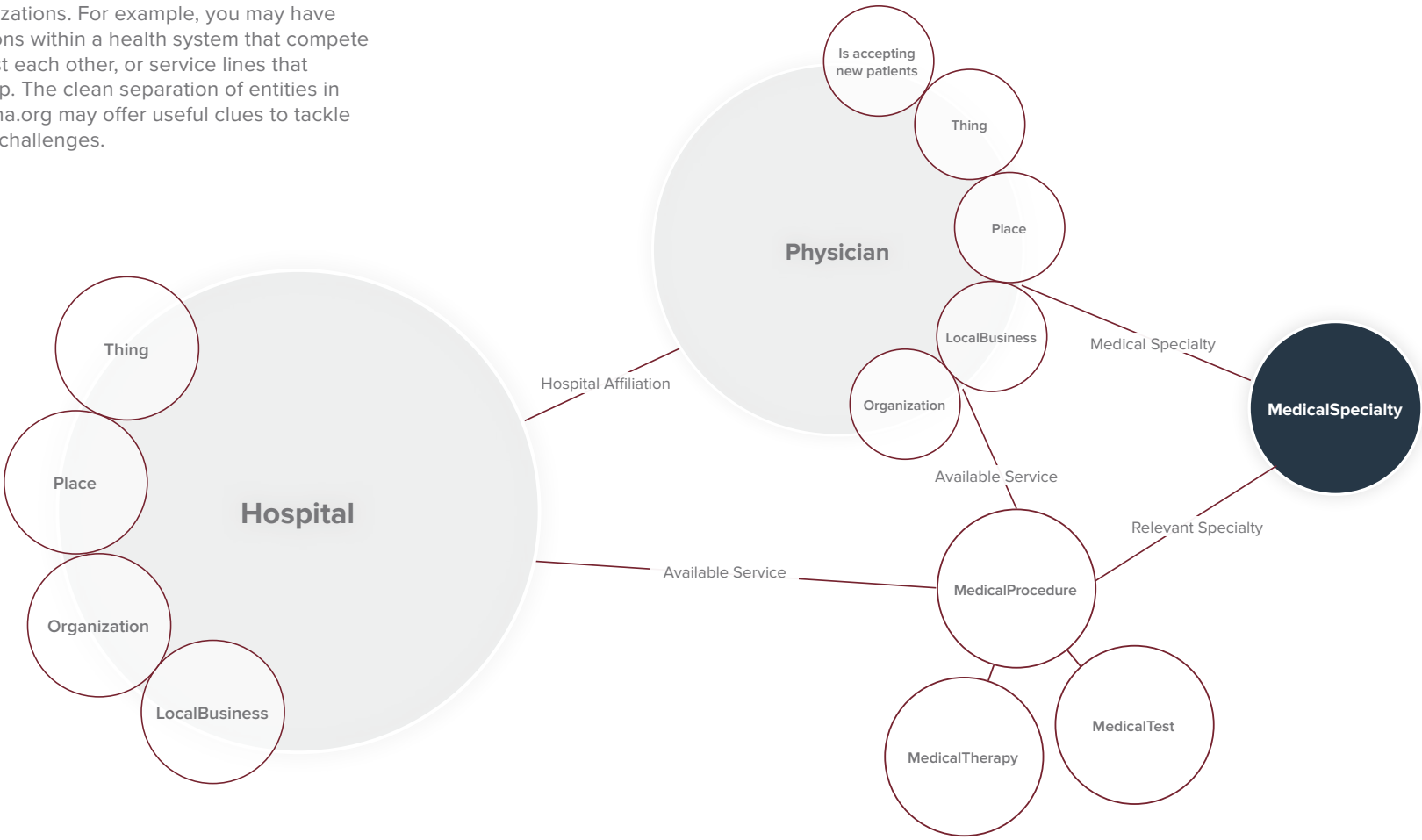


Chart 2: Schema.org Relationships for Healthcare

Schema.org Healthcare Entities

Now that you have a basic understanding of the potential of Schema.org, let's focus on the entities that are most relevant to healthcare marketers. Specifically, we'll look at six entities:

- Organization (specifically Hospital and MedicalOrganization)
- WebSite (plus Breadcrumbs and SiteNavigationElement)
- Videos
- LocalBusiness (specifically MedicalBusiness, Physician, and MedicalClinic)
- Reviews
- Events

We'll include examples for how each entity can help you enhance the search results for your healthcare organization.

Organization

The Organization entity controls important features within knowledge graphs and rich snippets. The main properties are (see images 5):

- **Logo.** This property controls the logo that appears.
- **Social Profiles.** The "SameAs" markup indicates some of the social profiles that you want associated with your organization.
- **Contact Point.** You can use this property to define phone numbers for customer service and billing support.

This entity is another opportunity to give Google the right information about your hospital, which will help keep your business listings up to date.

Website

In the WebSite entity, you can define:

- **PreferredSiteName.** In theory, you can define a short, friendly name for your site that Google would use in the green text under your HTML page title.
- **Sitelinks Searchbox.** You identify a site search function and that can help to influence whether that "sitelinks" search box shows (see image 6). The sitelinks searchbox may show anyway (at Google's whim), but the Schema.org markup will provide further encouragement for Google to show it.
- **Breadcrumbs.** You can add breadcrumb markup to influence some of the labeling in the results list. Rather than having the plain URL in the green text, you can include friendly text, such as "preparing for procedures" (see image 7). By editing this information, you can give searchers a better idea of the content on that page.
- **SiteNavigationElement.** You can define the main navigation to tell Google how each of those elements are part of your site navigation. While Google hasn't currently verified that they use this element, it's a simple way to provide an indication of this important information.

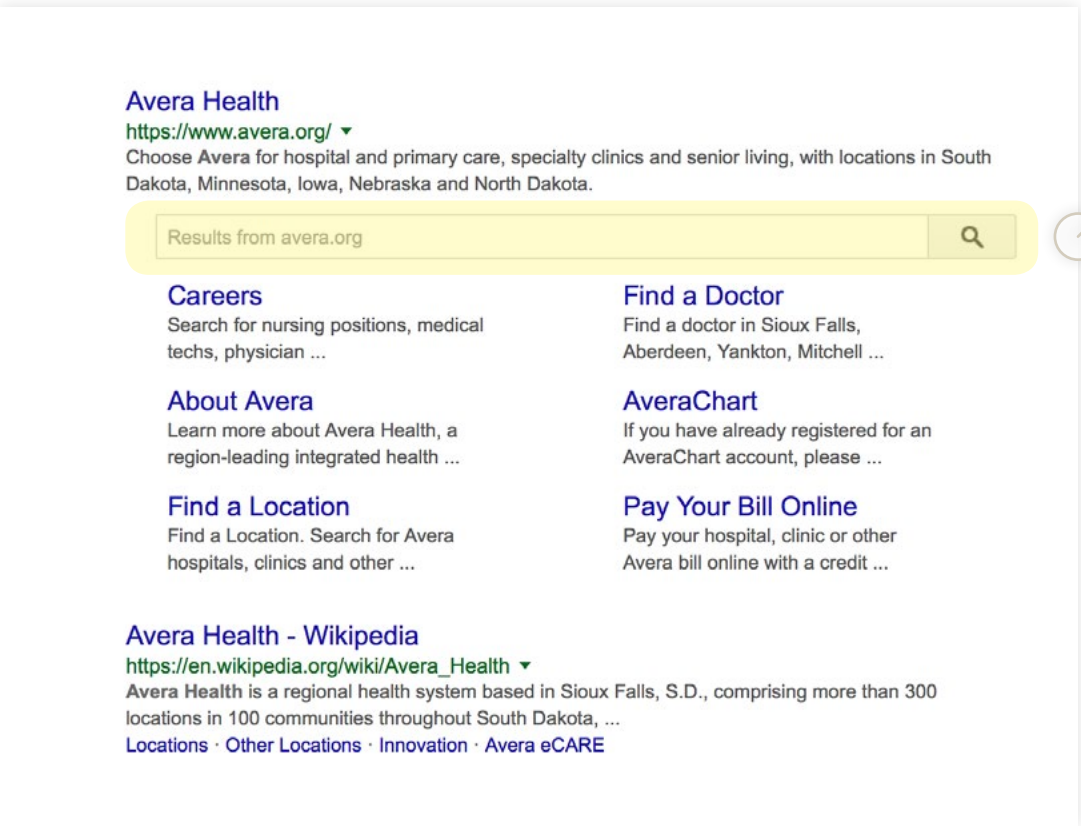


Image 6: SiteLinks Searchbox

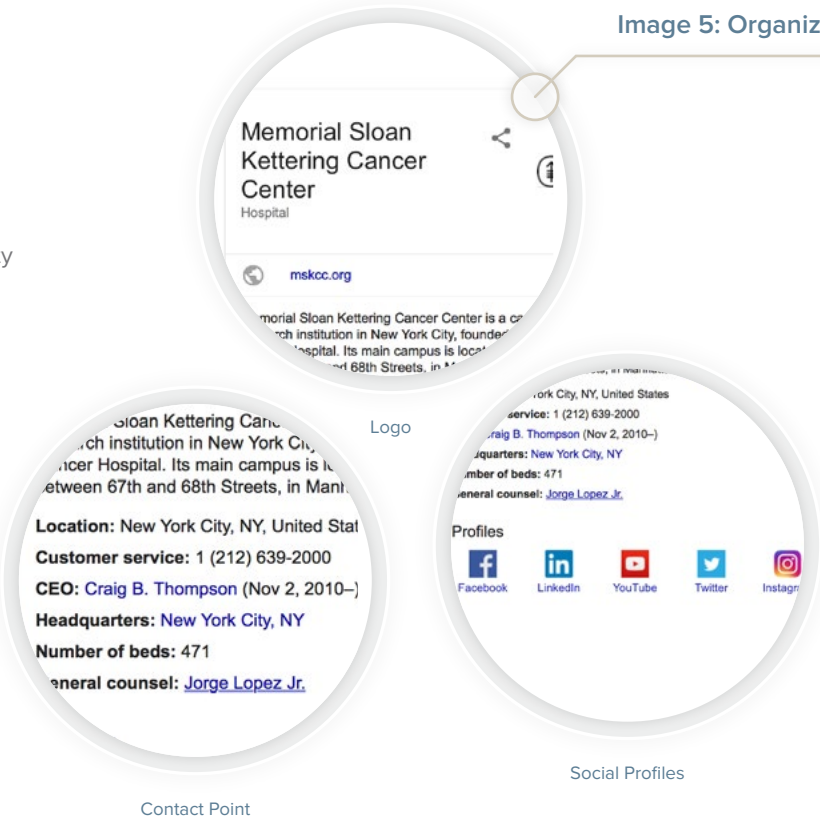


Image 5: Organization Entity



Image 7: Breadcrumbs

Videos

If you’re publishing videos with a branded experience on your site (they’re not hosted on YouTube), you may have experience with the challenge of getting your videos indexed and appearing on search results. Schema.org markup for video is one of the ways to accomplish that. It can help you to control the title and description that appears within the search results page (see image 8). In this example, you can tell the title and description were developed with SEO in mind.

LocalBusiness

LocalBusiness contains many sub-entities. For healthcare, it includes entities like Physician, MedicalOffice, and MedicalBusiness. Defining these entities can help keep your business listings up to date.

In the future, we may see additional capabilities with physicians, such as appointment scheduling. Google is running a closed beta with select organizations, including ZocDoc, an appointment-scheduling company, to start the scheduling process on the search results page (see image 9).

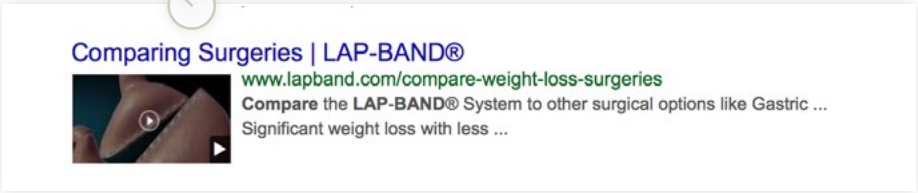
In addition, LocalBusiness is the parent entity that houses the reviews and aggregated reviews, such as star ratings.

Review

If you’re updating the Review entity, focus on the aggregate rating type. That’s the row of stars that appears in your listings and can have a huge impact on the click-through rate of your search results. As you can see in the example, having those five stars and 54 votes listed reflects well on the physician (see image 10). And it’s not only in the organic listing, but also in the local knowledge graph on the right hand side. They’re actually pulling in reviews from multiple websites.

Again, Google is creating the local knowledge graph on the fly. No webpage combines reviews from Gundersen Health System and HealthGrades in one place, but users can compare them side by side

Image 8: Videos



in Google's search results. And as you can see, it’s important to present your own data, because other sites may have a small number of reviews that skew the results.

Events

If you’re hosting health and wellness events, consider using the Event entity. This data can appear in rich snippets, rich cards, or other listings (see image 11). You see this across concerts and similar events. This is something to explore if you're doing a lot of event marketing.

Other Entities to Consider

- In addition to the entities described above, here are four others that you may find applicable, based on your particular strategies and tactics:
- 1. Courses.** If you’re doing training courses in addition to community events, you should consider the Course entity.
 - 2. Job Postings.** As we discussed earlier, job postings tend to be a popular content type. Many times, this is happening through your applicant tracking system. But it’s good to be aware of how these are appearing.
 - 3. Podcasts.** Podcasts seem to be getting big within healthcare, as they are everywhere else. Google is doing more to incorporate those in search results, especially on mobile.
 - 4. Recipes.** As we discussed earlier, recipes are a huge area of content on Google. If you’re trying to be competitive with healthy recipes, this is one entity to consider.

Image 10: Review

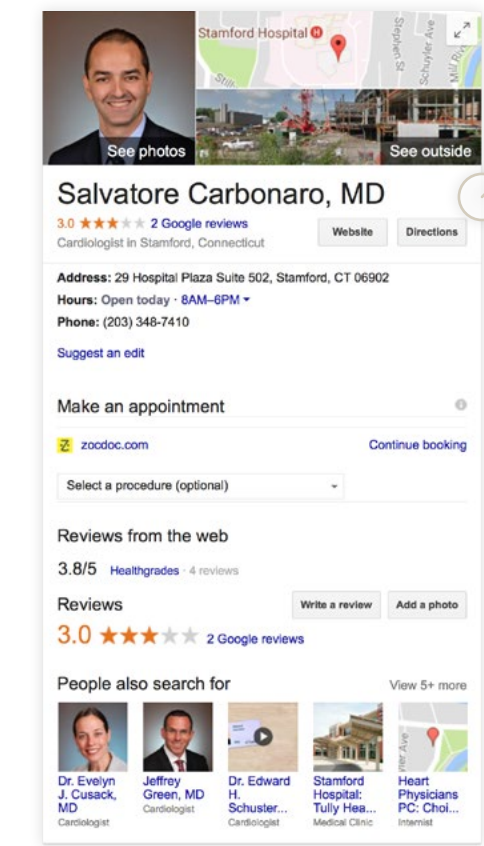


Image 9: LocalBusiness

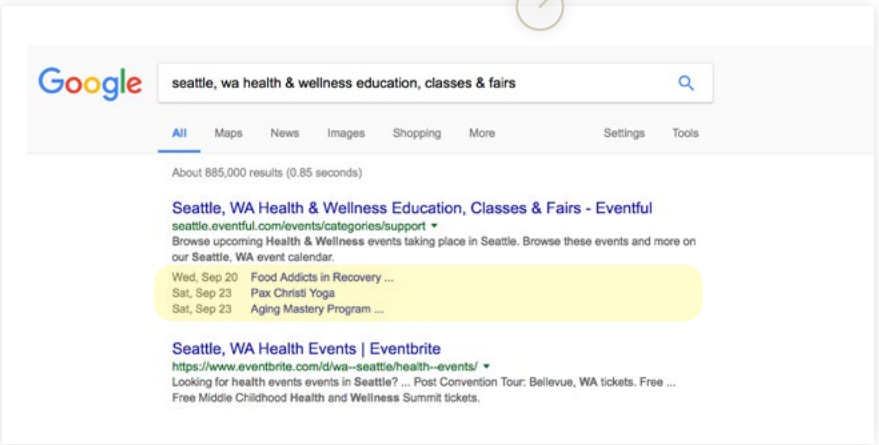
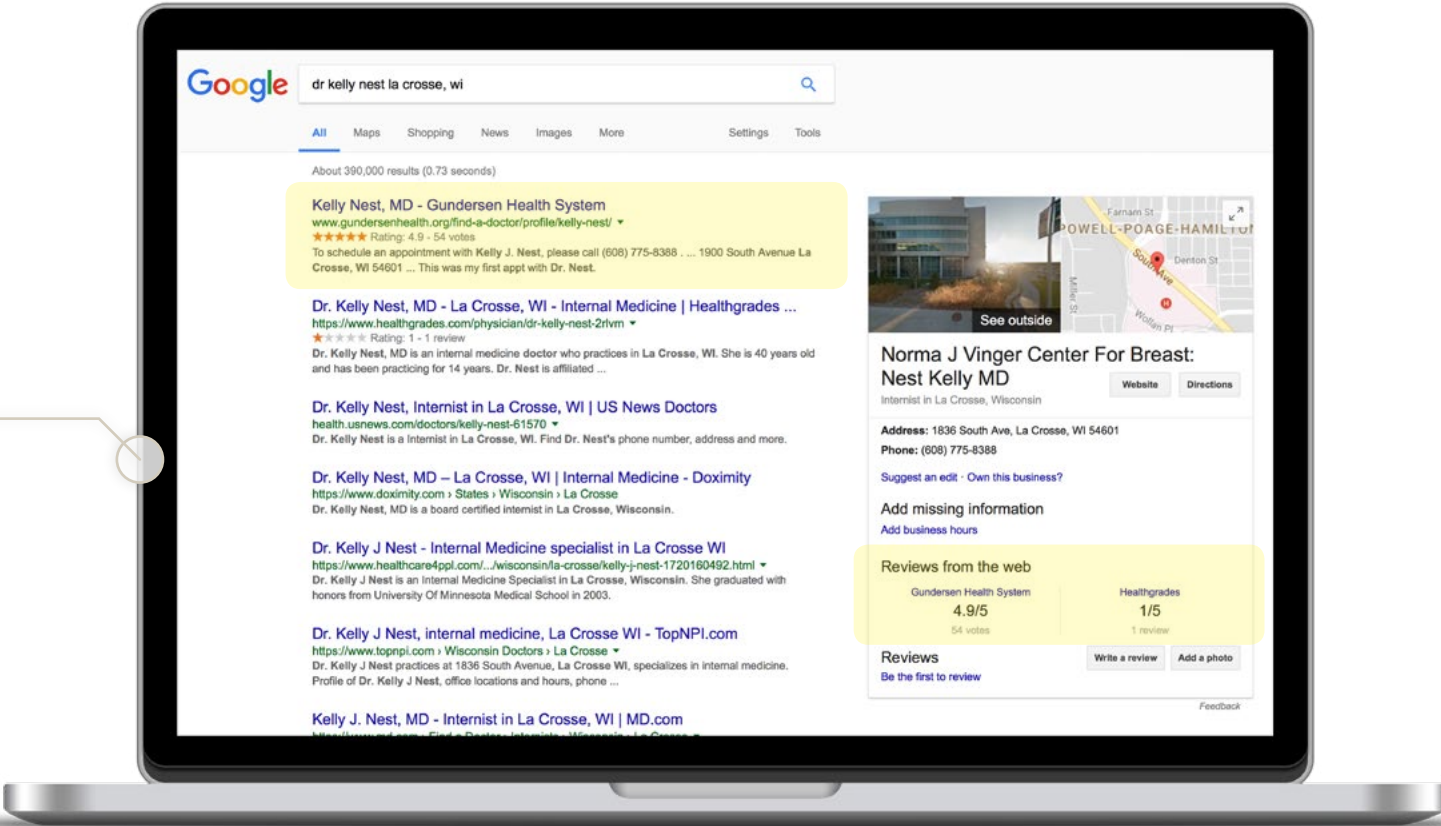


Image 11: Events



Implementing Schema.org

This section gets a little more technical as we'll discuss the different formats you can implement and how you can test your code.

Choosing a Schema.org Format

When you choose to describe content using Schema.org entities, you can use a number of formats. We distinguish two broad types of approaches: those that publish separate data, and those that are annotations to existing content within the HTML markup of a webpage. In the first type, the major language is a JavaScript format called JavaScript Object Notation for Linked Data (JSON-LD). JSON-LD is a way to define an object that represents links among different kinds of data.

For annotations in HTML, the two most common formats are Microdata and Resource Description Framework (RDFa). Using these two formats you can annotate content within your HTML markup, describing it using Schema.org.

For example, see the three languages with code describing the entity type of a website (see image 12). Each example includes the website's name, an alternate name, and the website's URL.

How to Pick a Language

JSON-LD, Microdata, and RDFa each have pros and cons. For example, if you're publishing a long page with JSON-LD, you're essentially duplicating that content, which increases the weight of the page and could impact page performance. Whereas with Microdata and RDFa, you're adding less code because you're marking up the HTML that's already there.

You will likely need to consider how the data you want to map to Schema.org standards gets published onto the page and what options

exist for duplicating the data or modifying markup. In some cases it may be much easier to publish JSON-LD, while in others modifying markup to add Microdata or RDFa may be the best (or only) option.

Google strongly recommends JSON-LD whenever possible. Although they support other formats, and will recognize them when indexing your site, Google has removed much of the documentation about Microdata and RDFa from their site. So, Google is seriously headed down the path of JSON-LD and wants to bring everybody along.

In contrast, Bing supports Microdata and RDFa. That may be fueling Google's choice to push forward with a newer format and create a competitive advantage, but that's only speculation.

A good indicator of where things are headed is to look at changes in what everyone else is doing on their sites. The chart on the bottom right compares the usage of the three markup languages based on the [Common Crawl](#). Common Crawl is a big set of data that's updated periodically, like a census of the web, and includes millions of sites. The data looks at the three formats for 2015, 2016, and the increase from year to year.

Microdata has been the most common format, and it increased significantly from 2015 to 2016. RDFa is the least common and it also grew the least. JSON-LD's growth outpaced the other two and is catching up to Microdata. With Google's support behind JSON-LD, it seems likely to surpass Microdata very quickly.

You can actually have both JSON-LD and Microdata going on the same page at the same time. They're not exclusive in any way. But it can be a lot more headache to maintain and can add to the overall page weight if there's a lot of content.

Useful Testing Tools

If you're adding Schema.org markup to your site, you want to become friends with [Google Search Console](#). Formerly called Google's "Webmaster Tools," Google Search Console enables you to monitor the information

Google finds when they crawl your site and the information they're finding. You can see the status of different types of structured data on the site, as well as Rich Cards that Google detects. You can get a sense of how many items Google is finding and if they're encountering any errors as they're parsing that information.

Google Search Console links into Google's other really useful tool, the [Structured Data Testing Tool](#). The tool breaks down all of the data, field by field, of what's on the page. And you can do this from a URL. You can even copy and paste markup in there if you're doing development and testing.

You can enter any URL, so you can test your site and look at what other sites are doing. If you see something cool on the search results page, put that URL into this tool and see what Schema.org entities it has on the page.

Image 12: Schema.org Formats

Publishing Separate Data

JSON-LD

```
{
  "@context": "http://schema.org",
  "@type": "WebSite",
  "name": "Your WebSite Name",
  "alternateName": "An alternative name for your WebSite",
  "url": "http://www.your-site.com"
}
```

Annotating HTML markup

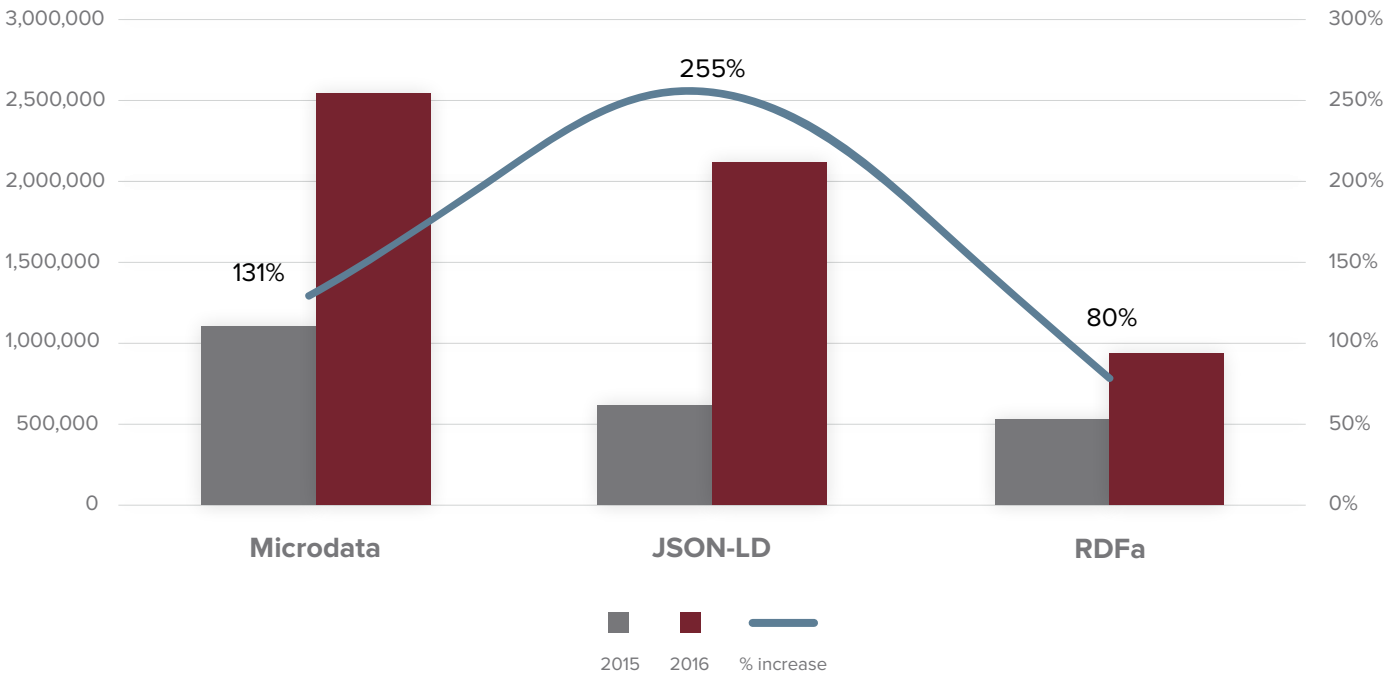
Microdata

```
<head itemscope itemtype=http://schema.org/WebSite>
<title itemprop=name>Your WebSite Name</title>
<link rel="canonical" href="https://example.com/" itemprop=url>
```

RDFa

```
<div vocab="http://schema.org/" typeof="WebSite" >
  <meta property="url" href="https://example.com/" />
  <meta property="name" value="Your WebSite Name" />
</div>
```

Domains Using Structured Data: 2015 vs. 2016



To choose the right language for your organization, analyze which search engine you value. Look at what your CMS supports. Consider your goals. Do some digging into the documentation on [Google](#), [Bing](#), and [Schema.org](#).

Schema.org is the Future

As search engines continue to evolve, it's important to make sure you provide a good brand experience on search results pages. Schema.org will help search engines understand your website's information and will help your organization stand out in knowledge graphs, rich snippets, local knowledge graphs, and more.

Implementing Schema doesn't need to be an “adopt it all at once” kind of process. It's something that most organizations are tackling as part of an ongoing process of improving the overall health of their websites.

To get started, look at some of your strategic priorities or your current projects, and add Schema.org to those. For example, if you're focusing on physicians, that may be a great opportunity to weave in some of the physician markup. Or if you need to clean-up page titles and descriptions—some of those basic elements within the page—add Schema.org as you're working through your site.

As you're working on Schema.org, keep in mind that the entities and properties that search engines support are always subject to change. Regardless, Schema.org is an important part of a major trend that will continue to grow. If you're looking to reach more health consumers online and provide a consistent brand experience across the patient journey, Schema.org is definitely something you can't ignore.

If you're looking for more information as you embark on this journey, check out these resources:

- [Schema.org website](#)
- [Google's documentation](#)
- [Bing's documentation](#)

About Geonetric

Geonetric helps healthcare brands thrive through effective marketing and distinctive websites. As a marketing agency and software developer with deep technical and creative expertise, Geonetric provides hospitals, health systems, and medical groups with a healthcare-specific content management system, hosting, marketing strategies, and creative services optimized for the unique needs of the healthcare industry.

If you're ready to find a partner to help with Schema.org or SEO, let's talk!

www.geonetric.com

855.848.7407

hello@geonetric.com



GEONETRIC[®]