

An Evolving Art Form: The Future of SEO

As we have noted throughout this book, SEO is about leveraging your company's assets through search engine–friendly content creation and website development (on-site) and targeted content promotion (off-site) in order to increase exposure and earn targeted traffic from organic search results. Therefore, the ultimate objective of the SEO professional is to make best use of organic search traffic as determined by various business goals of a website by guiding organizations through SEO strategy development, implementation, and ongoing measurement of SEO efforts.

This role will change as technology evolves, but the fundamental objectives will remain the same as long as “search engines” with nonpaid search results exist. The complexity of search will continue to grow, as all search engines seek to locate and index all of the world's digital information. As a result, we expect various levels of expansion and evolution in search within the following areas:

- Continued rapid expansion of mobile search, and with it, voice search, as the world continues to increase its demand for this capability.
- Large-scale investments in improved understanding of entities and relationships (semantic search) to allow search to directly answer more and more questions.
- Growth in social search as users begin to leverage social networks to discover new and interesting content and solutions to problems from their friends.
- Indexation of multimedia content, such as images, video and audio, including a better understanding of the content of these types of files.
- Indexation of data behind forms (something that Google already does in some cases—for example, if you use First Click Free).
- Continued improvements in extraction and analysis of JavaScript and AJAX-based content.

- Increased localization of search.
- Expanded personalization capabilities.

Mobile search is already driving an increasing demand for linguistic user interfaces, including voice recognition-based search. Voice search greatly improves the ease of use and accessibility of search on mobile devices, and this technology will continue to evolve and improve. In October 2011, Apple released the iPhone 4s with Siri. While most of its capabilities were already present in Google Voice Actions, Siri introduced a more conversational interface, and also showed some personality.

Business deals also regularly change the landscape. For example, on July 29, 2009, Microsoft and Yahoo! signed a far-reaching deal that resulted in Yahoo! retiring the search technology that powers Yahoo! Search and replacing it with Microsoft's Bing technology.¹ Bing also came to an agreement with Baidu to provide the English-language results for the Chinese search engine.²

Many contended that this deal would result in a more substantive competitor for Google. With Microsoft's deep pockets as well as a projected combined Bing/Yahoo! U.S. market share of just over 30%, Bing could potentially make a formidable competitor. However, six years later, it is not clear that there has been a significant shift in the search landscape as a result.

Since that time, it's been rumored that Yahoo! has undertaken projects (codenamed Fast-Break and Curveball) aimed at getting it back into the search game, but its lack of adequate search technology makes it seem unlikely any major changes will be made soon. But Yahoo! is not quitting: on November 20, 2014, it was announced that Yahoo! had reached a deal with Mozilla to replace Google as the default search engine for the Firefox web browser.³

Perhaps the bigger shift may come from the continuing growth of Facebook, which reports having 1.44 billion monthly active users as of April 2015 (http://bit.ly/fb_passes_1_44b), including about half the population of the United States and Canada.⁴ Facebook apps boast significant numbers of users as well; among the most popular are

1 Steve Lohr, "Microsoft and Yahoo Are Linked Up. Now What?", New York Times, July 29, 2009, http://bit.ly/ms_yahoo_linked_up

2 Michael Bonfils, "Bada Bing! It's Baidu Bing – English Search Marketing in China," Search Engine Watch, August 3, 2011, http://bit.ly/baidu_bing

3 Alexei Oreskovic, "Yahoo Usurps Google in Firefox Search Deal," Reuters, November 20, 2014, http://bit.ly/yahoo_usurps_google

4 Jim Edwards, "'Facebook Inc.' Actually Has 2.2 Billion Users Now — Roughly One Third Of The Entire Population Of Earth," Business Insider, July 24, 2014, <http://www.businessinsider.com/facebook-inc-has-22-billion-users-2014-7>

WhatsApp (500 million), Instagram (200 million), and Messenger (another 200 million).

As we noted in [Chapter 8](#), Bing's Stefan Weitz suggests that 90% of people use their friends to help them make one or more decisions every day, and 80% of people use their friends to help them make purchasing decisions.⁵ In addition, as we showed in [Chapter 8](#), Google+ has a material effect on Google's personalized search results. If Google can succeed in growing Google+, the scope of this impact could grow significantly over the next few years.

Third Door Media's Danny Sullivan supports the notion that some amount of search traffic may shift to social environments. In [a July 2011 interview with Eric Enge](#), he said: "I think search has his cousin called discovery, which is showing you things that you didn't necessarily know you wanted or needed, but you are happy to have come across. I think social is very strong at providing that."

Neither Stefan Weitz nor Danny Sullivan believes traditional web search is going away, but it seems likely that there will be some shifts as people discover other ways to get the information they want on the Web.

In fact, Stefan Weitz believes that search will get embedded more and more into devices and apps, and is focusing much of Bing's strategy in this direction, partly because the search engine recognizes that it won't win head to head with Google in what Weitz calls the "pure search space."⁶

These developments and many more will impact the role SEO plays within an organization. This chapter will explore some of the ways in which the world of technology, the nature of search, and the role of the SEO practitioner will evolve.

The Ongoing Evolution of Search

Search has come a long way, and will continue to progress at an increasingly rapid pace. Keeping up with these changes, the competitive environment, and the impact of new technology provides both a challenge and an opportunity.

The Growth of Search Complexity

Search has been evolving rapidly over the past decade and a half. At [the WSDM conference](#) in February 2009, Google Fellow Jeff Dean provided some interesting metrics that tell part of the story:

⁵ Emil Protalinski, "Bing Adds More Facebook Features to Social Search," ZDNet, May 16, 2011, <http://www.zdnet.com/blog/facebook/bing-adds-more-facebook-features-to-social-search/1483>

⁶ Brid-Aine Parnell, "Microsoft's Bing Hopes to Bag Market Share with ... Search Apps," The Register, November 4, 2014, http://bit.ly/bing_search_apps

- Google search volume had grown 1,000 times since 1999.
- Google has more than 1,000 times the machines it had in 1999.
- Latency dropped from less than 1,000 ms to less than 200 ms.
- Index update latency improved by about 10,000 times. Whereas updates took Google months in 1999, by 2009 Google was detecting and indexing changes on web pages in just a few minutes.

These are staggering changes in Google's performance power, and these are from six years ago. And of course this is just part of the changing search environment. Some of the early commercial search engines, such as Web Crawler, InfoSeek, and Alta Vista, launched in the mid-1990s. At that time, web search engines' relevancy and ranking algorithms were largely based on keyword analysis. This was a simple model to execute and initially provided pretty decent results.

However, there was (and is) too much money in search for such a simple model to stand. Spammers began abusing the weakness of the keyword algorithms by stuffing their pages with keywords, and using tactics to make them invisible to protect the user experience. This led to a situation in which the people who ranked first in search engines were not those who deserved it most, but were in fact those who understood (and could manipulate) the search algorithms the best.

By 1999, Google launched, and the next generation of search was born. Google was the search engine that most effectively implemented the concept of citation analysis (or link analysis). As we outlined earlier in the book, link analysis counted links to a website as a vote for its value. More votes represent more value, with some votes being worth more than others (pages with greater overall link authority have more "juice" to vote).

This created a situation that initially made the spammers' job more difficult, but the spammers began to catch up with this advance by purchasing links. With millions of websites out there, many of them with little or no revenue, it was relatively easy for the spammer to approach a site and offer it a nominal amount of money to get a link. Additionally, spammers could implement bots that surfed the Web, finding guest books, blogs, and forums, and leaving behind comments with links in them back to the bot owner's site.

The major search engines responded to this challenge as well. They took two major steps, one of which was to build teams of people who worked on ways to detect spamming and either discount it (by not attributing value) or penalize it. The other step was to implement an analysis of link quality that goes deeper than just the notion of PageRank. Factors such as anchor text, relevance, and trust became important as well. These factors also helped the search engines in their war against spam.

But the effort to improve search quality as well as fight spammers continued. Historical search result metrics, such as how many clicks a particular listing got and whether the user was apparently satisfied with the result she clicked on, have made their way into search algorithms. In 2008, then-Yahoo! Chief Scientist Jan O. Pederson wrote a [position paper](#) that advocated use of this type of data as follows:

Search engine query logs only reflect a small slice of user behavior—actions taken on the search results page. A more complete picture would include the entire click stream; search result page clicks as well as offsite follow-on actions.

This sort of data is available from a subset of toolbar users—those that opt into having their click stream tracked. Yahoo! has just begun to collect this sort of data, although competing search engines have collected it for some time.

We expect to derive much better indicators of user satisfaction by considering the actions post click. For example, if the user exits the clicked-through page rapidly [then] one can infer that the information need was not satisfied by that page.

Throughout the last few years, Google, Bing, and Facebook have invested in making use of social signals as well. As we discussed in [Chapter 8](#), links and mentions on social media sites could potentially be used to corroborate to the search engines that your content is worthy of ranking, though this does not appear to be in use today (other than the impact of Google+ on personalized search).

Consider also the notion of author authority. While Google killed the `rel="author"` tag, it can still create associations between a website and an author. If the author begins to establish a high level of authority, this could become a strong signal to the benefit of the site, even if that authority is established elsewhere.

In May 2007, Google made a big splash with the announcement of Universal Search (delivering what we have referred to in this book as “blended” search results). This was the beginning of the integration of all types of web-based data into a single set of search results, with data from video, images, news, blogs, and shopping search engines all being integrated into a single search experience.

This was only the very beginning of improved indexing for multimedia content. Google already has the ability to recognize the content of images in many cases. As we discussed in [Chapter 10](#), [Google Images](#) already allows you to drag an image into the search box and it will try to recognize that image’s content.

Search engines also can make use of other data sources, such as domain registry data to see who owns a particular website. In addition, they have access to analytics data, data from their browsers, data from their ad networks, data from their web search toolbars, and data from free WiFi and Internet access distribution to track actual web

usage on various websites. Other data sources will come into play too, such as data from wearables (e.g., Google Glass), embedded devices, and even your thermostat.⁷ Although no one knows how, or how much, the search engines use this type of data, these are additional information sources at their disposal.

Search engines continue to look for more ways to improve the quality of search results, and of the overall user experience within search. Google's efforts toward personalization allow it to look at a user's search history to get a better idea of what results will best satisfy that user. In 2008, Danny Sullivan summarized this entire evolution into four phases:⁸

- Search 1.0: keywords and text
- Search 2.0: link analysis
- Search 3.0: integration of vertical results
- Search 4.0: personalization

So, what makes up Search 5.0? Increased use of social media data appears to be one major possibility. The “wisdom of the crowds” may be becoming a factor in ranking, as discussed in [Chapter 8](#). Mike Grehan talks about this in his paper, “[New Signals to Search Engines](#)”. He summarizes the state of web search as follows:

We're essentially trying to force elephants into browsers that do not want them. The browser that Sir Tim Berners-Lee invented, along with HTML and the HTTP protocol, was intended to render text and graphics on a page delivered to your computer via a dial-up modem, not to watch movies like we do today. Search engine crawlers were developed to capture text from HTML pages and analyze links between pages, but with so much information outside the crawl, is it the right method for an always-on, ever-demanding audience of self producers?

Universal search was a step that acknowledged part of this problem by making all types of data available through web search. But many of these data types do not provide the traditional text-based signals that search engines rely on. Here is more from Mike Grehan's paper:

Signals from end users who previously couldn't vote for content via links from web pages are now able to vote for content with their engagement with

7 Marcus Wohlsen, “What Google Really Gets Out of Buying Nest for \$3.2 Billion,” *Wired*, January 14, 2014, <http://www.wired.com/2014/01/googles-3-billion-nest-buy-finally-make-internet-things-real-us/>

8 Danny Sullivan, “Search 4.0: Social Search Engines & Putting Humans Back in Search,” *Search Engine Land*, May 28, 2008, <http://searchengineland.com/search-40-putting-humans-back-in-search-14086>.

that content. You can expect that these types of signals will become a significant factor in the future.

User engagement and interaction with content provides the search engines with information about the value of content beyond links and on an array of data types (such as images and video), and this gives the engines more tools to improve search quality. This type of data was first confirmed as a factor in the rankings of videos on sites such as YouTube, which has become the third largest search engine for the United States (according to comScore, Hitwise, and Nielsen Online, and behind the combined Bing/Yahoo! results), and was confirmed as a factor in general web search in 2011. YouTube's ascent in search volume is particularly interesting because it is not a general web search engine, but one that focuses on a specific vertical—videos. This speaks to demand shifts taking place among the consumers of search results.

At the end of the day, the best results are likely to be provided by the best sites (there are exceptions; for example, for some search queries, the best results may be “instant answers”). The technology that has driven today's engines was based on two ranking signals: good keyword targeting and good links, but user experience and engagement are now significant factors too.

More data collection means more opportunity to win, even if your site doesn't conform flawlessly to these signals, and a better chance that if there is only one indicator you're winning, you could be in big trouble. Keywords and links will likely remain the primary ranking factors for the next few years, but the trend toward search engines using the new signals is steadily gaining momentum and strength.

Following these advances, what will be next? Artificial intelligence (AI) already plays a huge role in search. As Google's director of research Peter Norvig indicated to Eric Enge in [an October 2011 interview](#):

If you define AI as providing a course of action in the face of uncertainty and ambiguity, based on learning from examples, that's what our search algorithm is all about.

Further evidence of Google's fascination with AI can be seen in Google cofounder Larry Page's March 2014 TED presentation, [“Where's Google Going Next?”](#). In that presentation, he discussed how Google was applying machine learning to allow a computer to do things like play video games or, on their own, extrapolate the concept and nature of an entity like a cat. Google's goal is to apply that kind of learning to enable machines to understand the nature of information and how to present it to searchers.

One example of this is how Google Translate works. It examines and analyzes millions of real-world translations to learn how one language translates into another. It learns by example, rather than trying to learn an artificial set of rules. Learning a language based on grammar does not work, because language is far too dynamic and changing.

But learning from real-world usage does work. Using this technology, Google can offer instant translation across 58 languages.

Voice search works much the same way. Historically, speech recognition solutions did not work very well and required users to train the system to their voice. Google Voice uses a different approach, as noted by Peter Norvig in that same interview:

[For] Voice Search, where you speak your search to Google, we train this model on around 230 billion words from real-world search queries.

Google's Dominance

Thousands of posts, news articles, and analysis pieces have covered the central topic of battling Google's dominance in web search, but few have discussed the most telling example of the search giant's superiority. Many believe that the key to Google's success, and more importantly, a key component in its corporate culture, is its willingness and desire to return information relevant to a searcher's query as soon as possible.

Some also believe that Google's biggest impact in the search engine market is its advertising platform, which is the world's largest. By expanding its search capabilities, it is able to create a more enticing advertising platform through AdWords, AdSense, and its embeddable Google search box. Its incredible infrastructure allows for site speed and crawling depth that is unmatched by any other engine.

However, it goes a bit deeper than that. In late 2008, tests were performed in which users were asked which search engine's results they preferred for a wide variety of queries—long-tail searches, top-of-mind searches, topics about which their emotions ranged from great passion to total agnosticism. They were shown two sets of search results and asked which they preferred (see [Figure 15-1](#)).

Numerous tests like this have been performed, with a wide variety of results. In some, the brands are removed so that users see only the links. Testers do this to get an idea of whether they can win from a pure "quality" standpoint. In others, the brands remain to get an unvarnished and more "real-world" view. And in one particular experiment—performed many times by many different organizations—the results are swapped across the brands to test whether brand loyalty and brand preference are stronger than qualitative analysis in consumers.

It is this last test that has the most potentially intriguing results. Because in virtually every instance where qualitative differences weren't glaringly obvious, Google was picked as the best "search engine" without regard for the results themselves (see [Figure 15-2](#)).

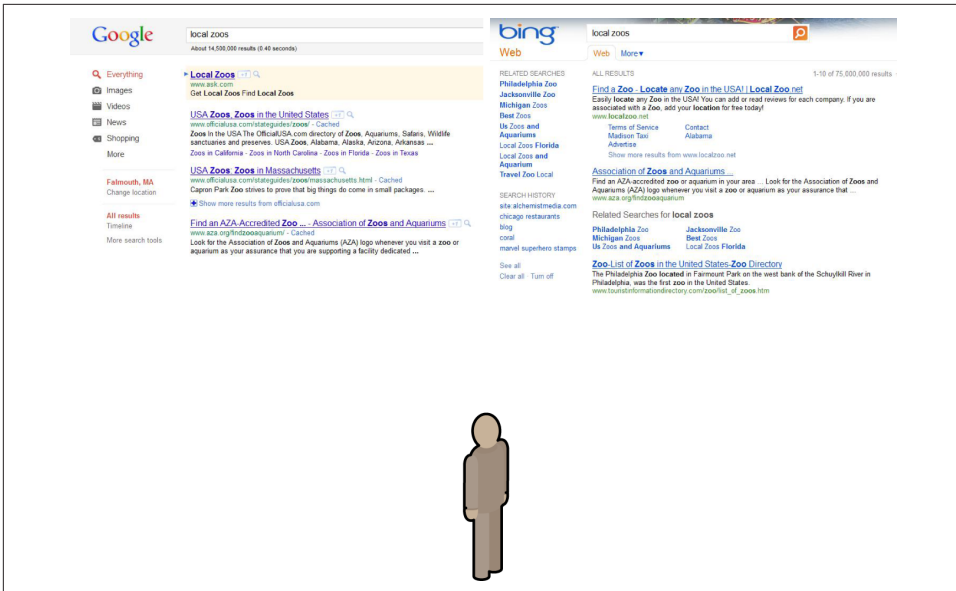


Figure 15-1. Comparing Google and Bing results



Figure 15-2. Results indicating that users may have a strong emotional preference for Google

Fundamentally, testers find (again and again) that the brand preference for Google outweighs the logical consideration of the search results quality.

Search engines that plan to take market share from Google are going to have to think differently. If Microsoft or a start-up search engine wants to take market share, it's going to have to think less like a technology company trying to build a better mouse-trap and more like a brand trying to win mindshare from a beloved competitor. How did Pepsi take share away from Coke? Or Toyota from Ford? That is beyond the scope of this book, but it is a process that can take more than a great idea or great technology. It requires a massive psychological shift in the way people around the world perceive the Google brand relative to its competition.

One strategy that Bing has embarked on is to leverage Facebook data in the Bing results; this data is shown in the right rail of the results page as a separate set of results. On October 13, 2010, the two companies signed a deal that provides Bing with substantial access to Facebook's data on shares, likes, and more.⁹

In addition, as noted earlier, Bing's Stefan Weitz talked about Bing's focus on embedding search technology into devices and apps:

The question is, where is search really going? It's unlikely we're going to take share in [the pure search] space, but in machine learning, natural language search...and how we can make search more part of living. For us, it's less about Bing.com, though that's still important. It's really about how we can instead weave the tech into things you're already doing.

Google's social media network, Google+, isn't the runaway success Google executives hoped it would be. Google purportedly continues to invest heavily in Google+¹⁰, yet in February 2014, the *New York Times* likened Google+ to a ghost town, citing Google stats of 540 million "monthly active users," but noting that almost half don't visit the site. A significant redirection with Google+ is likely inevitable

Also consider the official Google mission statement: "Google's mission is to organize the world's information and make it universally accessible and useful." It is already moving beyond that mission. For example, Google and NASA are working on new networking protocols that can work with the long latency times and low bandwidth in space.

Google is also pursuing alternative energy initiatives, which clearly goes beyond its mission statement. One example is its investments in self-driving cars. In addition, Google has ventures in office productivity software with Google Docs. These two ini-

⁹ Bing Blogs, "Bing Gets More Social with Facebook," October 13, 2010, <http://blogs.bing.com/search/2010/10/13/bing-gets-more-social-with-facebook/>

¹⁰ Kurt Wagner, "New Google+ Head David Besbris: We're Here for the Long Haul (Q&A)," Re/code, October 7, 2014, http://bit.ly/david_besbris

tatives have little to do with SEO, but they do speak to how Google is trying to expand its reach.

Another recent, now cancelled, venture for Google was its exploration into wearable technology, primarily in the form of **Google Glass**. Though Google Glass fell out of favor with much of the technology community, Glass was likely intended only as a first foray into the realm of wearable tech.¹¹ Despite the fact that Glass itself may have been released too soon, and in too much of a “beta” format, it dominated the early wearable search technology market. As wearable technology becomes more pervasive, expect Google to reintroduce another iteration of Google Glass in a bid to take a leading role in that market just as it has in mobile technology with Android.

Another potential future involves Google becoming a more general-purpose pattern-matching and searching engine. The concept of performing pattern matching on text (e.g., the current Google on the current Web) is only the first stage of an evolving process. Imagine the impact if Google turns its attention to the human genome and creates a pattern-matching engine that revolutionizes the way in which new medicines are developed. And the plethora of potential applications of pattern matching in the “real world” (such as disaster preparedness and logistics) is astounding; Google is uniquely poised to capitalize on this global opportunity with its understanding of our physical world from the ground (e.g., by Google’s self-driving cars), from in our homes (e.g., Google’s Nest acquisition), and from the sky via satellites and drones (e.g., Google’s acquisitions of Skybox Imaging and Titan Aerospace).

More Searchable Content and Content Types

The emphasis throughout this book has been on providing the crawlers with textual content semantically marked up using HTML. However, the less accessible document types—such as multimedia, content behind forms, and scanned historical documents—are increasingly being integrated into the search engine results pages (SERPs) as search algorithms evolve in the ways that the data is collected, parsed, and interpreted. Greater demand, availability, and usage also fuel the trend.

Engines Will Make Crawling Improvements

The search engines are breaking down some of the traditional limitations on crawling. Content types that search engines could not previously crawl or interpret are being addressed.

¹¹ Rachel Metz, “Google Glass Is Dead; Long Live Smart Glasses,” MIT Technology Review, November 26, 2014, http://bit.ly/google_glass_dead

In May 2014, Google announced that it had substantially improved the crawling and indexing of CSS and JavaScript content. Google can now render a large number of web pages with the JavaScript turned on so that its crawlers see it much more like the average user would. In October 2014, Google updated its Webmaster Guidelines to specifically advise that you do not block crawling of JavaScript and CSS files.

Despite these improvements, there are still many challenges to Google fully understanding all of the content within JavaScript or CSS, particularly if the crawlers are blocked from your JavaScript or CSS files, if your code is too complex for Google to understand, or if the code actually removes content from the page rather than adding it. Google still recommends that you build your site to “**degrade gracefully**”, which essentially means to build the site such that all of your content is available whether users have JavaScript turned on or off.

Another major historical limitation of search engines is dealing with forms. The classic example is a search query box on a publisher’s website. There is little point in the search engine punching in random search queries to see what results the site returns. However, there are other cases in which a much simpler form is in use, such as one that a user may fill out to get access to a downloadable article.

Search engines could potentially try to fill out such forms, perhaps according to a protocol where the rules are predefined, to gain access to this content in a form where they can index it and include it in their search results. A lot of valuable content is currently isolated behind such simple forms, and defining such a protocol is certainly within the realm of possibility (though it is no easy task, to be sure). This is an area addressed by Google in a **November 2011 announcement**. In more and more scenarios you can expect Google to fill out forms to see the content that exists behind them.

Engines Are Getting New Content Sources

As we noted earlier, Google’s stated mission is “to organize the world’s information and make it universally accessible and useful.” This is a powerful statement, particularly in light of the fact that so much information has not yet made its way online.

As part of its efforts to move more data to the Web, in 2004 Google launched an initiative to scan in books so that they could be incorporated into a **Book Search** search engine. This became the subject of a lawsuit by authors and libraries, but a **settlement was reached in late 2008**. In addition to books, other historical documents are worth scanning. To aid in that Google acquired reCAPTCHA (e.g., see <http://www.google.com/recaptcha>), and in December 2014, Google announced a major enhancement to how reCAPTCHA works, with the goal of making it much more user-friendly.

Similarly, content owners retain various other forms of proprietary information that is not generally available to the public. Some of this information is locked up behind log-ins for subscription-based content. To provide such content owners an incentive to

make that content searchable, Google came up with its First Click Free program (discussed earlier in this book), which allows Google to crawl subscription-based content.

Another example of new sources is metadata, in the form of markup such as Schema.org, microformats, and RDFa. This type of data, which is discussed in “[CSS and Semantic Markup](#)” on page 315, is a way for search engines to collect data directly from the publisher of a website. Schema.org was launched as a joint initiative of Google, Bing, and Yahoo! to collect publisher-supplied data, and the number of formats supported can be expected to grow over time.¹²

Another approach to this would be allow media sites and bloggers to submit content to the search engines via RSS feeds. This could potentially speed indexing time and reduce crawl burden at the same time. One reason why search engines may not be too quick to do this, though, is that website publishers are prone to making mistakes, and having procedures in place to protect against those mistakes might obviate the benefits.

However, a lot of other content out there is not on the Web at all, and this is information that the search engines want to index. To access it, they can approach the content owners and work on proprietary content deals, and this is also an activity that the search engines all pursue.

Another direction they can go with this is to find more ways to collect information directly from the public. Google Image Labeler was a program designed to do just this. It allowed users to label images through a game where they would work in pairs and try to come up with the same tags for the image as the person they were paired with. Unfortunately, this particular program was discontinued, but other approaches like it may be attempted in the future.

Multimedia Is Becoming Indexable

Content in images, audio, and video is currently not easily indexed by the search engines, but its metadata (tags, captioning, descriptions, geotagging data) and the anchor text of inbound links and surrounding content make it visible in search results. Google has made some great strides in this area. In [an interview with Eric Enge](#), Google’s director of research Peter Norvig discussed how Google allows searchers to drag an image from their desktop into the [Google Images search box](#) and Google can recognize the content of the image.

Or consider <http://www.google.com/recaptcha>. This site was originally used by Google to complete the digitization of books from the Internet Archive and old editions of the

¹² Google Official Blog, “Introducing Schema.org: Search Engines Come Together for a Richer Web,” June 2, 2011, http://bit.ly/intro_schema_org.

New York Times. These have been partially digitized using scanning and OCR software. OCR is not a perfect technology, and there are many cases where the software cannot determine a word with 100% confidence. However, reCAPTCHA is assisting by using humans to figure out what these words are and feeding them back into the database of digitized documents.

First, reCAPTCHA takes the unresolved words and puts them into a database. These words are then fed to blogs that use the site's CAPTCHA solution for security purposes. These are the boxes you see on blogs and account sign-up screens where you need to enter the characters you see, such as the one shown in [Figure 15-3](#).

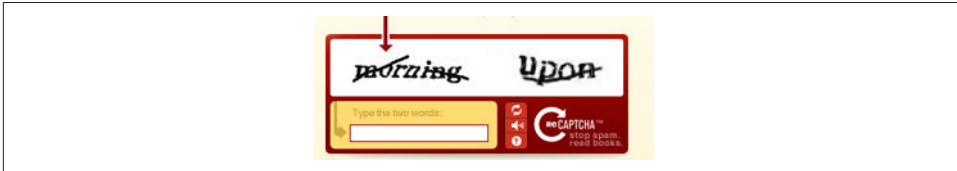


Figure 15-3. *ReCAPTCHA screen*

In this example, the user is expected to type in *morning*. However, in this case, Recaptcha.net is using the human input in these CAPTCHA screens to translate text that wasn't recognized by OCR. It makes use of this CAPTCHA information to improve the quality of this digitized book.

Google used this methodology for years, but has since changed reCAPTCHA to focus more on images instead.¹³ This new approach is image based, and is intended to help Google with its computer vision projects.

Similarly, speech-to-text solutions can be applied to audio and video files to extract more data from them. This is a relatively processing-intensive technology, and it has historically had trouble with many accents and dialects, so it has not yet been universally applied in search. Apple's Siri and Google Voice are leading the charge in addressing this issue. In addition, if you upload a video to YouTube, you can provide a search-indexable caption file for it, or request that Google use its voice recognition technology to attempt to autocaption it.

The business problem the search engines face is that the demand for information and content in these challenging-to-index formats is increasing exponentially. Search results that do not accurately include this type of data will begin to be deemed irrelevant or wrong, resulting in lost market share and declining ad revenues.

¹³ Frederic Lardinois, "Google's reCAPTCHA (Mostly) Does Away With Those Annoying CAPTCHAs," TechCrunch, December 3, 2014, http://bit.ly/googles_recaptcha.

The dominance of YouTube is a powerful signpost of user interest. Users want engaging, multimedia content—and they want a lot of it. For this reason, developing improved techniques for indexing such alternative content types is an urgent priority for the search engines.

Interactive content is also growing on the Web, with technologies such as AJAX. In spite of the indexing challenges these technologies bring to search engines, their use is continuing because of the experience they offer broadband users. The search engines are hard at work on solutions to better understand the content wrapped up in these technologies as well.

Over time, our view of what is “interactive” will likely change dramatically. Two- or three-dimensional first-person shooter games and movies will continue to morph and become increasingly interactive. Further in the future, these may become full immersion experiences, similar to the holodeck on *Star Trek*. You can also expect to see interactive movies where the audience influences the plot with both virtual and human actors performing live. These types of advances are not the immediate concern of today’s SEO practitioner, but staying in touch with where things are headed over time can provide a valuable perspective.

More Personalized, Localized, and User-Influenced Search

Personalization efforts have been under way at the search engines for some time. As we discussed earlier in this book, the most basic form of personalization is to perform an IP location lookup to determine where the searcher is located, and tweak the results based on the searcher’s location. However, the search engines continue to explore additional ways to expand on this simple concept to deliver better results for each user. It is not yet clear whether personalization has given the engines that have heavily invested in it better results overall or greater user satisfaction, but their continued use of the technology suggests that, at the least, their internal user satisfaction tests have been positive.

Indeed, Google has continued to expand the factors that can influence a users’ personalized search. For example, one major signal it uses is the user’s personal search history. Google can track sites a user has visited, groups of related sites the user has visited, whether a user has shared a given site over social media, and what keywords the user has searched for in the past. All of these factors may influence the given personalized search engine results page.

User Intent

As just mentioned, Google personalized results are tapping into user intent based on previous search history, and serving up a mix not just of personalized “blue links” but

of many content types, including maps, blog posts, videos, and local results. The major search engines already provide maps for appropriate location searches and the ability to list blog results based on recency as well as relevancy. It is not just about presenting the results, but about presenting them in the format that maps to the searcher's intent.

User Interactions

One area that will see great exploration is how users interact with search engines. As the sheer amount of information in its many formats expands, users will continue to look to search engines to be not just a search destination, but also a source of information aggregation whereby the search engine acts as a portal, pulling and updating news and other content based on the user's preferences.

Marissa Mayer, then Google's VP of Location and Local Services (now CEO of Yahoo!), made a particularly interesting comment that furthers the sense that search engines will continue their evolution beyond search:

I think that people will be annotating search results pages and web pages a lot. They're going to be rating them, they're going to be reviewing them. They're going to be marking them up...

Indeed, Google already offers users the ability to block certain results. Mayer's mention of "web pages" may be another reason why the release of **Google Chrome** was so important. Tapping into the web browser might lead to that ability to annotate and rate those pages and further help Google identify what content interests the user. As of February 2014, **StatCounter** showed that Chrome's market share had risen to an impressive 44%.

Chris Sherman, executive editor of Search Engine Land, offered up an interesting approach that the search engines might pursue as a way to allow users to interact with them and help bring about better results:

[F]ind a way to let us search by example—submitting a page of content and analyzing the full text of that page and then tying that in conjunction with our past behavior...

New Search Patterns

This is all part of increasing the focus on the users, tying into their intent and interests at the time of search. Personalization will make site stickiness ever more important. Securing a position in users' history and becoming an authoritative go-to source for information will be more critical than ever. Winning in the SERPs will require much more than just optimizing for position, moving toward an increased focus on engagement.

Over time, smart marketers will recognize that the attention of a potential customer is a scarce and limited quantity. As the quantity of information available to us grows, the amount of time we have available for each piece of information declines, creating an attention deficit. How people search, and how advertisers interact with them, may change dramatically as a result.

In 2008, *The Atlantic* published an article titled “[Is Google Making Us Stupid?](#)”. The thrust of this article was that Google was so powerful in its capabilities that humans need to do less (and less!). Google has made huge advances since this article, and this is a trend that will continue. After all, who needs memory when you have your “life-stream” captured 24/7 with instant retrieval via something akin to Google desktop search or when you have instant perfect recall of all of human history?

These types of changes, if and when they occur, could transform what today we call SEO into something else, where the SEO of tomorrow is responsible for helping publishers gain access to potential customers through a vast array of new mechanisms that currently do not exist.

Growing Reliance on the Cloud

Cloud computing is transforming how the Internet-connected population uses computers. Oracle founder Larry Ellison’s vision of thin-client computing may yet come to pass, but in the form of a pervasive Google operating system and its associated, extensive suite of applications. Widespread adoption by users of cloud-based (rather than desktop) software and seemingly limitless data storage, all supplied for free by Google, will usher in a new era of personalized advertising within these apps.

Google is actively advancing the mass migration of desktop computing to the cloud, with initiatives such as [Google Docs](#), [Gmail](#), [Google Calendar](#), [Google App Engine](#), and [Google Drive](#). These types of services encourage users to entrust their valuable data to the Google cloud. This brings them many benefits (but also concerns around privacy, security, uptime, and data integrity). In May 2011 Apple also made a move in this direction when it announced [iCloud](#), which is seamlessly integrated into Apple devices.

One simple example of a basic application for cloud computing is the notion of backing up all your data. Most users don’t do a good job of backing up their data, making them susceptible to data loss from hard drive crashes and virus infections. Companies investing in cloud computing will seek to get you to store the master copy of your data in the cloud, and keep backup copies locally on your devices (or not at all). With this approach you can more easily access that information from multiple computers (e.g., at work and home).

Google (and Apple) benefits by having a repository of user data available for analysis—which is very helpful in Google’s quest to deliver ever more relevant ads and search

results. It also provides multiple additional platforms within which to serve advertising. Furthermore, regular users of a service such as Google Docs are more likely to be logged in a greater percentage of the time when they are on their computers.

The inevitable advance of cloud computing will offer more and more services with unrivaled convenience and cost benefits, compelling users to turn to the cloud for their data and their apps.

Increasing Importance of Local, Mobile, and Voice Search

New forms of vertical search are becoming increasingly important. Search engines have already embraced local search and mobile search, and voice-based search is an area in which all the major engines are actively investing.

Local Search

Local search was an active business before the advent of the Internet, when Yellow Pages from your phone company were the tool of choice. As the World Wide Web gained prominence in our lives, the Yellow Pages migrated online. These types of local search sites have themselves evolved to leverage more of the unique nature of the Web. Some of the major players are [CitySearch](#), [Local.com](#), [YellowPages.com](#), and [Superpages](#).

As the major search engines evolved, they integrated sophisticated mapping technology into their systems, and then began to map in local business data as well. You can find these local search engines at <http://maps.google.com/>, <http://local.yahoo.com>, and <http://www.bing.com/maps>. These search engines map in some of their business data from the same types of data sources that drive the Yellow Pages websites, such as [Acxiom](#), [InfoUSA](#), and [Localeze](#), but they also supplement that data by crawling the Web.

One of the big challenges facing the local search engines is *map spam*, the proliferation of spam results in local search. Here are two blog posts that provide examples:

- [“Illusory Laptop Repair—A Most Elegant Google Places Hack”](#)
- [“Google not Good Enough in Controlling Local Map Spam”](#)

Resolving these issues is obviously critical for the search engines and is a major area of investment for them. We believe that local search is an area whose importance will continue to grow. With a very large number of searches having local intent, this is potentially a major frontier for developments in search.

Over time, you can expect that these problems will be worked out. Local search is already extremely powerful, but you can anticipate that the search engines will control the flow of customers in cyberspace and the real world, because customers will be finding merchants through search. You can also expect that more and more of these local searches will take place on mobile devices. Success in local search will be what makes or breaks many businesses.

Consumers are becoming increasingly reliant on local search and its auxiliary services—street maps, directions, satellite views, street views, 3D visualizations (Google Earth), enhanced listings, user reviews, and ratings. Through Google Mobile’s image recognition (formerly Google Goggles), Google even has the capability of searching based on photos you take with your mobile device. Augmented reality, where we see metadata juxtaposed on top of the physical world via a “looking glass,” is around the corner, and Google Glass is just a hint of what is to come.

With the help of social apps such as **Yelp**, **FourSquare**, and Facebook, mobile users can locate nearby friends, and search for special offers and discounts when they “check in” at certain stores or restaurants. Such tools guide consumers’ movements and their decision-making processes; this means the search engines and other search services are in a powerful position to manage the flow of customers from the Web to brick-and-mortar businesses. So, in a way, the search engines are in the logistics business, building the connection between online and offline.

Mobile Search

According to the June 2015 Ericsson Mobility Report, there were 7.1 billion world-wide mobile subscribers in 2014, with a forecast of 9.2 billion subscribers by 2020. (<http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf>).¹⁴ As a result, the opportunity for mobile search has begun to grow rapidly. A Google study showed that search engines were the most visited sites on mobile devices, with 77% of those surveyed reporting their use.¹⁵ The same study showed that 9 out of 10 searchers took action as a result of conducting a search.

In addition, as of October 2014, Google began experimenting with using the mobile friendliness of a site as a ranking factor, and in May 2015 Google made Mobile Friendliness a significant ranking factor impacting results for users searching from Smart-

¹⁴ More People Around the World Have Cell Phones Than Ever Had Land-Lines,” Quartz, February 25, 2014, <http://qz.com/179897/more-people-around-the-world-have-cell-phones-than-ever-had-land-lines/>

¹⁵ Julie Batten, “Newest Stats on Mobile Search,” ClickZ, May 23, 2011, http://bit.ly/newest_mobile_stats

phones.¹⁶ This provides strong incentive to have a mobile search strategy in place (see “The Mobile Landscape” on page 635 for more on this)!

The convenience of being able to get the information you need while on the go is just way too compelling. Why be tied to a desk if you do not have to be? Further, in many countries, freedom from a desktop is a necessity because users have much less personal space than their U.S. counterparts.

The ever-expanding versatility and power of tablet and mobile devices—from indispensable utility apps, to immersive multimedia players, to massively multiplayer online games, to paradigm-shifting hardware advances such as the iPhone’s multi-touch display, proximity sensor, GPS, and gyroscopes—will fuel this growth. The *network effect*, whereby the value of the network grows by the square of the size of the network (Metcalfe’s Law), gives further incentive for users to migrate to their mobile devices as more and more apps allow them to interact with their peers in increasingly more interesting ways.

The small keyboard/typing surface is currently a limitation, but Apple’s and Google’s voice-based solutions have already made great strides in replacing the keyboard as the input device of choice. The advent of the Linguistic User Interface (LUI) will continue to revolutionize mobile search, and with the LUI comes a whole new set of skills that will be required of the SEO practitioner.

In addition, wearable devices such as the Apple Watch are gaining traction. Voice technology is improving to the point where the absence of a keyboard will not be a significant limitation.

Voice Recognition Search

When users are mobile they must deal with the limitations of their mobile device, specifically the small screen and small keyboard. These make web surfing and mobile searching more challenging than they are in the PC environment.

Voice search is a great way to improve the mobile search experience. It eliminates the need for the keyboard, and provides users with a simple and elegant interface. Speech recognition technology has been around for a long time, and the main challenge has always been that it requires a lot of computing power.

Several examples already exist:

16 Barry Schwartz, “Google May Add Mobile User Experience To Its Ranking Algorithm,” Search Engine Land, October 8, 2014, <http://searchengineland.com/google-may-add-mobile-user-experience-ranking-algorithm-205382>

Google Voice Search

This is a free service from Google that enables you to perform a Web search based on text-to-speech voice queries instead of traditional text-typed queries.

Google Mobile App for iPhone

Included in this application is voice searching capability. You can speak your query into the application and the results are displayed on-screen.

Apple's Siri

What made Siri a significant step forward after its launch was its use of more natural human speech (Google offers similar capabilities now). This was the start of a movement toward the device learning the user, rather than vice versa.

Microsoft Cortana

In April 2014, Microsoft unveiled its personal assistant offering, known as Cortana. Cortana is being promoted as offering more capabilities for learning your personal preferences than the competing options. Currently, it is designed to work only on Windows Phones.

Processing power continues to increase, even on mobile devices, and the feasibility of this type of technology is growing. This should be another major area of change in the mobile search landscape.

In addition, it is reasonable to expect that voice recognition technology will be applied to the actual recorded content of audio and video files, to determine content and aid in ranking these media types within search results—something that will likely lead to video “script” optimization as an added component of video SEO.

Increased Market Saturation and Competition

One thing you can count on with the Web is continued growth. Despite its constantly growing index, a lot of the pages in Google may be low-quality or duplicate-content-type pages that will never see the light of day. The Web is a big place, but one where the signal-to-noise ratio is very low.

One major trend emerges from an analysis of Internet usage statistics. **According to Miniwatts Marketing Group**, 84.9% of the North American population uses the Internet, so there is not much room for growth there. In contrast, Asia, which already has the most Internet users (1.4 billion) has a penetration rate of only 34.7%. Other regions with a great deal of opportunity to grow are Africa, the Middle East, and Latin America.

This data tells us that in terms of the number of users, North America is already approaching saturation. Europe has some room to grow, but not that much. However, in Asia, you could see two times that growth, or 2 to 3 billion users! The bottom line is

that a lot of Internet growth in the coming decade will be outside North America, and that will provide unique new business opportunities for those who are ready to capitalize on that growth.

With this growth has come an increasing awareness of what needs to be done to obtain traffic. The search engines are the dominant suppliers of traffic for many publishers, and will continue to be for some time to come. For that reason, awareness of SEO will continue to increase over time. Here are some reasons why this growth has continued:

The Web outperforms other sales channels

When organizations look at the paths leading to sales and income (a critical analysis whenever budgets are under scrutiny), the Web almost always comes out with one of two assessments. Either it is a leading sales channel (especially from an ROI perspective), or it is the area with the greatest opportunity for growth. In both scenarios, digital marketing (and, in correlation, SEO) take center stage.

It is the right time to retool

Established companies frequently use down cycles as a chance to focus attention inward and analyze themselves. Consequently, there's a spike in website redesigns and SEO along with it.

Paid search drives interest in SEO

Paid search spending is still reaching all-time highs, and when companies evaluate the cost and value, there's a nagging little voice saying, "75%+ of the clicks do not even happen in the ads; use SEO."

SEO is losing its stigma

Google is releasing SEO guides, Microsoft and Yahoo! have in-house SEO departments, and the "SEO is BS" crowd have lost a little of their swagger and a lot of their arguments. No surprise—solid evidence trumps wishful thinking, especially when times are tough.

Marketing departments are in a brainstorming cycle

A high percentage of companies are asking the big questions: "How do we get new customers?" and "What avenues still offer opportunity?" Whenever that happens, SEO is bound to show up near the top of the "to be investigated" pile.

Search traffic will be relatively unscathed by the market

Sales might drop and conversion rates might falter a bit, but raw search traffic isn't going anywhere. A recession doesn't mean people stop searching the Web, and with broadband adoption rates, Internet penetration, and searches per user consistently rising, search is no fad. It is here for the long haul.

Web budgets are being reassessed

We've all seen the news about display advertising falling considerably; that can happen only when managers meet to discuss how to address budget concerns. Get 10 Internet marketing managers into rooms with their teams and at least 4 or 5 are bound to discuss SEO and how they can grab that "free" traffic.

Someone finally looked at the web analytics

It is sad, but true. When a downturn arrives or panic sets in, someone, maybe the first someone in a long time, checks the web analytics to see where revenue is still coming in. Not surprisingly, search engine referrals with their exceptional targeting and intent matching are ranking high on the list.

Although more and more people are becoming aware of these advantages of SEO, there still remains an imbalance between paid search and SEO. The SEMPO Annual State of Search Survey (http://bit.ly/2015_state_of_search; membership is required to access the report) includes information suggesting that as much as 90% of the money invested in search-related marketing is spent on PPC campaigns and only about 10% goes into SEO.

This suggests that either SEO could see some growth to align budgets with potential opportunity, or firms that focus solely on SEO services had better diversify. SEO budgets continue to expand as more and more businesses better understand the mechanics of the Web. In the short term, PPC is easier for many businesses to understand, because it has more in common with traditional forms of marketing. Ultimately, though, SEO is where the most money can be found, and the dollars will follow once people understand that.

SEO as an Enduring Art Form

Today, SEO can be fairly easily categorized as having five major objectives:

- Make content accessible to search engine crawlers.
- Find the keywords that searchers employ (understand your target audience) and make your site speak their language.
- Build content that users will find useful, valuable, and worthy of sharing. Ensure that they'll have a good experience on your site to improve the likelihood that you'll earn links and references.
- Earn votes for your content in the form of editorial links and social media mentions from good sources by building inviting, shareable content and applying classic marketing techniques to the online world.

- Create web pages that allow users to find what they want extremely quickly, ideally in the blink of an eye.

Note, though, that the tactics an SEO practitioner might use to get links from editorial sources have been subject to rapid evolution. We now turn to content marketing instead of link building. In addition, a strong understanding of how the search engines measure and weight social engagement signals is increasingly important to SEO professionals.

One thing that you can be sure about in the world of search is change, as forces from all over the Web are impacting search in a dramatic way. To be an artist, the SEO practitioner needs to see the landscape of possibilities for an individual website, and pick the best possible path to success. This currently includes social media optimization expertise, local search expertise, video optimization expertise, an understanding of what is coming in mobile, and more. That's a far cry from the backroom geek of the late 1990s.

No one can predict what the future will bring and what will be needed to successfully market businesses and other organizations on the Web in 2 years, let alone 5 or 10. However, you can be certain that websites are here to stay for a long time, and that websites are never finished and need continuous optimization just like any other direct marketing channel. SEO expertise will be needed for a long time—and no one is better suited to map the changing environment and lead companies to success in this new, ever-evolving landscape than today's SEO practitioner.

The Future of Semantic Search and the Knowledge Graph

In **Chapter 6**, we explored the state of semantic search and the Knowledge Graph as we know it today. All the search engines are continuing to investigate these types of technologies in many different ways, though Google is clearly in the lead. The Knowledge Vault is just one of many initiatives that Google is pursuing to make progress in this area.

Part of the objective is to develop a machine intelligence that can fully understand how people evaluate the world, yet even this is not sufficient. The real goal is to understand how *each* human being evaluates the world, so that the results can be fully personalized to meet each individual's needs.

Not only do search engines want to give you the perfect answer to your questions, they also want to provide you with opportunities for exploration. Humans like to conduct research and learn new things. Providing all of these capabilities will require a special type of machine intelligence, and we are a long way from reaching those goals.

There are many components that go into developing this type of intelligence. In the near future, efforts focus largely on solving specific problems. For example, one such problem is maintaining the context of an ongoing conversation. Consider the following set of queries, starting with *where is the empire state building?* in [Figure 15-4](#).

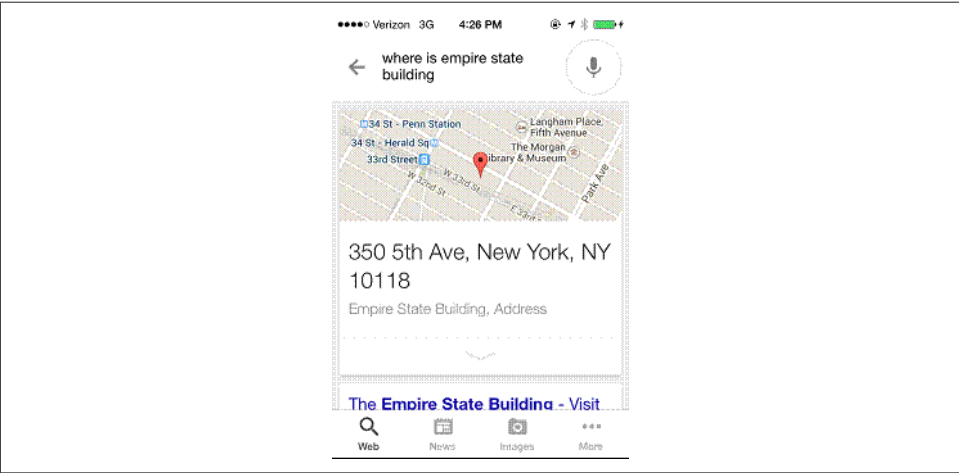


Figure 15-4. Response to the query “where is the empire state building?”

Notice how the word *the* was dropped in the query display. [Figure 15-5](#) shows what happens when you follow this query with the one on *pictures*.

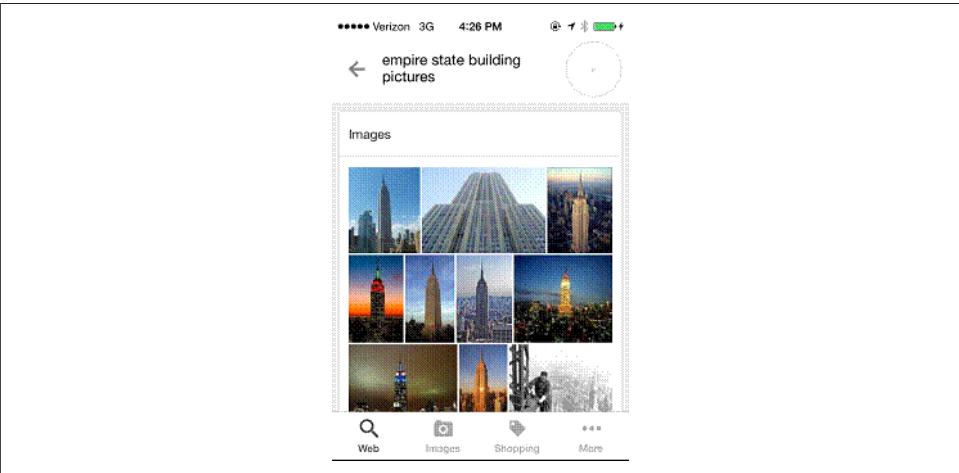


Figure 15-5. Response to the query “pictures”

Notice again how the query was modified to *empire state building pictures*. Google has remembered that the prior query was specific to the Empire State Building, and did

not require us to restate that. This query sequence can continue for quite some time. **Figure 15-6** shows the result when we now ask *who built it?*

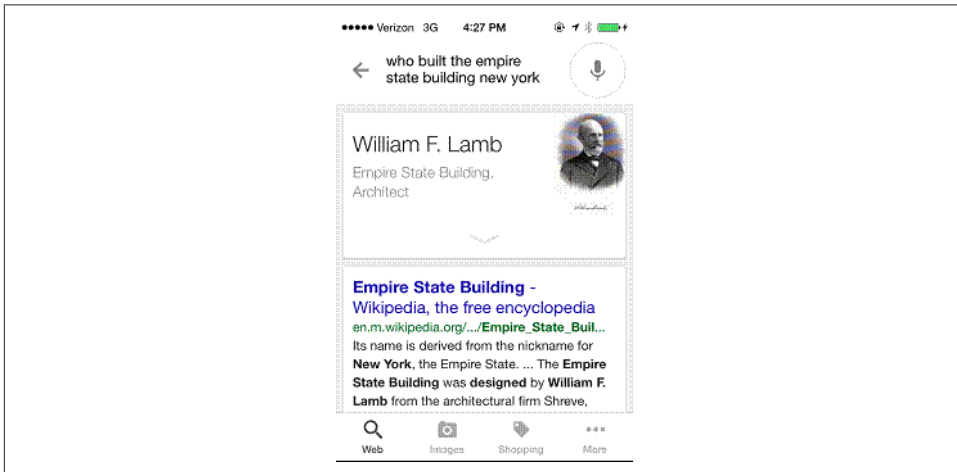


Figure 15-6. Response to the query “who built it?”

Once again, the query was dynamically modified on the fly, and Google has remembered the context of the conversation. **Figure 15-7** shows what happens when we now try the query *restaurants*.

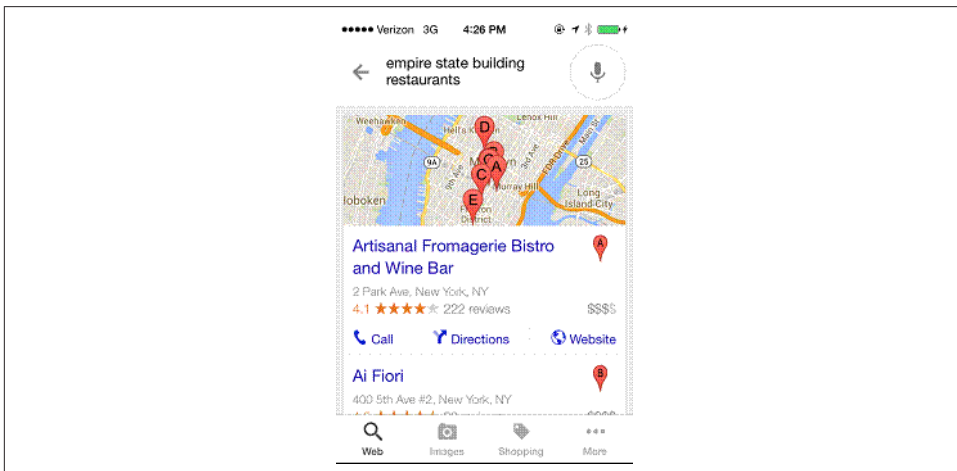


Figure 15-7. Response to the query “restaurants”

Finally, we can follow this query with the more complex query *give me directions to the third one*, as shown in **Figure 15-8**.

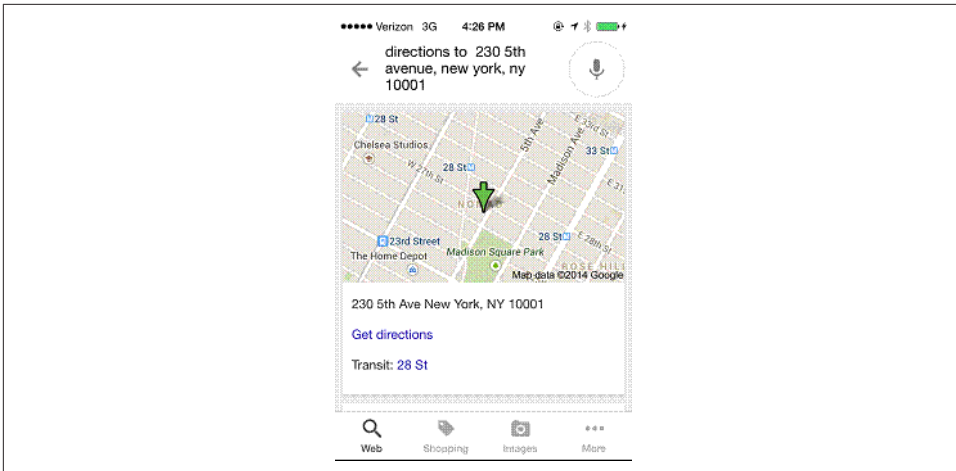


Figure 15-8. Response to the query “give me directions to the third one”

This entire sequence of queries is quite complicated, capped off by Google’s understanding of the concept of the “third one” in the final query. Even though this is very sophisticated, it is nonetheless an example of a point solution to a specific problem.

To truly model human thought, search engines will need to build machines that can reason like humans, are able to perceive the world around them, understand how to define objectives and make plans to meet them, and can independently work to expand their knowledge.

Many disciplines are involved in developing artificial intelligence, such as computer science, neuroscience, psychology, philosophy, and linguistics. Even just understanding linguistics is a major challenge, as there are thousands of different languages in the world, and this by itself multiplies the complexity of the task.

The computing power to take on these challenges does not yet exist, so developing expanding capabilities is a major piece of the puzzle. For example, Google is pursuing efforts to build a quantum computer.¹⁷

In the near term, we can expect changes in search results to come in the form of more point solutions to specific problems. As the understanding of how to model human intelligence expands, and as processing power grows with it, we may see much more significant changes, perhaps 5 to 10 years down the road.

¹⁷ Tom Simonite, “Google Launches Effort to Build Its Own Quantum Computer,” MIT Technology Review, September 3, 2014, <http://www.technologyreview.com/news/530516/google-launches-effort-to-build-its-own-quantum-computer/>.

Conclusion

SEO is both art and science. The artistic aspect of SEO requires dynamic creativity and intuition; the search engine algorithms are too complex to reverse-engineer every aspect of them. The scientific aspect involves challenging assumptions, analyzing data, testing hypotheses, making observations, drawing conclusions, and achieving reproducible results. These two ways of thinking will remain a requirement as SEO evolves into the future.

In this chapter, we conveyed some sense of what is coming in the world of technology, and in particular, search. Although the previous decade has seen an enormous amount of change, the reality is that it has simply been the tip of the iceberg. There's a lot more change to come, and at an ever-increasing (exponential) rate. If the Law of Accelerating Returns holds, we're in for a wild ride.

In this fast-moving industry, the successful SEO professional has to play the role of early adopter. The early adopter is always trying new things—tools, tactics, approaches, processes, technologies—to keep pace with the ever-evolving search engines, ever-increasing content types, and the ongoing evolution of online user engagement.

It is not enough to adapt to change. You will need to embrace it and evangelize it. Many in your (or your client's) organization may fear change, and steering them through these turbulent waters will require strong leadership. Thus, the successful SEO professional also has to play the role of change agent.

The need for organizations to capture search mindshare, find new customers, and promote their messaging will not diminish anytime soon, and neither will the need for searchable, web-based, and instantaneous access to information, products, and services. This ability—to generate traction and facilitate growth by connecting the seeker and provider—is perhaps the most valuable skill set on the Web today. And although there is the possibility that the search engines could eventually be called “decision,” “dilemma,” or even “desire” engines, the absolute need for the understanding of and interactions between both the psychological and the technological natures of search will ensure that SEO as a discipline, and SEO professionals, are here to stay.