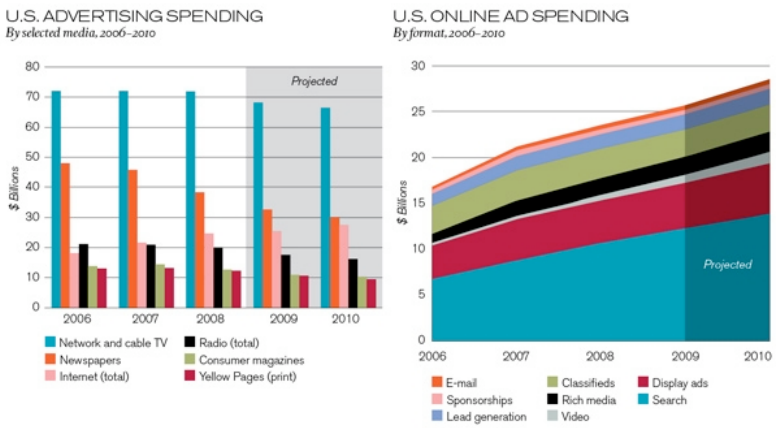
Google: Search, Online Advertising, and Beyond…

First draft version last modified: June 7, 2009 – comments welcome – john.gallaugher@bc.edu
Note: this is an earlier version of the chapter. All chapters updated after July 2009 are now hosted (and still free) at http://www.flatworldknowledge.com. For details see the 'Courseware' section of http://gallaugher.com

INTRODUCTION

Google has been called a one-trick pony[1], but as tricks go, it's got an exquisite one. Google's 'trick' is matchmaking – pairing Internet surfers with advertisers and taking a cut along the way. This cut is substantial – over $21 billion in 2008. In fact, as *Wired's* Steve Levy puts it, Google's matchmaking capabilities may represent 'the most successful business idea in history'[2]. For perspective, consider that as a ten-year-old firm, and one that had been public for less than five years, Google had already grown to earn more annual advertising dollars than *any* U.S. media company. No television network, magazine group, or newspaper chain brings in more ad bucks than Google. And none is more profitable. While Google's stated mission is "to organize the world's information and make it universally accessible and useful", advertising drives profits and lets the firm offer most of its services for free.

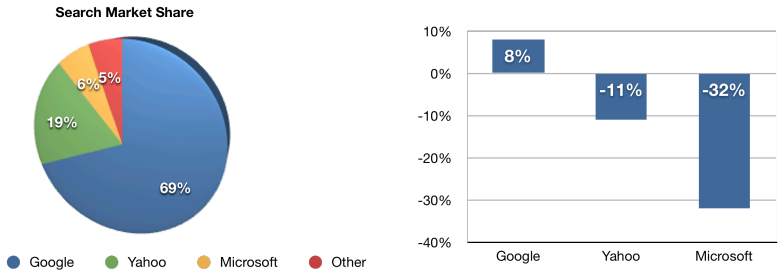


Online Advertising represents the only advertising category trending with positive growth. Search captures the most online ad dollars, and Google dominates search advertising[3]

As more people spend more time online, advertisers are shifting spending away from old channels to the Internet; and Google is swallowing the lion's share of this funds transfer[4]. By some estimates Google has 76 percent of the search advertising business[5]. Add to that Google's lucrative AdSense network that serves ads to sites ranging from small time bloggers to the *New York Times*, plus markets served by Google's acquisition of display ad leader DoubleClick, and

[1] Li 2009
[2] Levy 2009
[3] Chart needs to be re-created or permissions received for commercial product. Accessed from: http://www.technologyreview.com/business/22122/page3/
[4] Pontin 2009
[5] Sherman 2009

the firm controls in the neighborhood of 70 percent of *all* online advertising dollars[6]. Google has the world's strongest brand[7] (it's name is a verb – *just Google it*). It is regularly voted among the best firms to work for in America (twice topping *Fortune's* list). And its share of the search market has consistently grown while its next two biggest competitors have shrunk.



Search Market Share, Year End 2008, and Change in Market Share '07-'08 (Data Source: Hitwise)[8]

Wall Street has rewarded this success. The firm's *market capitalization* or *market cap* (the value of the firm calculated by multiplying it's share price by the number of shares) makes Google the most valuable media company on the planet. By early 2009, Google's market cap was greater than that of News Corp (which includes Fox, MySpace, the *Wall Street Journal*), Disney (including ABC, ESPN, theme parks, Pixar), Time Warner (*Fortune*, *Time*, *Sports Illustrated*, CNN, Warner Bros.), Viacom (MTV, VH1, Nickelodeon), CBS, and the New York Times – *combined*! Not bad for a business started by two twenty-something computer science graduate students. By 2007 that duo, Sergei Brin and Larry Page, were billionaires, tying for fifth on the Forbes 400 list of wealthiest Americans.

Genius Geeks & Plum Perks

Brin and Page have built a talent magnet. At the Googleplex, the firm's Mountain View, California headquarters, geeks are lavished with perks that include on-site laundry, massage, carwash, bicycle repair, free haircuts, state of the art gyms, and Wi-Fi equipped, biodiesel-powered shuttles that ferry employees between Silicon Valley and the San Francisco Bay-area. The Googleplex is also pretty green. The facility gets 30 percent of its energy from solar cells, representing the largest corporate installation of its kind at the time of deployment[9].

The firm's quirky tech-centric culture is evident everywhere. A T-Rex skeleton looms near the volleyball court. Hanging from the lobby ceiling is a replica of SpaceShipOne, the first commercial space vehicle. And visitors to the bathroom will find 'testing on the toilet', coding problems or other brainteasers to keep gray-matter humming while seated on one of the firm's $800 remote-controlled Japanese commodes. Staff also enjoy an A-list lecture series attracting luminaries ranging from celebrities to heads of state.

And of course there's the food – all of it free. The firm's founders felt that no employee should be more than 100 feet away from nourishment, and a tour around Google offices will find espresso bars, snack nooks, and fully stocked beverage refrigerators galore. There are 11 gourmet cafeterias on site, the most famous being "Charlie's Place", first run by the former executive chef for the Grateful Dead.

[6] Baker 2008
[7] Rao 2009
[8] Shankland 2009 – *Note to Editor: Gallaugher-created graphs
[9] Weldon, 2007

CEO Eric Schmidt says the goal of all this is "to strip away everything that gets in our employees' way"[10], and the perks, culture, and sense of mission have allowed the firm to assemble one of the most impressive rosters of technical talent anywhere. The Googleplex is like a well-fed Manhattan project, and employee ranks have included: 'Father of the Internet' Vint Cerf; Hal Varian, the former Dean of the U.C. Berkley School of Information Systems; Kai-Fu Lee, the former head of Microsoft Research in China; and Andy Hertzfeld, one of the developers of the original Macintosh user interface.

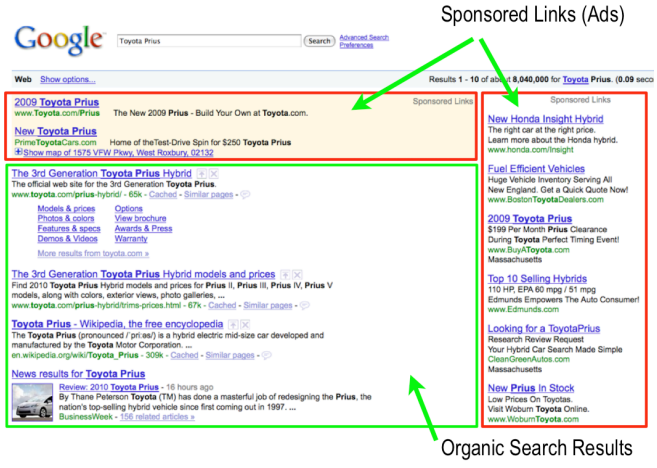
Engineers find Google a particularly attractive place to work, in part due to a corporate policy of offering "20 percent time", the ability work the equivalent of one day a week on new projects that interest them. It's a policy that has fueled innovation. Google Vice President Marissa Mayer (who herself regularly ranks among *Fortune's* most powerful women in business) has stated that roughly half of Google products got their start in 20 percent time[11].

Studying Google gives us an idea of how quickly technology-fueled market disruptions can happen, and how deeply these disruptions penetrate various industries. We'll also study the underlying technologies that power search, online advertising, and customer profiling. We'll explore issues of strategy, privacy, fraud, and discuss other opportunities and challenges the firm faces going forward.

UNDERSTANDING SEARCH

Before diving into how the firm makes money, let's first understand how Google's core service, search, works.

Perform a search (or *query*) on Google or another search engine, and the results you'll see are referred to by industry professionals as *organic*, or *natural* search. Search engines use different algorithms for determining the order of organic search results, but at Google the method is called *PageRank* (a bit of a play on words, it ranks web pages, and was initially developed by Google co-founder Larry Page). Google does not accept money for placement of websites in organic search results. Instead, PageRank results are a kind of popularity contest. Web pages that have more pages *linking to them* are ranked higher.



[10] Wolgemuth 2008
[11] Casnocha, 2009

The query for 'Toyota Prius' above triggers organic search results, flanked top and right by sponsored link advertisements

The process of improving a page's organic search results is often referred to as *Search Engine Optimization*, or *SEO*. SEO has become a critical function for many marketing organizations since if a firm's pages aren't near the top of search results, customers may never discover its site.

Google is a bit vague about the specifics of precisely how PageRank has been refined, in part because many have tried to game the system. The less scrupulous have tried creating a series of bogus web sites, all linking back to the pages they're trying to promote (this is called *link fraud*, or *spamdexing*, and Google actively works to uncover and shut down such efforts). We do know that links from some websites carry more weight than others. For example, links from websites that Google deems as 'influential', and links from most .edu web pages, have greater weight in PageRank calculations than links from run-of-the-mill .com sites.

Spiders and Bots and Crawlers – Oh My!

When performing a search via Google or another search engine, you're not actually searching the web. What really happens is that the major search engines make what amounts to a *copy* of web, storing and indexing the text of online documents on their own computers. Google's index considers over 1 trillion URLs.[12] The upper right-hand corner of a Google query shows you just how fast a search can take place (in the example above, rankings from over 8 million results containing the term "Toyota Prius" were delivered in less than a tenth of a second).

To create these massive indexes, search firms use software to crawl the web and uncover as much information as they can find. This software is referred to by several different names: *software robots*, *bots*, *spiders*, *web crawlers*. They all pretty much work the same way. In order to make its websites visible, every online firm provides a list of all of the public, named servers on its network (the *Telecommunications* chapter explains this as DNS or Domain Name Service listings). For example, Yahoo has different servers that can be found at www.yahoo.com, sports.yahoo.com, weather.yahoo.com, finance.yahoo.com, etc. Spiders start at the first page on every public server and follow every available link, traversing a web site until all pages are uncovered.

Google will crawl frequently updated sites, like those run by news organizations, as often as several times an hour. Rarely updated, less popular sites might only be re-indexed every few days. The method used to crawl the web also means that if a web page isn't the first page on a public server, or isn't linked to from another public page, then it'll never be found[13].

While search engines show you what they've found on their *copy* of web's contents, clicking a search result will direct you to the actual website, not the copy. But sometimes you'll click a result only to find that the web page doesn't match what the search engine found. This happens if a website was updated before your search engine had a chance to re-index the changes. In most cases you can still pull up the search engine's copy of the page. Just click the 'Cached' link below the result (the term *cache* refers to a temporary storage space used to speed computing tasks).

But what if you want the content on your website to remain off limits to search engine indexing and caching? Organizations have created a set of standards to stop the spider crawl, and all commercial

[12] Wright, 2009

[13] Most websites do have a link where you can submit a web page for indexing, and doing so can help promote the discovery of yoru content.

search engines have agreed to respect these standards. One way is to put a line of *HTML code* in the header of your web page to tell all software robots to either stop indexing a page, stop following links on the page, or stop offering old page archives in a cache. Users don't see this code, but commercial webcrawlers do. For those familiar with HTML code (the language used to describe a web page), the command to stop webcrawlers from indexing a page, following links, and listing archives of cached pages looks like this:

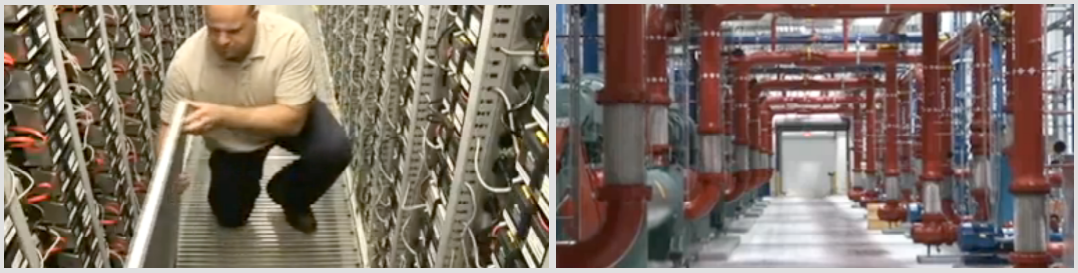
<META NAME="ROBOTS" CONTENT="NOINDEX, NOFOLLOW, NOARCHIVE">

There are other techniques to keep the spiders out, too. Website administrators can add a special file (called robots.txt) that provides similar instructions on how bots should treat the website. And a lot of content lies inside the '*dark web*', either behind corporate firewalls or inaccessible to those without a user account. Think private Facebook updates no one can see unless they're your friend – all of that is out of Google's reach.

What's It Take To Run This Thing?

Sergei Brin and Larry Page started Google with just four scavenged computers[14]. But in a decade, the infrastructure used to power the search sovereign has ballooned to the point where it is now the largest of its kind in the world[15]. Google doesn't disclose the number of servers it uses, but by some estimates, it runs over 1.4 million servers in over a dozen so-called *server farms* worldwide[16]. In 2008 the firm spent $2.18 billion on capital expenditures, with data centers, servers, and networking equipment eating up the bulk of this cost[17]. Building massive server farms to index the ever-growing web is now the cost-of-admission for any firm wanting to compete in the search market. This is clearly no longer a game for two graduate students working out of a garage.

The size of this investment not only creates a barrier to entry, it influences industry profitability, with market-leader Google enjoying huge economies of scale. Firms may spend the same amount to build server farms, but if Google has two thirds of this market (and growing) while Microsoft's search draws only about 1/10th the traffic, which do you think enjoys the better return on investment?



A Look at Google's Farm. Left: Google technician amidst racks of servers packed inside a shipping container. Right: the web of piping required to connect and cool a server farm's computing equipment[18]

[14] Liedtke, 2008
[15] Carr, 2006
[16] Katz, 2009
[17] Google, 2009
[18] Images from: Google Container Data Center Tour http://www.youtube.com/watch?v=zRwPSFpLX8I&feature=player_embedded. *Note to Editor: need to verify we can use these. Google has made the video publicly available similar images used in CNet article below

The hardware components that power Google aren't particularly special. In most cases the firm uses the kind of Intel or AMD processors, low-end hard drives, and RAM chips that you'd find in a desktop PC. These components are housed in rack-mounted servers about 3.5 inches thick, with each server containing two processors, eight memory slots, and two hard drives.

In some cases, Google mounts racks of these servers inside standard-sized shipping containers, each with as many as 1,160 servers per box (see image above)[19]. A given data center may have dozens of these server-filled containers all linked together. Redundancy is the name of the game. Google assumes individual components will regularly fail, but no single failure should interrupt the firm's operations (making the setup what geeks call *fault tolerant*). If something breaks, a technician can easily swap it out with a replacement.

Each server farm layout has also been carefully designed with an emphasis on lowering power consumption and cooling requirements. And the firm's custom software (much of it built upon open-source products) allows all this equipment to operate as the world's largest grid computer.

Web search is a task particularly well suited for the massively parallel architecture used by Google and its rivals. For an analogy of how this works, imagine that working alone, you need try to find a particular phrase in a one hundred-page document (that's a one-server effort). Next, imagine that you can distribute the task across five thousand people, giving each of them a separate sentence to scan (that's the multi-server grid). This difference gives you a sense of how search firms use massive numbers of servers and the divide-and-conquer approach of grid computing to quickly find the needles you're searching for within the web's haystack (for more on grid computing, see *Moore's Law and More*, and for more on server farms, see *Software in Flux*).

The Google Search Appliance and the Google Mini are hardware products that firms can purchase in order to run Google search technology within the privacy and security of an organization's firewall[20]

Google will even sell you a bit of its technology so that you can run your own little Google in-house, without sharing documents with the rest of the world. Google's line of search appliances are rack-mounted servers that can index documents within a corporation's web site, even specifying password and security access on a per-document basis. Selling hardware isn't a large business for Google, and other vendors offer similar solutions, but search appliances can be vital tools for law firms, investment banks, and other document-rich organizations.

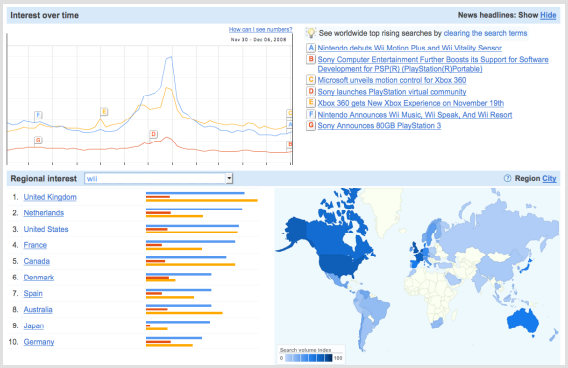
Trend-spotting with Google

Google not only gives you search results, it lets you see aggregate trends in what its users are searching for, and this can yield powerful insights. For example, by tracking search trends for flu symptoms, Google's *Flu Trends* website can pinpoint outbreaks one to two weeks faster than the Centers for Disease Control[21]. Want to go beyond the flu? Google's *Trends*, and *Insights for Search* services allow anyone to explore search trends, breaking out the analysis by region, category (image, news, product), date, and other criteria. Savvy managers can leverage these and similar tools for competitive analysis, comparing a firm, its brands, and its rivals.

[19] Shankland, 2009

[20] *Note to Editor: these are images taken from Google's online website promoting the products. Need to verify we can use them.

[21] Bruce, 2009.



Google Insights for Search can be a useful tool for competitive analysis and trend discovery. The chart above shows a comparison (over a 12-month period, and geographically) of search interest in the terms Wii, Playstation, and XBox

UNDERSTANDING THE INCREASE IN ONLINE AD SPENDING

For several years, Internet advertising has been the only major media ad category to show significant growth. There are three reasons driving online ad growth trends: 1) increasing user time online, 2) improved measurement and accountability, and 3) targeting.

American teenagers (as well as the average UK, Australia, and New Zealand web surfer), now spend more time on the Internet than watching television[22]. They're reading fewer print publications, and radio listening among the iPod generation is down 30 percent[23]. So advertisers are simply following the market. Online channels also provide advertisers with a way to reach consumers at work – something that was previously much more difficult to do.

Many advertisers have also been frustrated by how difficult it's been to gauge the effectiveness of traditional ad channels such as TV, print, and radio. This is reflected in the old industry saying "I know that half of my advertising is working – I just don't know which half". Well, with the Internet, now you know. While measurement technologies aren't perfect, advertisers can now count ad *impressions* (each time an ad appears on a web page), whether or not a user clicks on an ad, and the product purchases or other website activity that comes from those click-throughs[24].

Various technologies and techniques also make it easier for firms to target users based on how likely a person is to respond to an ad. In theory a firm can use targeting to spend marketing dollars only on those users deemed to be it's best prospects. Let's look at a few of these approaches in action.

SEARCH ADVERTISING

[22] MediaWeek, 2008; Hendry, 2008; BigMouthMedia, 2007
[23] Tobias, 2009
[24] See Pontin, 2009 for a more detailed overview of the limitations in online ad measurement.

The practice of running and optimizing search-engine ad campaigns is referred to as *search engine marketing* (*SEM*)[25]. SEM is a hot topic in an increasingly influential field, so it's worth spending some time learning how search advertising works on the Internet's largest search engine.

Roughly two-thirds of Google's revenues come from ads served on its own sites, and the vast majority of this revenue comes from search engine ads[26]. During Google's early years, the firm actually resisted making money through ads. In fact, while at Stanford, Brin and Page even coauthored a paper titled "The Evils of Advertising."[27] But when Yahoo and others balked at buying Google's search technology (offered for as little as $500,000), Google needed to explore additional revenue streams. It wasn't until two years after incorporation that Google ran ads alongside organic search results. That first ad, one for "Live Mail Order Lobsters", appeared just minutes after the firm posted a link reading 'See Your Ad Here"[28]).

The ads you'll see to the right (and sometimes top) of Google's organic search results appear as *keyword advertising*, meaning they're targeted based on a user's query. Advertisers bid on the keywords and phrases that they'd like to use to trigger the display of their ad. Linking ads to search was a brilliant move, since the user's search term indicates an overt interest in a given topic. Want to sell hotel stays in Tahiti? Link your ads to the search term "Tahiti Vacation".

Not only are search ads highly targeted, advertisers only pay for results. Ads appearing on Google search pages are billed on a *pay-per-click* (or *PPC*) basis, meaning that advertisers don't spend a penny unless someone actually clicks on their ad (note the term PPC is sometimes used interchangeably with the term *CPC* for *cost-per-click*).

Not Entirely Google's Idea

Google didn't invent pay-for-performance search advertising. A firm named GoTo.com (later renamed Overture) pioneered pay-per-click ads and bidding systems, and held several key patents governing the technology. Overture provided pay-per-click ad services to both Yahoo and Microsoft, but it failed to refine and match the killer combination of ad auctions and search technology that made Google a star. Yahoo eventually bought Overture and sued Google for patent infringement. In 2004 the two firms settled, with Google giving Yahoo 2.7 million shares in exchange for a 'fully paid, perpetual license' to over 60 Overture patents[29].

If an advertiser wants to display an ad on Google search, they can set up a Google AdWords advertising account in minutes, specifying just a single ad, or multiple ad campaigns that trigger different ads for different keywords. Advertisers also specify what they're willing to pay each time an ad is clicked, how much their overall ad budget is, and they can control additional parameters, such as the timing and duration of an ad campaign.

[25] Elliott, 2006
[26] Google Earnings Summary, Q1 2009
[27] Vise, 2008
[28] Levy, 2009
[29] Olsen, 2004

If no one clicks on an ad, Google doesn't make money, advertisers don't attract customers, and searchers aren't seeing ads they're interested in. So in order to create a winning scenario for everyone, Google has developed a precise ad ranking formula that rewards top performing ads by considering two metrics: the maximum *cost-per-click* (or *CPC*) that an advertiser is willing to pay, and the advertisement's *quality score* – a broad measure of ad performance. Create high quality ads and your advertisements might appear ahead of competition, even if your competitors bid more than you. But if ads perform poorly they'll fall in rankings, or even drop from display consideration.

Ad Rank = Maximum CPC x Quality Score
Formula used by Google to determine the rank order of sponsored links appearing on search results pages

The factors that go into determining an ad's quality score include the *click-through-rate* (*CTR*) for the ad, the overall history of click performance for the keywords linked to the ad, the relevance of an ad's text to the user's query, and Google's automated assessment of the user experience on the *landing page* – the web page displayed when a user clicks on the ad. This means that ads that don't get many clicks, ad descriptions that have nothing to do with query terms, and ads that direct users to generic pages that load slowly or aren't strongly related to the keywords and descriptions used in an ad, will all lower an ad's chance of being displayed[30].

When an ad is clicked, advertisers don't actually pay their maximum CPC, Google discounts ads to just one cent more than the minimum necessary to maintain an ad's position on the page. This means if you bid $1.00 per click, but the ad ranked below you bids 90 cents, you'll pay just 91 cents if the ad is clicked. Discounting was a brilliant move. No one wants to get caught excessively overbidding rivals, so discounting helps reduce the possibility of this so-called bidder's remorse. And with this risk minimized, the system actually encouraged higher bids![31]

Ad ranking and cost per click calculations take place as part of an automated auction that occurs *every time* a user conducts a search. Advertisers get a running total of ad performance statistics so that they can monitor the return on their investment and tweak promotional efforts for better results. And this whole system is automated for self-service – all it takes is a credit card and an ad idea, and you're ready to go.

How Much Do Advertisers Pay Per Click?

Google rakes in billions on what amounts to pocket change earned one click at a time. Most clicks bring in between 30 cents and one dollar. However, costs can vary widely depending on industry, current competition, and perceived customer value. The table below shows some of the highest reported CPC rates, but remember, any values fluctuate in real-time based on auction participants.

[30] Google, 2007
[31] Levy, 2009

Business/Industry	Keywords in the Top 25	Avg. CPC
Structured Settlements	2	$51.97
Secured Loans	2	$50.67
Buying Endowments	1	$50.35
Mesothelioma Lawyers	5	$50.30
DUI Lawyers	4	$49.78
Conference Call Companies	1	$49.64
Car Insurance Quotes	3	$49.61
Student Loan Consolidation	3	$49.44
Data Recovery	2	$49.43
Remortgages	2	$49.42

10 Most Expensive Industries for Keyword Ads (Source: WebPageFX Weekly)[32]

Since rates are based on auctions, top rates reflect what the market is willing to bear. As an example, law firms, which bring in big bucks from legal fees, decisions, and settlement payments, often justify higher customer acquisition costs. And firms that see results will keep spending. L.A.-based Chase Law Group has said that it brings in roughly 60 percent of its clients through Internet advertising[33].

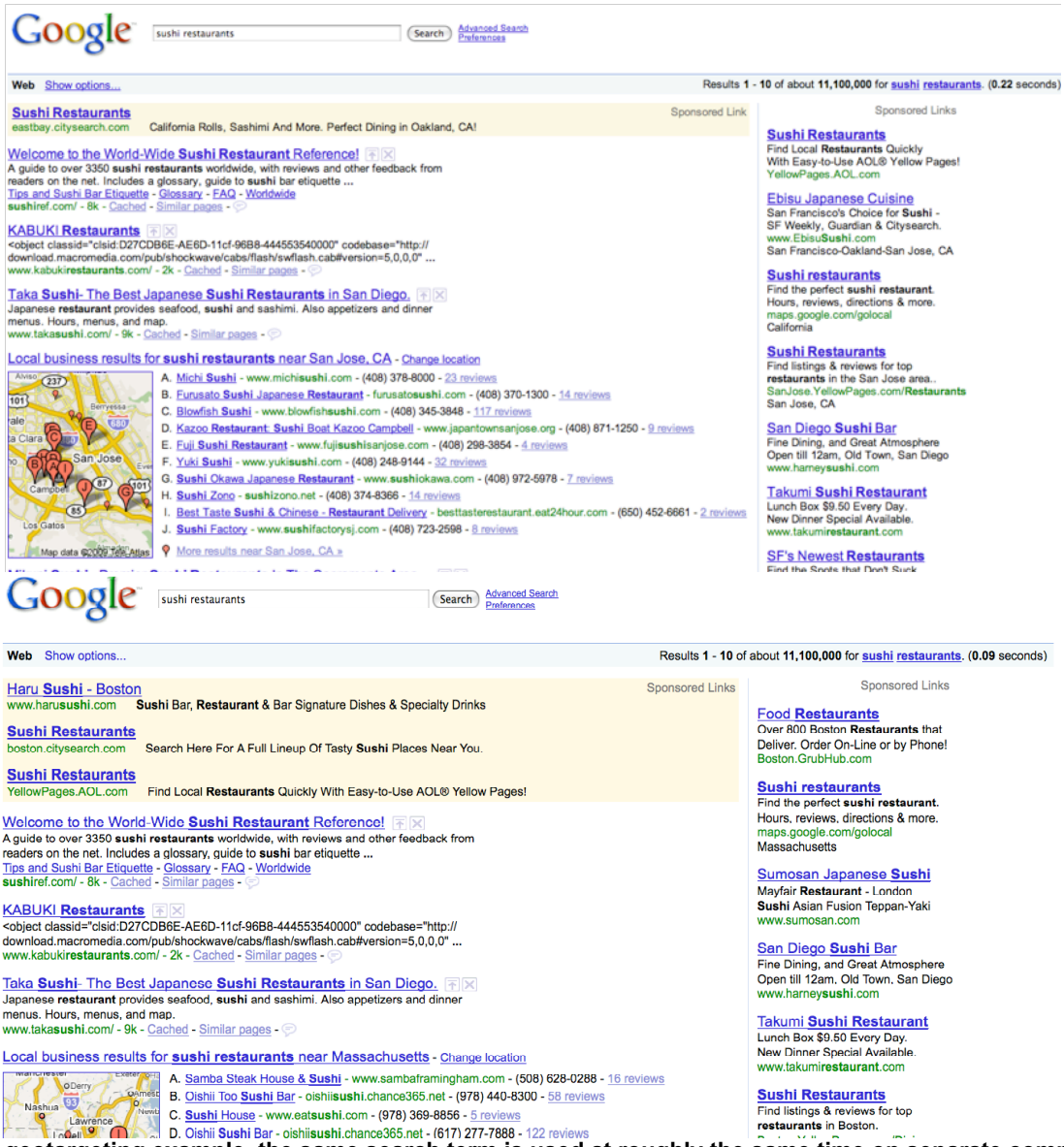
IP Addresses and Geotargeting

Geotargeting occurs when computer systems identify a user's physical location (sometimes called *geolocation*) for the purpose of delivering tailored ads or other content. On Google AdWords, for example, advertisers can specify that their ads only appear for web surfers located in a particular country, state, metropolitan region, or a given distance around a precise locale. They can even draw a custom ad-targeting region on a map, and tell Google to only show ads to users detected inside that space.

Ads in Google Search are geotargeted is based on *IP address*. Every device connected to the Internet has a unique IP address assigned by the organization connecting the device to the network. Normally you don't see your IP address (a set of four numbers, from 0 to 255, separated by periods, e.g. 136.167.2.220). But the range of IP addresses 'owned' by major organizations and Internet service providers (ISPs), is public knowledge. That means that in many cases it's possible to make an accurate guess as to where a computer, laptop, or mobile phone is located, simply by cross-referencing a device's current IP address with this public list.

For example, it's known that all devices connected to the Boston College network contain IP addresses starting with the numbers 136.167. If a search engine detects a query coming from an IP address that begins with those two numbers, it can be fairly certain that the person using that device is in the greater Boston area.

[32] Becket, 2009. Assembled from data at SpyFu.com.
[33] Mann, 2006.



In this geotargeting example, the same search term is used at roughly the same time on separate computers located in Silicon Valley area (top) and Boston (bottom). Note how geotargeting impacts results

IP addresses will change depending on how and where you connect to the Internet. Connect your laptop to a hotel's Wi-Fi when visiting a new city, and you're likely to see ads specific to that location. That's because your Internet service provider has changed, and the firm serving your ads has detected that you are using an IP address known to be associated with your new location.

Geotargeting via IP address is fairly accurate, but it's not perfect. For example, some Internet service providers may provide imprecise or inaccurate information on the location of their networks. Others might be so vague that it's difficult to make a best guess at the geography behind a set of numbers (values assigned by a multinational corporation with many locations, for example). And there are other ways locations are hidden, such as when Internet users connect to

proxy servers, third-party computers that pass traffic to and from a specific address without revealing the address of the connected users.

What's My IP Address?

While every operating system has a control panel or command that you can use to find your current IP address, there are also several websites that will quickly return this value (and a best guess at your current location). One such site is http://ip-adress.com (note the spelling has only one 'd'). Visit this or a similar site with a desktop, laptop, and mobile phone. Do the results differ and are they accurate? Why?

Geotargeting Evolves Beyound the IP Address

There are several other methods of geotargeting. Firms like Skyhook Wireless can identify a location based on its own map of Wi-Fi hotspots and nearby cell towers. Many mobile devices come equipped with *GPS* chips (identifying location via the *global positioning service* satellite network). And if a user provides location values such as a home address or zip code to a website, then that value might be stored and used again to make a future guess at a user's location.

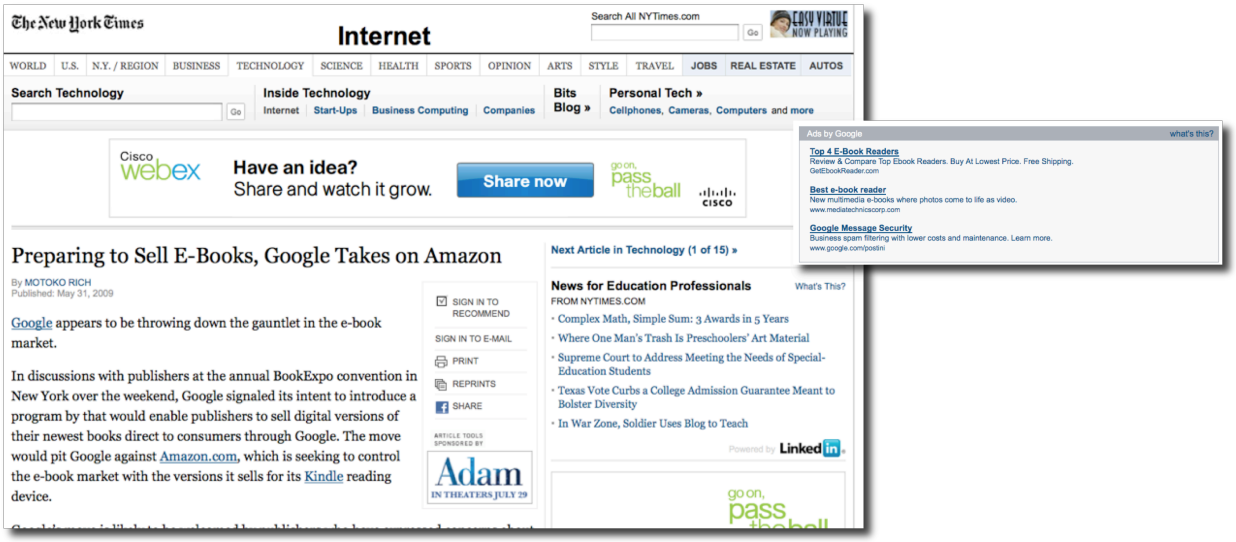
AD NETWORKS – DISTRIBUTION BEYOND SEARCH

Google runs ads not just in search, but also across a host of Google-owned sites like Gmail, Google News, and Blogger. It will even tailor ads for its map products and for mobile devices. But about 30 percent of Google's revenues come from running ads on websites that the firm doesn't even own.[34]

Next time you're surfing online, look around the different websites that you visit and see how many sport boxes labeled 'Ads by Google". Those websites are participating in Google's AdSense *ad network*, which means they're running ads for Google in exchange for a cut of the take. Participants range from small-time bloggers to some of the world's most highly trafficked sites. Google lines up the advertisers, provides the targeting technology, serves the ads, and handles advertiser payment collection. To participate, content providers just sign-up online, put a bit of Google-supplied HTML code on their pages, and wait for Google to send them cash (websites typically get about 70 to 80 cents for every AdSense dollar that Google collects[35]).

Google originally developed AdSense to target ads based on keywords automatically detected inside the content of a web page. A blog post on your favorite sports team, for example, might be accompanied by ads from ticket sellers or sports memorabilia vendors. AdSense and similar online ad networks provide advertisers with access to the long tail of niche websites; offering both increased opportunities for ad exposure, as well as more refined targeting opportunities.

[34] Google Earnings Results, Q1 2009
[35] Tedeschi, 2006.



The images above show advertising embedded around a story on the *New York Times* website. The page runs several ads provided by different ad networks. For example, the WebEx banner ad above the article's headline was served by AOL-owned Platform-A/Tacoda. The "Ads by Google" box appeared at the end of the article. Note how the Google ads are related to the content of the Times article

Running ads on your website is by no means a guaranteed path to profits. The Internet graveyard is full of firms that thought they'd be able to sustain their businesses on ads alone. But for many websites, ad networks can be like oxygen, sustaining them with revenue opportunities they'd never be able to achieve on their own.

For example, AdSense provided early revenue for the popular social news site Digg, as well as the multi-million dollar TechCrunch media empire. It supports Disaboom, a site run by physician and quadriplegic Dr. Glen House. And it continues to be the primary revenue generator for AskTheBuilder.com. That site's founder, former builder Tim Carter, had been writing a handyman's column syndicated to some 30 newspapers. The newspaper columns didn't bring in enough to pay the bills, but with AdSense he hit pay dirt, pulling in over $350,000 in ad revenue in just his first year![36]

[36] Rothenberg, 2008.

Tim Carter's 'Ask the Builder' website runs ads from Google and other ad networks. Note different ad formats surrounding the content. There's even an ad in the bottom of the video, served from Google-owned YouTube

Beware the Content Adjacency Problem

Contextual advertising based on keywords is lucrative, but like all technology solutions, it has its limitations. Vendors sometimes suffer from *content adjacency problems* when ads appear alongside text they'd prefer to avoid. In one particularly embarrassing example, a *New York Post* article detailed a gruesome murder where hacked up body parts were stowed in suitcases. The online version of the article included contextual advertising and was accompanied by… luggage ads.[37]

To combat embarrassment, ad networks provide opportunities for both advertisers and content providers to screen out potentially undesirable pairings based on factors like vendor, website, keyword, and category. Ad networks also refine ad-placement software based on feedback from prior incidents (for more on content adjacency problems, see the *Facebook case*).

Google launched AdSense in 2003, but Google is by no means the only company to run an ad network, nor was it the first to come up with the idea. Rivals include the Yahoo Publisher Network, Microsoft's adCenter, and AOL's Platform-A. Others, like Quigo don't even have a consumer website, yet manage to consolidate enough advertisers to attract high-traffic content providers such as ESPN, Forbes, Fox, and USA Today. Advertisers also aren't limited to choosing just one ad network. In fact, many content provider websites will serve ads from several ad networks (as well as exclusive space sold by their own sales force), oftentimes mixing several different offerings on the same page.

Ad Networks and Competitive Advantage

While advertisers can use multiple ad networks, there are several key strategic factors driving the industry. For Google, its ad network is a distribution play. The ability to reach more potential

[37] Overholt, 2007

customers across more websites attracts more advertisers to Google. And content providers (the websites that distribute these ads) want there to be as many advertisers as possible in the ad networks that they join, since this should increase the price of advertising, the number of ads served, and the accuracy of user targeting. If advertisers attract content providers, which in turn attract more advertisers, then we've just described network effects! More participants bringing in more revenue also help the firm benefit from scale economies - offering a better return-on-investment from its ad technology and infrastructure. No wonder Google's been on such a tear – the firm's loaded with assets for competitive advantage!

Google's Ad Reach Gets Bigger

While Google has the largest network specializing in distributing text ads, it had been a laggard in graphical display ads (sometimes called image ads). That changed in 2008, with the firm's $3.1 billion acquisition of display ad network and targeting company, DoubleClick. Now in terms of the number of users reached, Google controls both the largest text ad network, and the largest display ad network[38].

MORE AD FORMATS AND PAYMENT SCHEMES

Online ads aren't just about text ads billed in CPC. Ads running through Google AdSense, through its DoubleClick subsidiary, or on most competitor networks can be displayed in several formats and media types, and can be billed in different ways. The specific ad formats supported depend on the ad network, but can include: *image* (or *display*) *ads* (such as horizontally-oriented *banners*, smaller rectangular *buttons*, and vertically-oriented *skyscraper ads*), *rich-media ads* (which can include animation or video), *interstitials* (ads that run before a user arrives at a web-page's contents), and more. The industry trade group, the *Internet Advertising Bureau* (or *IAB*) sets common standards for display ads so that a single *creative* (the design and content of the advertisement) can run unmodified across multiple ad networks and websites[39].

And there are lots of other ways ads are sold, besides cost-per-click. Most graphical display ads are sold according to the number of times the ad appears (known as an *impression*, in industry-speak). Ad rates are quoted in *CPM*, meaning cost per thousand impressions (the M representing the roman numerical for one thousand). Display ads sold on a CPM basis, are often used as part of branding campaigns targeted more at creating awareness than generating click-throughs. Such techniques often work best for promoting products like soft drinks, toothpaste, or movies.

Cost-per-action (or CPA) ads pay whenever a user performs a specified action such as signing up for a service, requesting material, or making a purchase. *Affiliate programs* are a form of cost-per-action, where websites share a percentage of revenue earned when a user clicks-through and buys something from their website. Amazon.com runs the world's largest affiliate program, offering referring sites from 4 to 15 percent of sales generated from a click-through. Purists might not consider affiliate programs as advertising (rather than text or banner ads, Amazon's affiliates offer links and product descriptions that point back to Amazon's website), but these programs can be important tools in a firm's promotional arsenal.

[38] Baker, 2009.
[39] See IAB Ad Unit Guidelines for details

And rather than buying targeted ads, a firm might sometimes opt to become an exclusive advertiser on a site. A firm could, for example: buy access to all ads served a site's main page; it could secure exclusive access to a region of the page (such as the topmost banner ad); or it may pay to sponsor a particular portion or activity on a website (say a parenting forum, or a 'click-to-print' button). Such deals can be billed based on a flat rate, CPM, CPC, or any combination of metrics.

Some firms sell sponsorship opportunities. In the example on the left, the movie 500 Days of Summer sponsored the Printer-Friendly formatting of New York Times articles. The example on the right shows Abbott Laboratories' sponsorship of a Crohn's Heath Center on HealthCentral.com

Ads in Games?

As consumers spend more time in video games, it's only natural that these products become ad channels, too. Finding a sensitive mix that introduces ads without eroding the game experience can be a challenge. Advertising can work in racing or other sports games (in 2008, the Obama campaign famously ran virtual billboards in EA's *Burnout Paradise*), but ads make less sense for games set in the past, future, or on other worlds. Branding ads often work best, while click-throughs are typically not something you want disrupting your gaming experience.

Advertisers have also explored sponsorships of web-based and mobile games, while others have tried a sort of virtual product placement integrated into experiences. A version of the Sims, for example, included virtual replicas of real-world inventory from IKEA and H&M.

Obama Campaign's virtual billboard in EA's Burnout Paradise[40]

In-game ad-serving technology also lacks the widely accepted standards of web-based ads, so it's unlikely that ads designed for a Wii sports game could translate into a PS3 first-person shooter. Also, one of the largest in-game ad networks, Massive, is owned by Microsoft. Good if you want to run ads on XBox, but not exactly the firm that Nintendo or Sony want to play nice with. In-game advertising shows promise, but the medium is considerably more complicated than conventional website ads. That lowers relative ROI and will likely continue to constrain growth.

CUSTOMER PROFILING AND BEHAVIORAL TARGETING

Advertisers are willing to pay more for ads that have a greater chance of reaching their target audience, and online firms have a number of targeting tools at their disposal. Much of this

[40] Image Source: SocialMedia8 Slideshare Presentation: http://www.slideshare.net/socialmedia8/case-study-the-barack-obama-strategy

targeting occurs because whenever you visit a website, a behind-the-scenes software dialog takes place between web browser and web server that can reveal a number of pieces of information, including IP address, the type of browser used, the computer type, its operating system, and unique identifiers, called *cookies*.

And remember, *any* server that serves you content can leverage these profiling technologies. This means that you might be profiled not just by the website that you're visiting (e.g. nytimes.com), but also by any ad networks that serve ads on that site (e.g. Platform-A, DoubleClick, Google AdSense, Microsoft adCenter, etc.).

An IP address not only helps with geolocation, it can also indicate a browser's employer or university, which can be further matched with information such as firm size or industry. IBM has used IP targeting to tailor its college recruiting banner ads to specific schools (example: "There Is Life After Boston College, Click Here to See Why"). This targeting boasted click-throughs from 5 to 30 percent[41], vs. current rates well below 1 percent for most untargeted banner ads. DoubleClick once even served a banner that include a personal message for an executive at then-client Modem Media. The ad, reading "Congratulations on the twins, John Nardone", was served across hundreds of sites, but was only visible from computers on the Modem Media corporate network[42].

The ability to identify a surfer's computer, browser, or operating system can be used to target tech ads. For example, Google might pitch its Chrome browser to users detected running Internet Explorer, Firefox, or Safari; while Apple could target those "I'm a Mac" ads just to Windows users.

But perhaps the greatest degree of personalization and targeting comes from cookies. Visit a website for the first time, and in most cases, a behind-the-scenes dialog takes place that goes something like this:

> Server: *Have I seen you before?*
> Browser: *No.*
> Server: *Then take this unique string of numbers and letters (called a cookie). I'll use it to recognize you from now on.*

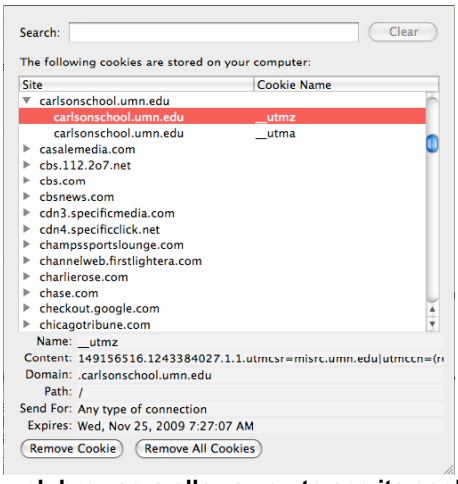
The cookie is just a line of identifying text, assigned and retrieved by a given web server, and stored in your browser. Upon accepting this cookie, your browser has been tagged, like an animal. As you surf around the firm's website, that cookie can be used to build a profile associated with your activities. If you're on a portal, like Yahoo, you might type in your zip code, enter stocks that you'd like to track, and identify the sports teams you'd like to see scores for. The next time you return to the website, your browser responds to the server's '*Have I see you before?*' question with the equivalent of '*Yes, you know me*', and it presents the cookie that the site gave you earlier. The site can then match this cookie against your browsing profile, showing you the weather, stock quotes, sports scores, and other info that it thinks you're interested in.

---

[41] Moss, 1999.
[42] Moss,1999

Cookies are used for lots of purposes. Retail websites like Amazon use cookies to pay attention to what you've shopped for and bought, tailoring web pages to display products that the firm suspects you'll be most interested in. Sites also use cookies to keep track of what you put in an online 'shopping cart', so if you quit browsing before making a purchase, these items will reappear the next time you visit. And many websites also use cookies as part of a 'remember me' feature, storing user IDs and passwords. Beware this last one! If you check the 'remember me' box on a public web browser, the next person who uses that browser is potentially using *your* cookie, and can log in as you!

An organization can't read cookies that it did not give you. So businessweek.com can't tell if you've also got cookies from forbes.com. But you can see all of the cookies in your browser. Take a look and you'll almost certainly see cookies from dozens of websites that you've never visited before. These are *third-party cookies* (sometimes called *tracking cookies*), and they are usually served by ad networks or other customer profiling firms.



The Preferences setting in most web browsers allows you to see its cookies. This browser has received cookies from several ad networks, media sites, and the U. Minnesota Carlson School of Management

By serving cookies in ads shown on partner sites, ad networks can build detailed browsing profiles that include sites visited, specific pages viewed, duration of visit, and the types of ads you've seen and responded to. And that surfing might give an advertising network a better guess at demographics like gender, age, marital status, and more. Visit a new parent site and expect to see diaper ads in the future, even when you're surfing for news or sports scores!

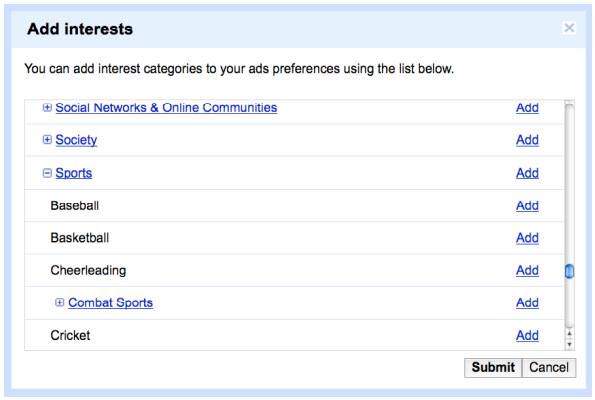
But What If I Don't Want a Cookie!

If all of this creeps you out, remember that you're in control. The most popular web browsers allow you to block all cookies, block just third party cookies, purge your cookie file, or even ask for your approval before accepting a cookie. Of course, if you block cookies, you block any benefits that come along with them, and some website features may require cookies to work properly. Also note that while deleting a cookie breaks a link between your browser and that website, if you supply identifying information in the future (say by logging into an old profile), the site might be able to assign your old profile data to the new cookie.

While the Internet offers targeting technologies that go way beyond traditional television, print, and radio offerings, none of these techniques is perfect. Since users are regularly assigned different IP addresses as they connect and disconnect from various physical and WiFi networks, IP targeting can't reliably identify individual users. Cookies also have their weaknesses. They're assigned by browsers, so if several people use the same browser, all of their web surfing activity may be mixed into the same cookie profile. Some users might also use different browsers on the same machine, or use different computers. Unless a firm has a way to match up these different cookies with user accounts or other user-identifying information, a site may be working with multiple, incomplete profiles.

PROFILING AND PRIVACY

Until 2009, Google hadn't used tracking cookies on its AdSense network. While AdSense has been wildly successful, contextual advertising has its limits. For example, what kind of useful targeting can firms really do based on the text of a news item on North Korean nuclear testing?[43] So in March 2009, the firm announced what it calls 'interest-based ads'. Google AdSense would now issue a third-party cookie and would track browsing activity across AdSense partner sites, and Google-owned YouTube. AdSense would build a profile, initially identifying users within 30 broad categories and 600 subcategories. Says one Google project manager; "We're looking to make ads even more interesting"[44].



Interest Categories for Google Interest-Based Advertising

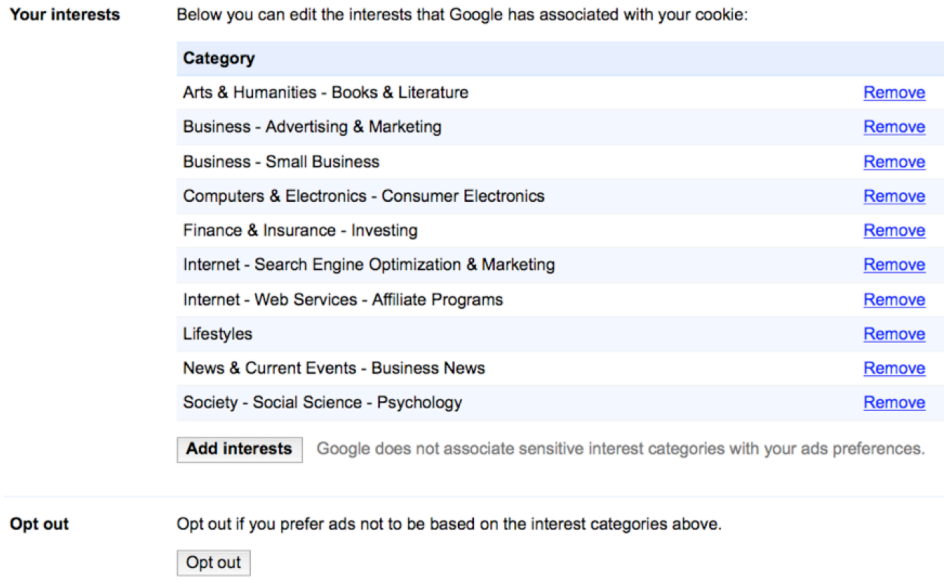
Of course, there's a financial incentive to do this, too. Ads deemed more interesting should garner more clicks, meaning more potential customer leads for advertisers, more revenue for websites that run AdSense, and more money for Google.

But while targeting can benefit web surfers, users will resist if they feel that they are being mistreated, exploited, or put at risk. Negative backlash might also result in a change in legislation. The U.S. Federal Trade Commission has already called for more transparency and user control in online advertising, and for requesting user consent (*opt-in*) when collecting sensitive data[45]. Mishandled user privacy could curtail targeting opportunities, limiting growth

[43] Singel, 2009
[44] Hof, 2009
[45] Singel, 2009

across the online advertising field. And with less ad support, many of the Net's free services could be suffer.

Google's roll-out of interest-based ads shows the firm's sensitivity to these issues. First, while major rivals have all linked query history to ad targeting, Google steadfastly refuses to do this. Other sites often link registration data (including user-submitted demographics such as gender and age) with tracking cookies, but Google avoids this practice, as well.



An example of one user's interests, as tracked by Google's "Interest-based Ads", and displayed in the firm's "Ad Preferences Manager"

Google has also placed significant control in the hands of users, with options at program launch that were notably more robust than those of its competitors[46]. Each 'interest-based ad' is accompanied by an "Ads by Google" link that will bring users to a page describing Google advertising, and which provides access to the company's "Ads Preferences Manager". This tool allows surfers to see any of the 30 broad categories and 600 subcategories that Google has assigned to that browser's tracking cookie. Users can remove categorizations, and even add interests if they want to improve ad targeting. The technology also avoids targeting certain sensitive topics, such as race, religion, sexual orientation, health, political or trade union affiliation, and certain financial categories[47].

Google also allows users to install a cookie that opts them out of interest-based tracking. And since browser cookies can expire or be deleted, the firm has gone a step further, offering a browser plug-in that will remain permanent, even if a user's opt-out cookie is purged.

Google, Privacy Advocates, and the Law

Google's moves are meant to demonstrate transparency in its ad targeting technology, and the firm's policies may help raise the collective privacy bar for the industry. While privacy advocates have praised

[46] Hansell, 2009
[47] Mitchell, 2009

Google's efforts to put more control in the hands of users, many continue to voice concern over what they see as the increasing amount of information that the firm houses[48]. For an avid user, Google could conceivably be holding e-mail (Gmail), photos (Picasa), a web surfing profile (AdSense and DoubleClick), medical records (Google Health), location (Google Latitude), appointments (Google Calendar), transcripts of phone messages (Google Voice), work files (Google Docs), and more.

Google insists and that reports portraying it as a data-hording Big Brother are inaccurate. The firm is adamant that user data exists in silos that aren't federated in any way, nor are employees permitted access to multiple data archives without extensive clearance and monitoring. Data is not sold to third parties. Activities in Gmail, docs, or most other services isn't added to targeting profiles. And any targeting is fully disclosed, with users empowered to opt-out at all levels[49]. But critics counter that corporate intensions and data use policies (articulated in a website's *Terms of Service*) can change over time, and that a firm's good behavior today is no guarantee of good behavior in the future[50].

Google does enjoy a lot of user goodwill, and it is widely recognized for its unofficial motto "Don't Be Evil". However some worry that even though Google might not be evil, it could still make a mistake, and that despite its best intensions, a security breach or employee error could leave data dangerously or embarrassingly exposed.

The wary point to AOL's release of search history on over 650,000 of its web searchers. These log files included queries such as: "how to tell your family you're a victim of incest", "surgical help for depression", "can you adopt after a suicide attempt", "gynecology oncologists in new york city", "how long will the swelling last after my tummy tuck", and perhaps most damning, queries that included specific names, addresses, and phone numbers. While AOL offered the data in a way that disguised individual user accounts, in many cases aggregate query detail contained terms so specific, they provided a strong indication of who conducted the searches[51].

While Google has never experienced a breach of that magnitude, it has suffered minor incidents, including a March 2009 gaffe in which the firm inadvertently shared some Google Docs with contacts who were never granted access to them[52].

Privacy advocates also worry that the amount of data stored by Google serves as one-stop shopping for litigators and government investigators. The counter argument points to the fact that Google has continually reflected an aggressive defense of data privacy in court cases. When Viacom sued Google over copyright violations in YouTube, the search giant successfully fought the original subpoena, which had requested user-identifying information[53]. And Google was the only one of the four largest search engines to resist a 2006 Justice Department subpoena for search queries[54].

Google is increasingly finding itself in precedent-setting cases where the law is vague. Google's 'Street View', for example, has been the target of legal action in the U.S., Canada, Japan, Greece, and the U.K. Varying legal environments create a challenge to the global rollout of any data-driven initiative[55].

---

[48] Helft, 2009
[49] Mitchell, 2009
[50] Mitchell, 2009
[51] Kawamoto and Mills, 2006
[52] Kincaid, 2009
[53] Mitchell, 2009
[54] Broche, 2006
[55] Sumagaysay, 2009

Ad targeting brings to a head issues of opportunity, privacy, security, risk, and legislation. Google is now taking a more active public relations and lobbying role to prevent misperceptions and to be sure its positions are understood. While the field continues to evolve, Google's experience will lay the groundwork for the future of personalized technology and provide a case study for other firms that need to strike the right balance between utility and privacy. Despite differences, it seems clear to Google, its advocates, and its detractors, that with great power comes great responsibility.

SEARCH ENGINES, AD NETWORKS, AND FRAUD

There's a lot of money to be made online, and this has drawn the attention of criminals and the nefarious. Online fraudsters may attempt to steal from advertisers, harm rivals, or otherwise dishonestly game the system. But bad guys beware - such attempts violate terms-of-service agreements and may lead to prosecution and jail time.

Studying ad-related fraud helps marketers, managers, and technologists understand potential vulnerabilities, as well as the methods used to combat them. This also builds tech-centric critical thinking, valuation, and risk assessment skills.

Some of the more common types of fraud that are attempted in online advertising include:

- *Enriching click fraud* – when site operators generate bogus ad clicks to earn PPC income.
- *Enriching impression fraud* – when site operators generate false page views (and hence ad impressions) in order to boost their site's CPM earnings.
- *Depleting click fraud* – clicking a rival's ads to exhaust their PPC advertising budget.
- *Depleting impression fraud* – generating bogus impressions to exhaust a rival's CPM ad budget.
- *Rank-based impression fraud* - on CPC ads where ad rank is based on click performance, fraudsters repeatedly search keywords linked to rival ads or access pages where rival ads appear. The goal is to generate impressions without clicks. This lowers the performance rank (quality score) of a rival's ads, possibly dropping ads from rank results, and allowing fraudsters to subsequently bid less for the advertising slots previously occupied by rivals.
- *Disbarring fraud* – attempting to frame a rival by generating bogus clicks or impressions that appear to be associated with the rival, in hopes that this rival will be banned from an ad network or punished in search engine listings.
- *Link fraud* or *spamdexing* - creating a series of bogus web sites, all linking back to a page, in hopes of increasing that page's results in organic search.
- *Keyword stuffing* – packing a web page with unrelated keywords (sometimes hidden in fonts that are the same color as a web page's background) in hopes of either luring users who wouldn't normally visit a web page, or attracting higher-value contextual ads.

Disturbing stuff, but firms have put their best geeks on the case. Widespread fraud would tank advertiser ROI and crater the online advertising market, so Google and rivals are diligently working to uncover and prosecute the crooks.

Busting the Bad Guys

On the surface, enriching click fraud seems the easiest to exploit. Just set up a web page and click like crazy. Each click should ring the ad network cash register, and a portion of those funds will be passed on to you – *ka ching*! But remember, each visitor is identified by an IP address, so lots of clicks from a single IP makes the bad guys easy to spot.

So organized crime tried to up the game, running so-called *click farms* to spread fraud across dozens of IP addresses. *The Times of India* uncovered one such effort where Indian housewives were receiving up to 25 cents for each ad click made on fraudster-run websites[56]. But an unusually large number of clicks from Indian IP addresses foiled these schemes, as well.

Fraudsters then moved on to use *zombie networks* - hordes of surreptitiously infiltrated computers, linked and controlled by rogue software[57]. To create zombie networks (sometimes called *bot nets*), hackers exploit security holes, spread viruses, or trick users into installing software that will lie dormant, awaiting commands from a central location. The controlling machine then sends out tasks for each zombie, instructing them to visit websites and click on ads in a way that mimics real traffic. Zombie bot nets can be massive. Dutch authorities once took down a gang that controlled some 1.5 million machines[58].

Scary, but this is where scale, expertise, and experience come in. The more activity an ad network can monitor, the greater the chance that it can uncover patterns that are anomalous. Higher click-through rates than comparable sites? Caught. Too many visits to a new or obscure site? Caught. Clicks that don't fit standard surfing patterns for geography, time, and day? Caught.

Sometimes the goal isn't theft, but sabotage. Google's Ad Traffic Quality Team backtracked through unusual patterns to uncover a protest effort targeted at Japanese credit card firms. Ad clicks were eventually traced to an incendiary blogger who incited readers to search for the Japanese word *kiyashinku* (meaning cashing credit, or credit-cards), and to click the credit card firm ads that show up, depleting firm search marketing budgets. Sneaky, but uncovered and shut down, without harm to the advertisers[59].

Search firm and ad network software can use data patterns and other signals to ferret out most other types of fraud, too, including rank-based impression fraud, spamdexing, and keyword stuffing. While many have tried to up the stakes with increasingly sophisticated attacks, large ad networks have worked to match them, increasing their anomaly detection capabilities across all types of fraud[60].

Click Fraud: How Bad Is It?

[56] Vidyasagar, 2004.
[57] Mann, 2006.
[58] Sanders, 2007. Members of Google's security team published an autopsy of one 100,000+ network. For technical details, see Daswani and Stoppleman, 2007.
[59] Jakobsson and Ramzan 2008.
[60] Jakobsson and Ramzan 2008.

Accounts on the actual rate of click fraud vary widely. Some third-party firms contend that nearly one in five clicks is fraudulent[61]. But Google adamantly disputes these headline-grabbing numbers, claiming that many such reports are based on site logs that reflect false data from conditions that Google doesn't charge for (e.g. double counting a double click, or adding up repeated use of the browser back button in a way that looks like multiple clicks have occurred). The firm also offers monitoring, analytics, and reporting tools that can uncover this kind of misperceived discrepancy.

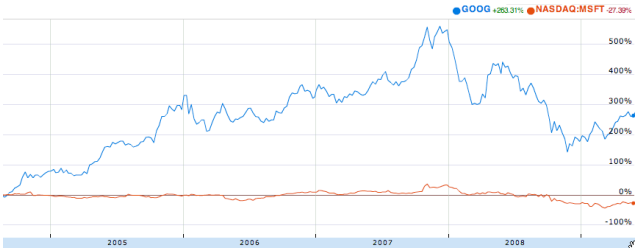
Google contends that all invalid clicks (mistakes and fraud) represent less than 10 percent of all clicks, that the vast majority of these clicks are filtered out, and that Google doesn't charge advertisers for clicks flagged as mistakes or suspicious[62]. In fact, Google says their screening bar is so high and so accurate that less than .02% of clicks are reactively classified as invalid and credited back to advertisers[63].

So who's right? While it's impossible to identify the intention behind every click, the market ultimately pays for performance. And advertisers are continuing to flock to PPC ad networks (and to Google in particular). While that doesn't mean that firms can stop being vigilant, it does suggest that for most firms, Google seems to have the problem under control.

THE BATTLE UNFOLDS

Google has been growing like gangbusters, but the firm's twin-engines of revenue growth, ads served on search and through its ad networks, will inevitably mature. And it will likely be difficult for Google to find new growth markets that are as lucrative as these. Emerging advertising outlets such as social networks and mobile have lower click-through rates than conventional advertising, suggesting that Google will have to work harder for less money.

For a look at what can happen when maturity hits, check out Microsoft. The House that Gates Built is more profitable than Google, and continues to dominate the incredibly lucrative markets served by Windows and Office. But these markets haven't grown much for over a decade. In developing nations, most Windows and Office purchases come not from growth, but when existing users upgrade or buy new machines. And without substantial year-on-year growth, the stock price doesn't move.



A comparison of roughly five years of stock price change – Google (GOOG) vs. Microsoft (MSFT)

For big firms like Microsoft and Google, pushing stock price north requires not just new markets, but *billion dollar* ones. Adding even $100 million in new revenues doesn't do much for firms bringing in $21 billion and $51 billion a year, respectively. That's why you see Microsoft

[61] Hamner, 2009.
[62] Lafsky, 2008.
[63] Jakobsson and Ramzan 2008.

swinging for the fences, investing in the uncertain, but potentially gargantuan markets of video games, mobile phone software, cloud computing (see *Software in Flux* chapter), music and video, and of course, search and everything else that fuels online ad revenue.

Search: Google Rules, but It Ain't Over

PageRank is by no means the last word in search, and offerings from Google and its rivals continue to evolve. Google supplements PageRank results with news, photos, video, and other categorizations (click the 'Show options…' link above your next Google search). Yahoo is continually refining its search algorithms and presentation (click the little down arrow at the top of the firm's search results for additional categorizations and suggestions). And Microsoft's third entry into the search market, Bing, sports nifty tweaks for specific kinds of queries. Restaurant searches in Bing are bundled with ratings stars, product searches show up with reviews and price comparisons, and airline flight searches not only list flight schedules and fares, but also a projection on whether those fares are likely go up or down. Bing also comes with a $100 million marketing budget, showing that Microsoft is serious about moving its search market share out of the single digits.

New tools like the Wolfram Alpha 'knowledge engine' (and to a lesser extent, Google Squared) move beyond web page rankings and instead aggregate data for comparison, formatting findings in tables and graphs. Web sites are also starting to wrap data in invisible tags that can be recognized by search engines, analysis tools, and other services. If a search engine can tell that a number on a restaurant's website is, for example, either a street address, an average entrée price, or the seating capacity, it will be much easier for computer programs to accurately categorize, compare, and present this information. This is what geeks are talking about when they refer to the *semantic web*. All signs point to an increasingly more useful Internet!

Both Google and Microsoft are on a collision course. But there's also an impressive roster of additional firms circling this space, each with the potential to be competitors, collaborators, merger partners, or all of the above. Yahoo, while wounded and shrinking, is still a powerhouse, ranking ahead of Google in some overall traffic statistics. Add in eBay, Facebook, Twitter, Amazon, SalesForce.com, Apple, Netflix, the video game industry, telecom and mobile carriers, cable firms, and the major media companies, and the next few years have the makings of a big, brutal fight.

Strategic Issues

Google's scale advantages in search, and its network effects advantages in advertising, were outlined earlier. The firm also leads in search/ad experience and expertise, and continues to offer a network reach that's unmatched. But the strength of Google's other competitive resources is less clear.

Within Google's ad network, there are *switching costs* for advertisers and for content providers. Google partners have set up accounts and are familiar with the firm's tools and analytics. Content providers would also need to modify web pages to replace AdSense or DoubleClick ads with rivals. But choosing Google doesn't cut out the competition. Many advertisers and content providers participate in multiple ad networks, making it easier to shift business from one firm to another. That likely means that Google will have to retain its partners by offering superior value.

Another vulnerability may exist with search consumers. While Google's brand is strong, switching costs for search users are incredibly low. Move from Google.com to Ask.com and you actually save three letters of typing!

Still, there are no signs that Google's search leadership is in jeopardy. So far users have been creatures of habit, returning to Google despite heavy marketing by rivals. And in Google's first decade, no rival has offered technology compelling enough to woo away the googling masses – the firm's share has only increased. Defeating Google with some sort of technical advantage will be difficult, since web-based innovation can often be quickly imitated.

The Google Toolbar helps reinforce search habits among those who have it installed, and Google pays the Mozilla foundation (the folks behind the Firefox browser) upwards of $66 million a year to serve as its default search option[64]. But Google's track record in expanding reach through distribution deals is mixed. The firm spent nearly $1 billion to have MySpace run AdSense ads, but Google has publicly stated that social network advertising has not been as lucrative as it had hoped (see the *Facebook Case*). The firm had also spent nearly $1 billion to have Dell pre-install its computers with the Google browser toolbar and Google desktop search products. But in 2009, Microsoft inked deals that displaced Google on Dell machines, and it also edged Google out in a 5-year search contract with Verizon Wireless[65].

How Big is Too Big?

Microsoft could benefit from embedding its Bing search engine into its most popular products (imagine putting Bing in the right-mouseclick menu alongside *cut*, *copy*, and *paste*). But with Internet Explorer market share above 65 percent, Office above 80 percent, and Windows at roughly 90 percent[66], this seems unlikely.

European anti-trust officials have already taken action against Redmond's bundling Windows Media Player and Internet Explorer with Windows. Add in a less favorable anti-trust climate in the U.S., and tying any of these products to Bing is almost certainly out of bounds. What's not clear is whether or not regulators would allow Bing to be bundled with less dominant Microsoft offerings, such as mobile phone software, XBox, and MSN.

But increasingly, Google is also an anti-trust target. Microsoft has itself raised anti-trust concerns against Google, unsuccessfully lobbying both U.S. and European authorities to block the firm's acquisition of DoubleClick.[67] Google was forced to abandoned a Fall 2008 search advertising partnership with Yahoo after the Justice Department indicated its intention to block the agreement. The Justice Department is also investigating a Google settlement with the Authors' Guild, a deal in which critics have suggested that Google scored a near monopoly on certain book scanning, searching, and data serving rights[68]. And yet another probe is investigating whether Google colluded with Apple, Yahoo, and other firms to limit efforts to hire away top talent[69].

[64] Shankland, 2008.
[65] Wingfield, 2009.
[66] Data source: Hitslink.com, May 2009; and Montalbano, 2009
[67] Broach, 2007; and Kawamoto and Broach, 2007.
[68] Wildstrom, 2009.
[69] Buskirk, 2009.

Of course, being big isn't enough to violate U.S. anti-trust law. Harvard Law's Andrew Gavil says: "You've got to be big, and you have to be bad […] You have to be both."[70] This may be a difficult case to make against a firm that has a history of being a relentless supporter of open computing standards. And as mentioned earlier, there is little forcing users to stick with Google – the firm must continue to win this market on its own merits. While Google may escape all of these investigations, increased anti-trust scrutiny is a downside that comes along with the advantages of market-dominating scale.

More Ads, More Places, More Formats

Google has been a champion of increased Internet access. But altruism aside, more net access also means a greater likelihood of ad revenue.

Google's effort to catalyze Internet use worldwide comes through on multiple fronts. In the US, Google has supported (with varying degrees of success) efforts to offer free WiFi in San Francisco and Mountain View. But most ambitiously, Google is also a major backer (along with Liberty Global and HSBC) of the O3b satellite network. O3b stands for 'the other 3 billion' of the world's population who currently lack Internet access. O3b hopes to have 16 satellites circling the globe by late 2010, blanketing underserved regions with *low latency* (low delay), high-speed Internet access[71]. With Moore's Law dropping computing costs as world income levels rise, Google hopes to empower the currently disenfranchised masses to start surfing. Good for global economies, good for living standards, and good for Google.

Another way Google can lower the cost of surfing is by giving mobile phone software away for free. That's the thinking behind the firm's Android offering. With Android, Google provides mobile phone vendors with a Linux-based operating system, supporting tools, standards, and an application marketplace akin to Apple's AppStore. Android itself isn't ad-supported – there aren't Google ads embedded in the OS. But the hope is that if handset manufacturers don't have to write their own software, the cost of wireless mobile devices will go down. And cheaper devices mean that more users will have access to the mobile Internet, adding more ad-serving opportunities for Google and its partner sites.

While Android started as a mobile phone operating system and software stack, its use has expanded, with it now serving as a Windows-alternative on low-cost, Internet-equipped laptops (netbooks). And a tailored Android could conceivably power net access in a host of devices, including televisions, set top boxes, and automobiles. Google has dabbled in selling ads for television (as well as radio and print). There may be considerable potential in bringing variants of ad targeting technology, search, and a host of other services across these devices.

Google has also successfully lobbied the U.S. government to force wireless telecom carriers to be more open, dismantling what are known in the industry as *walled gardens*. Before Google's lobbying efforts, mobile carriers could act as gatekeepers, screening out hardware providers and software services from their networks. Now, paying customers of carriers that operate over recently allocated U.S. wireless spectrum will have access to a choice of hardware, and less

[70] Lohr and Helft, 2009.
[71] Malik, 2008.

restrictive access to websites and services. And Google hopes this expands its ability to compete without obstruction.

YouTube

Then there's Internet video, with Google in the lead here, too. It's tough to imagine any peer-produced video site displacing YouTube. Users go to YouTube because there's more content, while amateur content providers go there seeking more users (classic two-sided network effects). This critical advantage was the main reason why in 2006, Google paid $1.65 billion for what was then just a 20 month old startup.

That popularity comes at a price. Even with falling bandwidth and storage costs, at 13 hours of video uploaded to YouTube *every minute*, the cost to store and serve this content is cripplingly large[72]. Credit Suisse estimates that in 2009, YouTube will bring in roughly $240 million in ad revenue, pitted against $711 million in operating expenses. That's a shortfall of more than $470 million. Analysts estimate that for YouTube to break even, it would need to achieve a ad CPM of $9.48 on each of the roughly 75 billion streams it'll serve up this year. A tough task. For comparison, Hulu (a site that specializes in offering ad-supported streams of television shows and movies) earns CPM rates of $30 and shares about 70 percent of this with copyright holders. Most user-generated content sports CPM rates south of a buck[73].

Adding to costs: video uploading is set to explode as more cell phones become net-equipped video cameras. Adding to revenue potential: viewing will also skyrocket as mobile devices and television sets ship with YouTube access. The firm is still experimenting with ad models – these include traditional banner and text ads, plus ads transparently layered across the bottom 20 percent of the screen, *pre-roll* commercials that appear before the selected video, and more. Google has both the money and time to invest in nurturing this market, and it continues to be hesitant in saturating the media with ads that may annoy users and constrain adoption.

Apps and Innovation

In 2007 the firm announced a tagline to sum up its intensions: "search, ads, and apps". Google iss king of the first two, but this last item hasn't matured to the point where it impacts the firm's financials.

Experimentation and innovation are deeply ingrained in Google's tech-centric culture, and this has led to a flood of product offerings. Google released more than 360 products in 2008, and another 120 in Q1 2009[74]. It's also cancelled several along the way, including Jaiku (which couldn't beat Twitter), Google Video (which was superseded by the YouTube acquisition), and a bunch more you've likely not heard of, like Dodgeball, Notebook, Catalog Search, and Mashup Editor[75].

---

[72] Nakashima, 2008
[73] Wayne, 2009.
[74] Shiels, 2009.
[75] Needleman, 2009.

What's Google Up To?

With all this innovation, it's tough to stay current with Google's cutting edge product portfolio. But the company does offer 'beta' releases or projects, and invites the public to try out and comment on its many experiments. To see some of these efforts in action, visit Google Labs at http://googlelabs.com. And to see a current list of more mature offerings, check out http://www.google.com/options/

Google's "Apps" are mostly web-based software-as-a-service offerings. Apps include an Office-style suite that sports a word processor, presentation tool, and spreadsheet, all served through a browser. While initially clunky, the products are constantly being refined. The spreadsheet product, for example, has been seeing new releases every two weeks, with features such as graphing and pivot tables inching it closer in capabilities to desktop alternatives[76]. And new browser standards, such as HTML 5, will make it even easier for what lives in the browser to mimic what you're currently using on your desktop.

Google also offers Gears, a product that allows compatible apps to be used offline when net access isn't available. That'll be critical as long as Internet access is less reliable than your hard drive, but online collaboration is where these products can really excel (no pun intended). Most Google apps allow not only group viewing, but also collaborative editing, common storage, and version control. Google's collaboration push also includes its wiki-like Google Sites tool, and a new platform called Wave, billed as a sort of next-step evolving beyond e-mail and instant messaging.

Unknown is how much money Google will make money off all of this. Consumers and small businesses have free access to these products, with usage for up to 50 users funded by in-app ads. But is there much of a market serving ads to people working on spreadsheets? Enterprises can gain additional, ad-free licenses for a fee. While users have been reluctant to give up Microsoft Office, many have individually migrated to Google's web-based e-mail and calendar tools. Google's enterprise apps group will now do the same thing for organizations, acting as a sort of outsourcer by running e-mail, calendar, and other services for a firm; all while handling upgrades, spam screening, virus protection, backup, and other administrative burdens. Arizona State University, biotech giant Genentech, and auto parts firm Valeo are among the Google partners that have signed on to make the firm's app offerings available to thousands[77].

It's not until considered in its entirety that one gets a sense for what Google has the potential to achieve. It's possible that increasing numbers of users worldwide will adopt light, cheap netbooks and other devices powered by free Google software (the Android OS, Google's Chrome browser). Productivity apps, e-mail, calendaring, and collaboration tools will all exist 'in the cloud', accessible through any browser, with files stored on Google's servers in a way that minimizes hard drive needs. Google will entertain you, help you find the information you need, help you shop, handle payment (Google Checkout), and more. And the firms you engage online may increasingly turn to Google to replace their existing hardware and software infrastructure with corporate computing platforms like Google Apps Engine (issues associated

[76] Girouard, 2009.
[77] Coughlin 2007; Hardy, 2008; and Claburn 2009

with Software as a Service and Cloud Computing models are covered in the *Software in Flux* chapter). All of this would be based on open standards, but switching costs, scale, and increasing returns from expertise could yield enormous advantages.

• • • • •

Studying Google allowed us to learn about search and the infrastructure that powers this critical technology. We've studied the business of ads, covering search advertising, ad networks, and ad targeting in a way that blends strategic and technology issues. And we've covered the ethical, legal, growth, and competitive challenges that Google and its rivals face. Studying Google in this context should not only help you understand what's happening today, it should also help you develop critical thinking skills for assessing the opportunities and threats that will emerge across industries as technologies continue to evolve.

• • • • •

About This Work

The goal of this project is to have an impact. At my university, we've bucked the national trend, tripling our Information Systems majors in the three years since we've adopted a business-focused IS teaching approach, and I'm delighted to share this content with you. I hope that Flatworld's free online copies and low-cost print versions encourage wide adoption of this material, and I hope that you and your students enjoy it. Please tell others, and thanks!

Comments & feedback are most welcome! Contact Info:

- E-Mail: john.gallaugher@bc.edu
- Draft Chapters, Cases, Slides, and Podcasts: http://gallaugher.com/chapters
- For updates, supporting articles, and commentary, sign up for the Week in Geek at: http://www.gallaugher.com
- And follow on Twitter at http://twitter.com/gallaugher

About the Author

John Gallaugher is a member of the Dept. of Information Systems in Boston College's Carroll School of Management. Prof. Gallaugher teaches courses and conducts research at the intersection of technology and strategy. An award-winning teacher, he leads the School's TechTrek programs, co-leads the Asian field study program, and has consulted to and taught executive seminars for several organizations including Accenture, Alcoa, Brattle Group, ING Group, Patni Computer Systems, Staples, State Street, and the U.S. Information Agency.

This reading is available to faculty for non-commercial use. Enjoy! More chapters and cases will follow in Professor Gallaugher's forthcoming book "Information Systems: A Manager's Guide to Harnessing Technology", to be published both free online and low-cost (less than $30) print version, by Flat World Knowledge (FlatWorldKnowledge.com).

REFERENCES

Attributor, "Get Your Fair Share of the Ad Network Pie", March 30, 2008.

Baker, L., "Google Now Controls 69% of Online Advertising Market", *Search Engine Journal*, March 31, 2008.

Becket, X., "10 Businesses with the Highest Cost Per Click", *WebPageFX Weekly*, Feb. 20, 2009

BigMouthMedia, "Brits spend more time online than watching TV", July 12, 2007.

Blodgett, H., "Google Sucks Life Out of Old Media: Check Out The 2007 Share Shift", *Silicon Alley Insider*, March 14, 2008.

Broache, A., "Judge: Google must give feds limited access to records", *CNet*, March 17, 2006.

Broach, A., "On Capitol Hill, Google and Microsoft spar over DoubleClick", *CNet*, Sept. 27, 2007.

Bruce, S., "Google says user data aids flu detection", *eHealthInsider*, May 25, 2009.

Bruene, J., "Advertising on the Net", *NetBanker*, June 9, 1997.

Buskirk, E., "Antitrust Probe to Review Hiring Practices at Apple, Google, Yahoo: Report", *Wired News*, June 3, 2009.

Casnocha, B., "Success on the Side", *The American: The Journal of the American Enterprise Institute*, April 24, 2009.

Carr, David F., "How Google Works", *Baseline*, July 6, 2006.

Claburn, T., "Google's Cloud Evangelism Converts Enterprise Customers", *InformationWeek*, May 13, 2009.

Coughlin, S., "Google's e-mail for universities ", *BBC News*, June 11, 2007.

Daswani, N., and Stoppleman, M., "The Anatomy of Clickbot.A", *Proceedings of the First conference on First Workshop on Hot Topics in Understanding Botnets*, Cambridge, MA, April 11-13, 2007.

Elliott, S., "More Agencies Investing in Marketing With a Click". *The New York Times*. March 14, 2006.

Gharachorloo, K., "Click Fraud: Anecdotes from the Front Line", Google Presentation, Sept. 14, 2007.

Girouard, D., “Google Inc. Presentation”, *Bank of America Technology Conference*, June 4, 2009.

Google, *Marketing and Advertising Using Google: Targeting Your Advertising to the Right Audience*, First Edition, 2007

Google (Press Release), “Google Announces Fourth Quarter And Fiscal Year 2008 Results”, Jan. 22, 2009.

Gross, D., “Birth of a Salesman”, *Fortune*, Aug. 8, 2005.

Hamner, S., “Pay-Per-Click Advertisers Combat Costly Fraud”, *The New York Times*, May 12, 2009.

Hardy, Q., “Google Muscles Further Into Business Software”, *Forbes*, Feb. 28, 2008.

Hendry, A, “Connected Aussies Spend More Time Online Than Watching TV”, Computerworld Australia, May 21, 2008.

Hof, R., “Behavioral Targeting: Google Pulls Out the Stops”, *BusinessWeek*, March 11, 2009.

Jakobsson, M., and Ramzan, Z., *Crimeware: Understanding New Attacks and Defenses*, Symantec Press, Cupertino, CA, 2008.

Katz, R., “Tech Titans Building Boom”, *IEEE Spectrum*, Vol. 46, No. 2, Feb. 1, 2009.

Kawamoto, D., and Mills, E., “AOL Apologizes for Release of User Search Data”, *CNet*, Aug. 7, 2006.

Kawamoto, D., and Broach, A., “EU extends review of Google-DoubleClick merger”, *CNet*, Nov. 13, 2007.

Kincaid, J., “Google Privacy Blunder Shares Your Docs Without Permission”, *TechCrunch*, March 7, 2009.

Lafsky, M., “Google and Click Fraud: Behind the Numbers”, *The New York Times*, Feb. 27, 2008.

Levy, S., “The Secrets of Googlenomics”, Wired, June 2009.

Li, C., “Why Google's One-Trick Pony Struggles to Learn New Tricks”, *Harvard Business Publishing*, May, 2009

Liedtke, M., “Google Reigns as World’s Most Powerful 10 Year Old”, *Associated Press*, Sept. 5, 2008.

Lohr, S., and Helft, M., "New Mood in Antitrust May Target Google", *The New York Times*, May 18, 2009.

Malik, O., "Google Invests in Satellite Broadband Startup", Sept. 9, 2008.

Mann, C., "How Click Fraud Could Swallow the Internet", *Wired*, Jan. 2006.

MediaWeek, "American Teenagers Spend More Time Online Than Watching Television", June 19, 2008

Mitchell, R., "What Google Knows About You", *Computerworld*, May 11, 2009.

Montalbano, E., "Forrester: Microsoft Office in no danger from competitors", *InfoWorld*, June 4, 2009.

Moss, M, "These Web sites know who you are", *ZDNet UK*, Oct. 13, 1999.

Nakashima, E., "YouTube Ordered To Release User Data", *Washington Post*, July 4, 2008

Navneet, "Google: Over 1 Million Servers and Counting", *PageTraffic Blog*, July 3, 2007.

Needleman, R., "Google Killing Jaiku, Dodgeball, Notebook, Other Projects", *CNet*, Jan. 14, 2009.

Olsen, S., "Google, Yahoo bury the legal hatchet", *CNet*, Aug. 9, 2004.

O'Reilly, T., "Google's Rich Snippets and the Semanatic Web", *O'Reilly Radar*, May 14, 2009.

Overholt, A., "Search for Tomorrow", *FastCompany*, Dec. 19, 2007.

Pontin, J., "But Who's Counting?", *Technology Review*, March/April, 2009

Rao, L., "Guess Which Brand is Now Worth $100 Billion?" *TechCrunch*, April 30, 2009.

Rothenberg, R., "The Internet Runs on Ad Billions", *BusinessWeek*, April 10, 2008.

Sanders, T., "Dutch Botnet Gang Facing Jail", *IT News Australia*, Jan. 18, 2007.

Shankland, S., "Thanks, Google: Mozilla Revenue Hits $75 Million", *CNet*, Nov. 19, 2008.

Shankland, S., "Google Conquers 2008 Search Market in U.S.", *CNet*, Jan. 14, 2009.

Shankland, S., "Google Unlocks Once-Secret Server", *CNet*, April 1, 2009.

Shiels, M., Google unveils 'smarter search', *BBC News*, May 13, 2009.

Sherman, C., “Report: Google Leads U.S. Search Advertising Market With 76% Market Share”, *Search Engine Land*, Jan. 20, 2009.

Singel, R., “Online Behavioral Targeting Targeted by Feds, Critics”, *Wired News*, June 3, 2009.

Sumagaysay, L., “Not everyone likes the (Google Street) View”, *Good Morning Silicon Valley*, May 20, 2009.

Tedeschi, B., “Google’s Shadow Payroll Is Not Such a Secret Anymore”, *The New York Times*, Jan. 16, 2006.

Tobias, M., “Newspapers under siege”, *Philstar*, May 18, 2009

Vascellaro, J., “Gap Widens in Online Advertising”, *Wall Street Journal*, Sept. 4, 2008.

Vidyasagar, N., “India's secret army of online ad 'clickers’”, *The Times of India*, May 3, 2004.

Vise, D., “Google’s Decade”, *Technology Review*, Sept. 12, 2008.

Wayne, B., “YouTube is Doomed”, Silicon Alley Insider, April 9, 2009.

Weldon, D., “Google’s Power Play”, *EnergyDigital*, Aug. 30, 2007.

Wildstrom, S., “Google Book Search and the Dog in the Manger”, *BusinessWeek*, April 18, 2009.

Wingfield, N., “Microsoft Wins Key Search Deals”, The Wall Street Journal, Jan. 8, 2009.

Wolgemuth, L., “Forget the Recession, I want a Better Chair”, *US News & World Report*, April 28, 2008

Wright, A., “Exploring a ‘Deep Web’ That Google Can’t Grasp”, *The New York Times*, Feb. 23, 2009.