

## CONTENTS

- Why does the richest information resource known to man return such crap?
- An infinite number of monkeys? Picking on Demand Media.
- What does Quora, the new online sensation, say about having quality content?
- Google has lived by the sword of advertising. Will they die by it?
- Note to capitalists: This is a very good time to start a new search engine company.

### WHY ARE WE HERE?

After 15,000 hours of hands-on research it's clear that most of the easily accessed Internet content, particularly what is returned from Google search results, is crap. Although generally acknowledged by people in the hard-core knowledge worker segment, this viewpoint is now becoming mainstream.<sup>1</sup>

So how did things get this bad? While advertising may be the "root of all evil," how deep it goes depends on the conditions. The number and size of the forces at work here are prodigious. Here are some of the more technical and content-related forces:

"**Search Engine Optimization,**" or SEO, takes advantage of the algorithms Google uses to game the system. By relying on computers to rank search results, we invite the endless contest between those that find ways to outsmart the algorithms and the search engines trying to return valuable information. Google sometimes make changes to their algorithms in an attempt to foil current strategies but at the same time opens up other loopholes for exploiters to use. Another negative side effect is that these changes sometimes upset search rankings that were previously good in what amounts to a big reshuffling. Google has much better algorithms for returning search results

but we are unlikely to see them implemented—more on that later.

**Dumb Clicks:** Generally, the more clicks something gets the higher it ranks. It doesn't matter if the clicks were of the "fooled you, there's no real content here" variety. (See the attached example of a common top-ranked page.) Because people tend to click on something near the top of the results this crap keeps getting clicks. Even a senior search engineer at Google observing user behavior said: "How can you not see that this is a spam page and click on it?!" Time only makes this pattern worse as this type of content "crowds out" everything else.

**Tricksters & Cheap Shots:** One example of this category is alternativeto.net. In this case, someone must have noticed a popular search technique of using "alternative to" as in "alternative to Photoshop." Suddenly most queries put in this way returned crappy results full of ads rather than useful information.<sup>2</sup> There is a legion of mostly small-time operators that use domain misspellings and other simple ploys to capture traffic and a few clicks on "parked" domains. This class of problem is fairly easy to solve but still pops up from time to time again like a disease you can't quite eradicate.

**Content Farms:** This is where players like Associated Content (acquired by Yahoo), About.com and

Content  
Reduced to  
Advertising &  
Promotion

1

<http://www.codinghorror.com/blog/2011/01/trouble-in-the-house-of-google.html>

<sup>2</sup> Google has recently figured out how to suppress this site more effectively but it took months.

Demand Media take the game to a new level. Content farms are harder to counter because they invest some money in “real” content to insert themselves into search results. We’ll save the analysis for the following section on Demand Media. Content farms may seem beneficial compared to the SEO villains and tricksters but that’s what makes them insidious.

**Syndication:** Many sites are so desperate for content they are willing to syndicate just about anything. Because they are often following some of the same SEO strategies and now provide even more links to the original content source, the ranking and ubiquity of a lousy piece of content solidifies like a plaque to block normal information flow.

**Filter Failure:** “Leakage” is when something that is supposed to protect us from garbage content begins to break down. Even paid-for services like CapitalIQ rely on automated filtering and eventually get penetrated by spam content that corrupts the feed. This is just another factor suggesting that effective filters in the future will require some level of human validation even if it also is automated. Some of those methods are described in the final section.

Two other more anthropological factors play a major role in the drive to lower quality content:

**Recency bias:** Good content has long been pushed out of focus by inferior versions because they are newer. This is true in long-standing categories like movies and books. While freshness has a value it ends up being counterproductive in many cases and results in a “reinventing of the wheel” in the case of factual, well reasoned and documented content. Good content is often not “sticky,” fades away quickly, and becomes hard if not impossible to find. Wikipedia is one example where this is not the case which is why so many people search it explicitly.

**All clicks are not created equally:** There’s a strong inverse relationship between intelligence / expertise / judgment / insight and the tendency to click. On Google “a click is a click,” so content Crapification

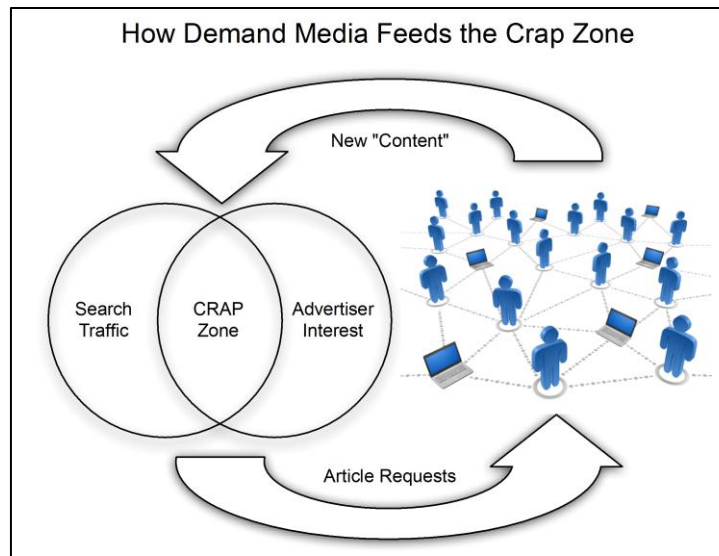
gets driven to the lowest common denominator. Steve Jobs probably makes fewer, more informed clicks and online decisions than Paris Hilton—but Paris and her ilk are what drive content rankings.

Solving some of these problems is actually easier than it would first appear, because harnessing the power of human behavior and adding more available information to the analysis can lead to excellent results. First we’ll look at the gorilla of the content farms.

## DEMAND MEDIA

The launch of the Demand Media IPO provides a detailed look at the massive machine whose purpose is to make money by serving content with a high affinity for advertisers and clicking consumers. First they figure out what people are searching for, intersect that with what advertisers are interested in, and then use a freelance team of thousands of content creators to write articles that will rank highly in search results and draw clicks on paid advertising links.

Although the company takes pains to demonstrate that they do, in fact, produce some decent content that provides complete and unique answers to some queries (like “how to draw a basket of fruit”), the vast majority is similar to what is shown in Exhibit A. The content is basically a page with a blurb of text surrounded by layers of advertising.



Fundamentally,

Demand Media is an upgraded, well-capitalized and professionally managed version of other content farms. They are not trying to fool anyone with their statement: “Our Mission is to fulfill the world’s demand for **commercially valuable** content” (emphasis added). Demand Media brings more technology and resources to this enterprise than the myriad “small time spammers” out there but, in the end, they exist only to push out content that gets clicks. The content is cursory, sometimes cut and pasted from elsewhere, and all the time buried

in piles of advertising. It's certainly not the rich, high quality content users want.

Demand Media is a giant sensing and production engine. They collect and analyze a huge amount of data and then use their production engine, DemandStudio, to publish 4,000 "articles" per day from over 10,000 writers. The company lists articles like "how to become a stockbroker" or "how to look sexy in a bathing suit" that can be written in exchange for a \$15 fee. Combine inane questions with uninformed articles by the masses and you have a recipe for a pile of insipid web pages. The last step insert ads on top, both sides, the bottom, and weaves them into the text to complete the exercise.

Of course, an elitist view of content can lead to missing some great investments. YouTube had plenty of garbage on it to start with. Many blogs and instant messaging traffic streams can be just as content-light as Demand Media sites. There are also companies like I Can Has Cheezburger, which has been extremely successful by dealing unabashedly with sites that provide silly content. Their sites comprise a vast network of distracting and useless sites like LOLcats. There's real money to be made online by giving people fun, distracting or entertaining content.

Unlike these other companies, however, Demand Media is not trying to entertain but instead to inform and do so with the sole intent of driving advertising clicks. This creates what can only be called spam.

Uninformed search algorithms like Google's will return enough of these results to fill the first page or two of a screen. Most of the time it's not obvious to users that the results are spam until they click on them. Google sees this as "Another happy user clicks on this link. We're the best!"

Whether one is happy about what Demand Media is doing is beside the point from an investment perspective. Similar arguments were made concerning QuinStreet (QNST - \$23) after their IPO, but in our view a year ago the shares were undervalued at \$13.<sup>3</sup> Demand Media has an attractive business even though we think innovations in search technologies may make parts of it harder to monetize.

---

<sup>3</sup> <http://blog.research2zero.com/2010/02/quinstreet-is-getting-more-interesting/>

Our Intrinsic Valuation model (see attachment) suggests a \$30 stock price.

QuinStreet took time to prove their model to investors. After a solid IPO early in 2010, the shares reached \$17 only to fall back to \$10 by the summer. Subsequently, however, the company has executed on their strategy and now the stock is more in line with their IV at \$23. Last year we saw similar stock dynamics for Ancestry.com and Higher One Holdings.

In order to shed more light on what will help improve information content on the web, it's worth looking at another player in the online content space that has been turning heads within the "tech-noscenti" and is beginning to enter the mainstream.

## DO YOU QUORA?

A new kid on the block has captured the attention of alpha geeks and early adopters and has certainly entered the hype cycle as the "next big thing." Quora starts with a focus on questions and answers, adds to it elements of social networking (like Twitter) plus a crowd-sourced ranking mechanisms found in Slashdot and made popular by Digg. This sounds a little esoteric at first but it ends up being effective at turning online information and the energies and knowledge of the many into productive results.

The crux of Quora is that it allows users to determine what questions to ask, answer them, and then provokes the community to rank the best ones so they are at the top. In this way they combine elements of Wikipedia with Digg; however, the ability to follow topics and have networks of followers adds another dimension.

There are a few reasons Quora is proving more useful than other online networks, so they are worth a closer look:

**Topical focus:** Quora uses questions and answers to provide a framework for information sharing. Users can submit new questions as well as answers to existing ones. There is an approval filter on questions to guard against multiple instances of the same question and "junk" questions.

**Notion of authority:** Quora has been blessed with a very high level of initial user. For example, answers are from individuals who are clear authorities. It's not unusual to see a founder or CEO of a

company answering questions like “how did ABC company get their name?” But anyone can end up with the best answer to any question by dint of knowledge and their effort to express it.

**Crowd filtering and ranking:** Questions tend to elicit multiple answers, and in addition to authority the users vote for the answers they think best address the question. In this way, the best information “bubbles up” to the top of the list and silly or wrong answers are pushed out of focus. Over time this leads to an inventory of “right” answers to interesting and useful questions. This is a filtering mechanism that works very well.

**Intersection with networks:** You can “bring your network” to Quora by following people you already know and also start following new ones. The combination of following people and following questions leads to a more productive filter that still allows for discovery.

There are some other benefits that result from the Quora approach. The first is that people stop adding their 2 cents when it’s not needed. After a question has been addressed for a while, the best answer is right on top and is typically quite good. This helps prevent subsequent visitors from adding content that doesn’t improve the answer and instead encourages them to vote, comment, answer other questions or ask new ones.

Quora is also a good place to build your network and meet new people because it’s based on shared interests, knowledge and constructive interaction. In other words, it’s about the content that people generate more than the jobs they’ve had or people they know. LinkedIn takes the latter approach and is not nearly as effective.

Is Quora a flash in the pan? That’s a question that has some people worried because they enjoy the service so much. Part of what has made Quora special is the quality of the dialog and the fact that so far the company has been willing to forgo a business model. Will a huge increase in the number of active users result in a “dumbing down” or a “smartening up and broadening out” of the content? Quora does have a few mechanisms that may protect them from the adverse effects of a very large user base but these will need to be tested.

The other main risk for Quora comes from within. Users are building something that may be of great value and the company will want to monetize it to stay in business. Advertising is the most common

path but doing so without jeopardizing the character, energy and quality of the site is a challenge. Right now the costs to run Quora are low and if they stay that way it may be possible to preserve the aspects that make the site special. Craigslist is an example of a site that was largely able to preserve their free and open model and only charge for specific things like job postings.

## WHAT CAN GOOGLE DO?

There are some deep concerns about Google being able to address the shortcomings of their search results. The focus on speed and general answers for search may fail to capture enough semantic processing to deliver improved results.

Of course the folks at Google know all about the problems we have described here, and have myriad solutions they could apply. Technologies like instant search make it clear how deep their technology reservoir is. **Google now finds itself facing the innovator’s dilemma.**

Google has been disruptive by giving away valuable services “for free” since they could earn billions of dollars in advertising fees. Today that’s a huge and very profitable business for Google. If the path to continue to delight users happens to involve showing far fewer ads, is it one that Google can afford to take?

Based on recent statements, it sounds as if Google will make noise about improving results but is unlikely to do anything radical. As the market leader they can afford to move slowly, at least for now. However, there are a few fairly simple but powerful changes Google could implement to improve results:

**User-defined site rankings:** A simple way to provide positive and negative feedback on sites would make a big difference. A heavily bookmarked site generally has good content. Offering the opposite would be sensible. There are many sites for which I would check the box “never show search results from this site.” Google could make this easy.

**Leverage social and professional networks:** There are social and professional graphs emerging that also go a long way to defining some notion of authority or at least superior judgment. Google already can show me a search result that has been bookmarked by someone I’m following on Twitter.

Ranking sites bookmarked or tweeted by people one follows would be a helpful option.

**Penalize sites with high advertising content:**

This is crux of Google's problem. Users want more content and less advertising. We've reached the point where pages have more ads than content and viewers are beginning to find other places to look. Google isn't in danger yet but if they don't do something about the advertising proportions of returned results they'll lose share. It might take years, but once a decline starts it's hard to stop. [Look how many years it took for Microsoft Internet Explorer to shrink from having the vast majority of market share to now less than 50% and falling.]

**More collaborative filtering:** Taking bookmarking and combining it with search and networks would allow Google to show users new sites in results lists that are more likely to fit the nature of the searcher. This would generally improve the quality of new sites being discovered.

Making enough changes to produce high quality results would probably put a dent in advertising revenues.

The race is on in terms of who will do the best job integrating social graphs and networks into filtering information and search results. One likely outcome is that search traffic will spread out more evenly among more players and be less concentrated in Google. It's already more natural to search Yelp if you are looking for restaurant reviews, Amazon if you want to buy a book, eBay if you want to buy a used motorcycle, Foodily if you want to find recipes, and Craigslist if you are looking for an apartment. This trend will continue and erode some of the power Google has in the market.

## TIME FOR A NEW SEARCH ENGINE

Internet content management will be an unending cycle of massive undifferentiated expansion followed by the tools and attenuation need to turn it into real intelligence. Recognizing that it's a cycle scientifically and realizing that business models end

up getting embedded into the existing players means that starting a new search engine company is a good idea.

However, "search engine" is probably not the right term and should be replaced with something more active like "content finder" or, more charmingly, "librarian."

As noted, users are already searching on individual sites that tend to offer either the general content they like (as in National Public Radio or The Economist) or better answers to specific questions (like Yelp for restaurants and Amazon for books.)

There's a "curation" craze on right now that should die out and be replaced by intelligence based on actions (bookmarks) and relationships (social networks) that can scale more effectively.

This is a huge global market that offers extremely high returns on invested capital. Why aren't we seeing more investment in better information finders? We've seen a few like Blekko and Wolfram but there should be more. It turns out it only costs about \$25M to create a fully functioning, scalable search engine. Adding some intelligence to the results and offering improved results would create a large multiple of value on that investment.

There are also some existing companies that deserve a second look. For example, a search property like InfoSpace has a small market share but enough of a presence to do something interesting in this market and improve their fortunes. Other companies like Oracle, IBM and HP are in a strong position to create technology and online services that rival Google. We're due for a little disruption in search.

## ABOUT RESEARCH 2.0

We have been building a new and unique brand of emerging technology research since 2005 with some new characteristics:

**Open:** Most research is seen and used by a small fraction of the interested audience. Our research is distributed broadly to more institutions, investors, company managers and individuals because we leverage the power of the Internet.

**Long Term Intrinsic Value:** We use a more reliable model for determining company valuation that is very effective over multi-quarter periods. It also avoids the noise of small quarterly fluctuations. It enables investors to exploit volatility.

**Focus on Emerging Technology:** Our coverage starts with technologies that are in the early stages of commercialization and stretches to those that are driving the bulk of industry and market growth. We avoid technologies that have reached a plateau or are declining.

**Interactive Platform:** Today, people are embracing more online and frequently updated information sources. Our research is targeted to the online community with distribution by email, blogs, syndication and social and professional networks.

**Professional & Independent:** We measure our success by the quality of our work and the independence of our research process.

Research 2.0 is based in Boston, Massachusetts with locations in Paris, New York and Connecticut. Visit our website for more information and to contact us: [www.research2zero.com](http://www.research2zero.com).

Attachment A

Demand Media

DMD

11-Feb-11

Dec YE	2008	2009	2010	2011	2012	2013	2014	2015	2016	DMD	Ticker
YoY Change \$		28.2	51.6	75.0	100.0	125.0	125.0	125.0	200.0	Nasdaq	Exchange
Total Revenue	\$170	\$198	\$250	\$325	\$425	\$550	\$675	\$800	\$1,000	26%	Rev Growth
YoY Growth		16.6%	26.0%	30.0%	30.8%	29.4%	22.7%	18.5%	25.0%	\$17.00	Current Price
COGS %	57.7%	57.7%	48.0%	47.7%	47.0%	47.0%	47.0%	47.0%	47.0%	83.0	Shares Out
COGS \$	\$98.2	\$114.5	\$120.0	\$155.0	\$199.8	\$258.5	\$317.3	\$376.0	\$470.0		
Gross Profit	\$72.0	\$83.9	\$130.0	\$170.0	\$225.3	\$291.5	\$357.8	\$424.0	\$530.0	1%	Avg. Dilution
Gross Margin	42.3%	42.3%	52.0%	52.3%	53.0%	53.0%	53.0%	53.0%	53.0%	\$1,411	Cap (M)
S&M %	9.0%	10.1%	9.6%	9.4%	9.2%	9.0%	8.8%	8.6%	8.4%	\$40	Cash
S&M \$	\$15.3	\$20.0	\$24.0	\$30.55	\$39.10	\$49.50	\$59.40	\$68.80	\$84.00	\$0	Debt
R & D %	8.4%	10.9%	10.0%	10.0%	9.8%	9.6%	9.4%	9.2%	9.0%	30%	Tax Rate
R & D \$	\$14.3	\$21.7	\$25.0	\$32.50	\$41.65	\$52.80	\$63.45	\$73.60	\$90.00	22.5	P/E Multiple
G & A %	16.5%	14.4%	14.4%	14.0%	13.5%	13.0%	12.5%	12.0%	11.5%	15%	Discount Rate
G & A \$	\$28.1	\$28.5	\$36.0	\$45.5	\$57.4	\$71.5	\$84.4	\$96.0	\$115.0		
Operating Margin	8.4%	6.9%	18.0%	18.9%	20.5%	21.4%	22.3%	23.2%	24.1%	\$29.59	Intrinsic Value
Operating Income	\$14	\$14	\$45	\$61	\$87	\$118	\$151	\$186	\$241	74%	Up/Downside
Taxes	\$4.3	\$4.1	\$13.5	\$18.4	\$26.1	\$35.3	\$45.2	\$55.7	\$72.3		
Tax Rate	30%	30%	30%	30%	30%	30%	30%	30%	30%		
Net Income	\$10	\$10	\$32	\$43	\$61	\$82	\$105	\$130	\$169		
Net Margin	6%	5%	13%	13%	14%	15%	16%	16%	17%		
Market Value Using P/E	\$226	\$216	\$709	\$968	\$1,372	\$1,854	\$2,371	\$2,923	\$3,796		
Cash Position			\$40	\$83	\$144	\$226	\$332	\$462	\$630		
Shares (M)	83	83	84	85	86	86	87	88	89		
Period Share Price	\$3	\$3	\$8	\$11	\$16	\$21	\$27	\$33	\$43		
PV of MV 4 Years Out	\$785	\$1,060	\$1,355	\$1,671	\$2,170						
PV of Cash 4 Years Out	\$82	\$129	\$190	\$264	\$360						
PV MV + Cash	\$867	\$1,189	\$1,545	\$1,935	\$2,531						
PV Value Per Share	\$10.44	\$14.33	\$18.43	\$22.86	\$29.59						