

Is Relevance Relevant? Market, Science, and War: Discourses of Search Engine Quality

Elizabeth Van Couvering

Department of Media and Communications
London School of Economics

In the face of rising controversy about search engine results—that they are too restrictive, too comprehensive, lacking in certain areas, over-represented in others—this article presents the results of in-depth interviews with search engine producers, examining their conceptions of search engine quality and the implications of those conceptions. Structuration theory suggests that the cultural schemas that frame these discourses of quality will be central in mobilizing resources for technological development. The evidence presented here suggests that resources in search engine development are overwhelmingly allocated on the basis of market factors or scientific/technological concerns. Fairness and representativeness, core elements of the journalists' definition of quality media content, are not key determiners of search engine quality in the minds of search engine producers. Rather, alternative standards of quality, such as customer satisfaction and relevance, mean that tactics to silence or promote certain websites or site owners (such as blacklisting, whitelisting, and index "cleaning") are seen as unproblematic.

doi:10.1111/j.1083-6101.2007.00354.x

Introduction

An ongoing debate about search engine quality is gathering pace in both academic journals and public forums. The debate is not by any means unified and often seems contradictory. Search engines have been criticized for providing too much access to information—for example, by indexing sites promoting child pornography, anorexia, or hate (Finkelstein, 2004; Sullivan, 2004), as well as by releasing information about their users to governments or to a wider audience (Barbaro & Zeller, 2006). Conversely, search engines have been criticized for not providing enough access to information—for example, by poorly indexing dynamic sites, failing to refresh their own index (Bergman, 2001; Cothey, 2004), or by restricting access to certain websites when viewed from certain nations, as in the debate over the censorship of results in China (Amnesty International, 2006; Johnson, 2005; Marshall, 2005).

Search engines have also been criticized for providing biased information, that is, results that prioritize and include more American, more commercial, and more

popular sites than might be expected by taking a representative sample (Kleinberg & Lawrence, 2001; Lawrence & Giles, 1999; Mowshowitz & Kawaguchi, 2005; Vaughan & Thelwall, 2004). They have been criticized for providing access to harmful information, by directing users to spyware products and other damaging software (Edelman, 2006). Finally, search engines have been criticized simply for their lack of transparency—in other words, for the fact that no one really knows how the results come to take the form that they do, since the indexing and ranking algorithms are closely guarded trade secrets (Hargittai, 2000; Introna & Nissenbaum, 2000; Machill, Neuberger, & Schindler, 2003). These ongoing controversies reflect a lack of consensus about the nature and quality of the information that search engines ought to provide.

Quality is an important normative issue, yet despite its importance, it is problematic to study empirically. Any measures that a researcher might choose to examine would, of necessity, be controversial. By examining search engine producers, this article highlights accepted and contested views of search engine quality within the community of practice in which search engines are produced. Specifically, it asks:

RQ1: How do search engine producers conceive of quality?

This work builds on structuration theory, which suggests that conceptions of quality are part of larger collective schemas that affect how resources are allocated on a larger scale. Thus, the study also asks:

RQ2: What are the implications of these conceptions of quality for the future development of search engines?

The primary aim of this article is to articulate the unspoken assumptions of search engine producers regarding their views in the debate about search engine quality. The article begins by articulating the concept of the technological schema, a discursive formation through which technology is given meaning and which is used to mobilize other resources such as money and labor. The second part of the article discusses the schemas in evidence in a series of interviews conducted with search engine producers, primarily high-level engineering staff. Each of the schemas identified ascribes meaning to search engine technology in a different way, using different definitions of quality. The third part of the article examines how the different definitions of quality inherent in the technological schemas are used strategically by the search engine producers to control the development of search engine technology. Finally, the article considers how these technological schemas constrain both the possible interpretations of quality and the mobilization of resources around alternate frameworks by which search engine quality might be assessed.

Theoretical Background

The key theoretical underpinning of this article is the concept of ‘structuration,’ a reflexive social theory drawn from the work of Anthony Giddens (1984), which proposes that people’s actions and discourses influence social systems but are also in

turn influenced by them. From this perspective, technology can be seen to be an element of both system and social action, which are interacting dynamically. Several different frameworks and terminologies have been used to describe this interaction, including, from the science and technology studies (STS) perspective, the “technological frame” (Bijker, 1995) and “technological frame of reference” (Orlikowski & Gash, 1994); and from social and cultural theory, “interpretative schemes” (Giddens, 1984), “interpretative repertoires” (Potter & Wetherell, 1987), and “schemas” (Sewell, 1992).

The starting point for this article is Giddens’s “interpretive scheme,” which is a way of accounting for and explaining the world. According to Giddens, “‘interpretative schemes’ are the modes of typification incorporated within actors’ stocks of knowledge” (1984, p. 29). These interpretative schemes help to govern the allocation of resources and thereby reinforce or potentially change larger structures. Sewell (1992) helpfully develops Giddens’s concepts by suggesting how culture functions to develop social structures (which Giddens terms “rule-resource sets”). According to Sewell (1992), while material objects (such as money or buildings) have an independent existence, their simple existence is not enough to create value or power. For that, cultural “schemas” are needed. To take one example, paper money has no value without its meaning as a medium of exchange. Resources like money and technology, therefore, “embody cultural schemas” (Sewell, 1992, p. 19), but not unambiguously. The value of a resource is dependent upon the ways in which cultural schemas can mobilize that and other resources.

Both Bijker (1995) and Orlikowski and Gash (1994) acknowledge a debt to Giddens in developing their concepts. According to Bijker, a “technological frame” refers to the “shared cognitive frame that defines a relevant social group and constitutes members’ common interpretation of an artifact” (1995, pp. 125-126). Similarly, Orlikowski and Gash argue that a “technological frame of reference” governs the way in which people perceive technology as appropriate or inappropriate; thus their definition of the technological frame of reference “includes not only the nature and role of the technology itself, but the specific conditions, application, and consequences of that technology in particular contexts” (1994, p.178).

Returning to Giddens (1984) and Sewell’s (1992) arguments about the relationship between rules and resources, therefore, we might suggest that the “technological frame” or “technological frame of reference” is the way in which producers cognitively organize technology. This frame then not only helps to interpret technology and give it meaning, but also aids the producers in mobilizing other resources around their interpretation.

Viewing a technological frame in this way also allows us to account for the *strategic* use of different frames. Potter and Wetherell (1987) argue that each speech act is an achievement on the part of the speaker, and that different “interpretative repertoires” have different functions for speakers and will appear in different contexts. For example, Potter and Wetherell review an earlier study by Gilbert and Mulkay (cited in Potter & Wetherell, 1987, p. 147 ff) where scientists talk about

the quality of experimental science. In their investigations, Gilbert and Mulkey found two interpretative repertoires, which they named the “empiricist” repertoire and the “contingent” repertoire. The empiricist repertoire stressed the primacy of scientific data in accounting for the acceptance of one theory over another, while the contingent repertoire stressed external factors such as personal characteristics or political affiliations. Each repertoire in their study was deployed in particular circumstances—for example, in formal journal submissions versus informal contexts or within the same conversation to explain away contradictions arising from inconsistencies in a single repertoire. The strategic use of interpretative repertoires fits more closely with the empirical data in this study than the unified technological frame approach of either Bijker (1995) or Orlikowski and Gash (1994).

This article uses the nomenclature of a *technological schema* to denote a cultural schema (in Sewell’s sense) that includes an important technological element, used to account for actions and strategically to mobilize resources. In turn, following the duality of structuration, the technological schema itself meaningfully both constrains and enables the agency of the actors.

Methodology

As discussed above, this article analyzes the technological schemas of search engine producers. The analysis was approached through a series of in-depth interviews with search engine producers, primarily senior engineers and technical executives involved with directing code development (see Table 1).

While not large, the sample is comprehensive in that it includes interviews from producers in all the major search engine companies and some of the minor ones (interviewees were currently working for or had worked for companies including Google, Yahoo!, MSN, Ask Jeeves, AOL, Excite, Lycos, Infoseek, and WebCrawler, among others). These producers can be characterized as elite interviewees. The period of data collection was thus quite long—between November 2002 and May 2004—and each interview typically required several contacts prior to the interview itself. All interviewees requested anonymity; concern for the security of their jobs was acute for many. To protect their confidence, interviewees are identified by letter only, and all search engine names are anonymized and referred to in the text as Engine 1, Engine 2, etc.

Four strategies were used to reach appropriate interviewees: formal approaches to the companies, locating interviewees through existing contacts, locating interviewees directly through research on technical websites and forums (e.g., the social networking site orkut), and recommendations from other interviewees. In all cases, the approach began with emails or faxes that introduced the project and were addressed to the potential interviewee or, in the case of formal approaches, to the head of the search engineering group or the executive responsible for the search division. These were then followed up with phone calls. Of the final tally of producer interviews, two were the result of formal letters, two were the result of the researcher’s

Table 1. Interviewees, job functions, and organizations

<i>Interviewee code</i>	<i>Job function (job titles are capitalized)</i>	<i>Organization</i>
A	Founder and lead programmer of early search engine	Currently unaffiliated; formerly major search engine provider
B	Chief Scientist	Major search provider
C	Project engineer for specialized search	Major search provider
D	Former Head of Operations at early major search engine. Developer of another early search engine.	Currently not working within search industry
E	Program Manager, Relevance	Major search provider
F	Program Manager, Search	Major search provider
G	Founder of small search engine. Founder and lead developer for early search engine; former fellow with special responsibility for search at a large media conglomerate.	Small specialist search provider; formerly several major search engine providers
H	Senior Vice President of Technology	Major search provider
I	Founder and Head of Engineering. Former senior engineer at major search engine.	Startup search provider
J	Project leader for specialized search	Major search provider
K	Director of Product Development, Europe	Major search provider

personal contact network, four were the result of contacts made from Internet research, and three were recommendations by another interviewee. All of those who agreed to participate were men. In part, this is due to the shortage of women in computer programming and high-tech fields; in part, it is due to the lack of women in leadership positions in large companies.¹ While my interviewees indicated that they did have women colleagues, all of those women declined to be interviewed. It is unfortunate that we do not know what perspectives women producers might have; this remains an area for future research.

The interviews were semi-structured, in-depth, and primarily conducted over the telephone. Interviewees were informed in advance of the general purpose of the research, were told the questions that would be asked, and were directed towards the interviewer's website in case of further questions and to establish the researcher's bona fides. The interview questions covered the interviewees' histories and their job functions and then turned to specific instances of changes to the search engine algorithm: What motivated the changes, who was involved, what results were forthcoming, etc. The interviewees were asked about search engine optimization and spam, and how they were dealt with within the organization. Finally, interviewees had the opportunity to comment on the research. Most interviews lasted approximately one hour, although some were closer to two hours. The shortest interview was 35 minutes.

Each interview text was then analyzed to develop a preliminary categorization of likely schemas. Strong and unusual verbs and nouns with particular connotations were the first indicators. For example, one interviewee said, “I fought in the search wars” (Interviewee A); the use of the verb “fought” and the noun “war” indicated a potential schema of conflict or war. Each transcript was reviewed in depth, and a final determination of major and minor schemas was made by the author. The distinction between major and minor schemas was operationalized as involving both quantity (major schemas were used far more frequently) and quality (minor schemas were used to frame shorter passages within the major schemas, rather than organizing long sections of discourse). The quotes assigned to a particular schema were then examined for the ways in which they answered the primary research question: How do search engine producers conceive of search engine quality?

Schemas of Search Quality

The results showed that two major schemas structure the development of search engine technology. The first I have chosen to call the *market* schema, because discourse in this schema refers mainly to business-related issues: costs, revenues, and competition. The second major schema I call the *science-technology* schema; its discourse is dominated by experiments, measures, proof, and utility. Minor schemas were also noted, primarily a *war* schema whose discourse was one of enemies and combat. Although the schemas are analytically distinct, in practice they were not mutually exclusive, and one of the striking elements of the interviews is the ways in which interviewees negotiate among the different schemas.

The Market Schema

During the interviews, each interviewee was asked to describe a time when there was a modification to the search engine he was working on, what the rationale for that modification was, and who was involved in the modification. This was intended specifically to bring out typical accounts and rationales for change, and implicitly quality, and the interviewees often described both general processes and specific incidents. The most common description and justification in all the interviews was with reference to business issues, including competition, revenues, and costs. My interviewees discussed the search engine as a commercial service, a product in a marketplace that is highly competitive. In the market schema, interviewees related decisions to alter the functionality or the display of the search engine to revenues, costs, or competitive goals.

Quality as Customer Satisfaction

The market schema links revenue to quality through the idea of “customer satisfaction.” This makes sense when considering that companies exist to create profit and wealth, so an increase in wealth is a quality business outcome. The belief that more customers lead to more revenue is unquestioned, and measures of customer

satisfaction are based on the idea that satisfied customers will recommend other satisfied customers, leading to increased revenue, whereas dissatisfied customers will leave and tell their friends, leading to a revenue decrease.

As discussed above, discourse that invokes the market schema stresses the business rationale for changes. An excellent example is the following quote from a senior engineer who relates the way in which changes to the search engine are discussed and developed before being implemented:

Well, I mean basically if a change is suggested there needs to be some kind of motive for it ... [gives examples of specific motives] and these things all drive towards market share, which of course is the ultimate goal, which leads to revenue, etc. (Interviewee B)

In this comment, the change is articulated within a market or business framework, where the strategic goals are leveraging assets, market distribution, and market share (the examples of motives that he gives). This is the “ultimate goal,” according to the interviewee, because it leads to revenue in a competitive environment.

Despite the fact that the discourse of the market predominated in the interviews, the interviewees seemed generally hostile to interference from other parts of the company in the search product, and in particular to the demands of advertising, as the quotes below illustrate:

Product managers come up with completely irrelevant types of features they want to see implemented. So, for instance, instead of focusing on core technology, they ask you to put in yet another link or yet another space for ads in the interface. (Interviewee C)

I had to sacrifice a portion of the homepage to promoting their stuff. Which was pathetic. (Interviewee A)

[I]t was clear to us that if we started to give too much weight to the advertisers in terms of our index, we would dilute the value of our product. (Interviewee D)

“Irrelevant” and “pathetic” it may be, but advertising is central to the search business (Van Couvering, 2004), and a reduction in advertising means a reduction in revenue. Nevertheless, in the years leading up to when these interviews took place, many companies decided to move away from search results that returned many results from the parent company or that included hidden advertisements. Some of the interviewees were involved in arguing for these changes and mentioned their role with pride. They argued for, and secured, a reduction in advertising using a quality argument based on customer satisfaction as their justification.

First, one interviewee articulates the early strategy behind showing many results from the parent company:

[D]on't send users, our customers away [from the search engine] ... [instead] send them into our portal so we can help to monetize them again and all that good stuff. (Interviewee E)

The business rationale for advertising is clear: Keep the users in the portal and show them as many ads as possible. He then goes on to say why his company no longer follows that strategy:

[W]e are moving away from that, as policy, because fundamentally, it doesn't work. As it turns out, if the portal has what you are looking for, you'll go there. You will, right? If it doesn't, then sending you there just pisses you off and you look like a shill, because you are. (Interviewee E)

By "monetizing" the customers in accordance with business demands—that is, by sending them to other products so that they can look at more ads—this interviewee argues that the search engine will "piss [users] off" and damage its own reputation by appearing "like a shill" or conman. The risk to customer satisfaction is clear. Another interviewee takes up the tale:

It was a tough decision to make, because it meant a big revenue impact in the short term. But taking the long-term view, we knew that if we didn't do that, we'd probably have dissatisfied customers who would not want to use our service or not recommend it to friends, or maybe even switch to another service. So we were sort of taking the long-term view. (Interviewee F)

The short-term revenue impact is justified by arguing for long-term revenue ("sort of," as this interviewee says). A third interviewee continues in the same vein, by saying that immediate monetization—"controlling where people go"—can be dangerous, because "the users will stop:"

You have to be subtle in controlling where people go. You can't just only show them your own content. You can't hit them over the head. But you can certainly influence ... It's tough, right? [There are] editorial concerns as to where you drive people. But it can obviously only be done without affecting perceived quality. If the user doesn't think they are getting the results they want, that won't fly. So you can't stick inferior products on the top of better ones, just because they are your products. The users will stop. They will object. (Interviewee G)

These quotes suggest that quality is linked to long-term satisfaction on the part of customers, which in turn is linked to revenue, the positive goal or norm of the market schema. The next important thing to remember is that it is the engineering production teams who define search quality (in terms of "relevance"), as is discussed in the next section.

The Science-Technology Schema

The second major schema identified, the science-technology schema, is characterized by discourse that includes experimentation, measurement, and proof (the more scientific aspect), and usefulness, feasibility, and design qualities such as "state-of-the-art" (the more technological aspect). The science constructed by this schema

is a positivist, experimental science that has objectivity as an essential norm. Technology is the application of this science to the problems of search and is focused on solutions and progress. Thus, in the science-technology schema, the search engine is both interesting in itself as a research object and as a potential solution to people's needs.

The science-technology schema is exemplified by quotes like the one below, where the interviewee, who is Chief Scientist for a major search engine, describes the procedure for making changes in their search engine algorithm:

Yeah, well we're constantly making changes. The key thing to understand is that search and indeed basically all Internet business is highly data-driven. One of the key components of what we do here is to develop a deep array of metrics with which we measure what is going on in the service. These are quality metrics. So a lot of the decision-making is really focused around observing deficiencies in some particular metric relative to where we'd like to be or relative to the competition, considering changes that would improve it, and often provedly improve it, because you can do a test and see what is the impact on the number of hits, and then know what would happen. So a lot of the work that I do is meant to be driven pretty objectively, and we tend to do that for most of what we do. (Interviewee B)

Here the interviewee gives as his reason for making changes elements of positive science: measures, observation, proof, and objectivity. It should be noted that the interviewee also speaks of measuring "relative to the competition;" this section will later examine the relationship between the science-technology schema and the market schema.

The next quote exemplifies a more technological point of view. Here the interviewee, the founder of a search engine that was successful in the late 1990s, is talking about how he came to develop the technology for that search engine:

The common thread is big problems ... I was working on all sorts of interesting things. One of the things I did from the very start was to try to make [Engine 1] useful to everybody, so that was a big effort ... Making sure that everybody can access it ... the analogy I was using at the time was, think of it as a pencil. You don't want your pencil to be some big complicated contraption that starts singing at you every time you pick it up. (Interviewee A)

The emphasis on "big problems" and things that are interesting for their own sake is an element of scientific research, but when the interviewee also goes on to talk about things being useful, accessible, practical, etc., the discourse becomes more applied or technological than scientific. It is worth pointing out that the way in which science and technology are constructed within this discourse is quite specific. The strong impression given is that science deals with measurable (if complex) facts that are causally linked, while the goal of technology is to use the knowledge of the causal links to enable the user of the technology to act on the world effectively and

efficiently. Both science and technology are, or should be, progressing. If x , then y . If only we can do $x+$, we should achieve $y+$.

As Interviewee B stated above, changes to the search engine in this schema are “meant to be driven pretty objectively.” The implication, of course, is that decisions are not always objective. In the next section, which discusses quality from the science-technology perspective, this contradiction is explored through a discussion of the concept of *relevance*.

Quality as Relevance

As discussed above, many interviewees equate search engine quality with “customer satisfaction.” A second major way to discuss quality that relates to the science-technology schema is *relevance*. The term “relevance” is borrowed into the discourse of search engine producers from information science, where it forms the bedrock of several traditional measures of information retrieval quality, including, for example, recall and precision. *Recall* in information retrieval refers to the proportion of relevant documents retrieved from the database. *Precision* measures the proportion of retrieved documents that are relevant (Singhal, 2001). These terms were developed for relatively small, relatively high-quality databases of documents, for example, news articles contained in Lexis-Nexis. In those cases, users do not want to miss any relevant documents (that is, they want their searches to have high recall), and they do not want to retrieve very many irrelevant documents (that is, they want their searches to have high precision). Metrics such as precision and recall are still part of quality-testing search engines, but what is contested is the underlying ability to categorize documents into relevant and irrelevant.

Discursively, relevance takes central stage when interviewees talk about changes related to technical quality, as the senior vice president for technology of a major search engine indicated when asked what motivated changes in his search engine:

Relevance. So relevance, freshness—I mean you can almost lump everything under relevance, but that’s such a big umbrella. (Interviewee H)

What is relevance? In a small, well-defined database, it is relatively easy to sort relevant from irrelevant documents. On the Web, this is not necessarily as simple. One interviewee commented that the standard of relevance has changed from when he began to work with information retrieval systems:

[W]here the systems used to only be the Dialogues and the Lexis-Nexises, you know, I think they strove for a more academic standard of relevance, where you define relevance as the relationship between the subject that is in the document with what the user is asking about. So it is sort of topical relevance. Whereas in the practical world where the search engines are reaching today, something being useful to the user and something where the user grabs the information and continues, has become, I think, more important and less emphasis on say, getting the best document. (Interviewee G)

In other words, as this interviewee says elsewhere, it is about “satisfying users.” Relevance has changed from some type of topical relevance based on an applied classification to something more subjective. Most of the interviewees defined a relevant document as a document that answered the user’s question or was what he or she wanted:

Really, it is the standard definition, which is, we are trying to answer people’s questions. Period. Relevance is when we actually return something that answers their question. (Interviewee E)

From a technical standpoint, then, the definition of a quality search engine is simple: If the search engine gives you results that answer your question, then the search engine has delivered a relevant response and the results are quality results.

The War Schema

The war schema refers to discourse that is characterized by talk of fighting, guarding, war, defenses, and the like. This schema is a minor schema; the words or phrases that characterize it occur in sections of the interviews that are primarily related to the market schema or the science-technology schema. For example, an interviewee will talk about an “arms race” between spammers and search engine producers, will refer to competition as a “tough battle,” etc. Nonetheless, military words and expressions occurred with enough frequency and sufficient clarity to warrant inclusion as a separate schema. For example, one interviewee said, characterizing his time developing search engines: “I fought in the search wars” (Interviewee A).

The war schema has little to say regarding search engine quality. In contrast to the other schemas, the war schema is focused on others, particularly the enemy (and, by contrast, the identity of the speaker). In this context, the enemy is twofold. First, the competition (other companies) is characterized as the enemy:

[W]e’re not trying to beat [Engine 2] and [Engine 3] at their game directly, I think that’s a very tough battle, they’ve got lots of bright people, very well-paid, working on this stuff. And to try to go head-to-head, say on search quality, is a very difficult thing. (Interviewee I)

This is a fairly commonplace usage: “battling” other companies, going “head-to-head,” etc. The animosity implicit to the war schema is quite impersonal when referring to these respected opponents.

However, animosity towards the second class of enemy, the guerilla fighters of spamming and hacking, is more direct. Here one interviewee describes them as trying to “get at you:”

There is also an adversarial aspect to it in that you have hackers and spammers trying to get at you. (Interviewee F)

Sometimes the enemy can be trying to “get at you” by threatening your revenue, and sometimes by threatening your technology.

Talk within this schema is about “beating” the opponent or enemy. In other words, decision-making is characterized not by any kind of appeal to hierarchy, consensus, or objective measure but rather by who can “win,” even though several interviewees likened it to an “arms race” in which no one was likely to come out on top. This particular metaphor, the “arms race,” was not used about competing with other businesses. Spammers were also likened to criminals, particularly fraudsters or conmen, and specifically contrasted with “honest” people.

The war schema is not only important in defining relationships with other actors, it also provides a reflection on the identity of the producers as they assume the role of guardian or protector of something precious—in this case, access to the Web. Interviewee D does this explicitly in this quote:

We considered search to be important, we considered it to be a service that people needed and wanted and it was up to us guardians to make sure that we gave them the best experience possible. (Interviewee D)

The war schema is important because it frames much of the discussion about people *outside* the search engine organization. In this schema, the guardians of search defend against the incursions of the other, whether those others are honored competitors or fraudulent spammers.

In summary, therefore, the search engine producers use two primary technological schemas to ascribe meaning to search technology. In the market schema, the search technology is part of a business; in the technology schema, it is a piece of engineering work. Each schema has a concomitant definition of quality: either “customer satisfaction” or “relevance.” The minor war schema characterizes search technology as a defense, either against the competition (market schema) or against those who would affect the results for their own ends (technology schema). Quality issues are not specifically referred to, but one could draw the tentative inference that secrecy (from the competition) or robustness (against spammers) might be indications of quality.

Technological schemas are not only a method of accounting for and explaining technology, but also have real consequences for development as they function as a device to mobilize other resources, such as cash, office space, extra personnel, etc. Which schema to use, at which point, to mobilize what resource, is therefore a strategic question for the actors, with ongoing implications. Having discussed the schemas and their definitions of quality, the next section examines these implications.

Strategic Use of Different Quality Discourses

This section analyzes the strategic aspect of the schemas evident in the interviews. First, it investigates the ways in which, discursively, producers construct their own identity and agency as they talk about their work with search engines. Second, it

focuses on the recursive relationship between relevance and customer satisfaction and the construction of these two terms in such a way as to empower producers. Finally, it examines how these major schemas constrain the expression of alternative quality schemas.

Identity and Agency

As interviewees discussed their work and accounted for their actions, they used language which either implicitly or explicitly reflected their own senses of identity and agency, or their ability to act. When framing actions within the market schema, interviewees constructed themselves as significantly constrained, in marked contrast to the empowered constructions of the science-technology schema.

The grammatical structure of the interviews shows that interviewees overwhelmingly refer to actions and descriptions as part of a collective corporate “we,” typically using the pronoun “I” only when discussing personal matters or when they are unsure of the agreed corporate version.

For example, one interviewee discussed what led to the decision for his company to build their own search technology instead of purchasing listings from a third party, which had been their previous strategy. In the following quote, in which the interviewee accounts for a change to the search engine that began before he joined the company, note the use of “I” when the interviewee is uncertain, in contrast to the corporate “we” when he returns to more familiar ground:

I don't have a lot of background on that, but I would imagine, personally, just observing the explosion of information online, partially comments from our own customers and business partners, partially an observation of customer dissatisfaction with our search experience. We measure that on a very regular basis, and we care a lot about what the end user tells us, and we knew that customers weren't happy with various aspects of the service. (Interviewee F)

The interviewee also describes the way in which the decision was finally made, which involved the whole management chain; he later clarifies that it went the whole way to the CEO of the company. The interviewee goes on to say that “while I had responsibility for designing a service that we could operate to bring in revenue, I don't have final say on a decision that's going to have significant revenue impact” (Interviewee F). His perceived sphere of action—the changes he can and cannot make to the search engine—are not positioned as relative to his technical competence or his ability as a leader, but rather relative to business factors. In this case, he accounts for his own agency through the market schema: Changes that have no or little revenue impact are within his sphere; those with significant impact are outside it.

While none of the interviewees identifies himself explicitly as a “Microsoft man” or a “Google man,” there are explicit professional identifications that engineers and researchers present in the interviews. One early search engine developer talks about why he began to work on search engines in the mid 1990s:

[T]here was a need, the need wasn't being met, our collective Internet experience was less as a result of it, and we wanted to fix that! You know, we're engineers, we fix things. (Interviewee D)

In this quote, the interviewee identifies specifically as an engineer and very emotionally expresses why he belongs: because engineers fix things so that our collective experience can be greater.

It was notable during the interviews that, while the market schema is more pervasive, the interviewees were most animated and excited when expressing themselves using the language of the science-technology schema. Their voices rose with excitement, they spoke more quickly, they engaged more with the interviewer, and they almost took on the role of educator. In short, they seemed to be more comfortable with this way of expressing themselves and seemed to identify more emotionally with this schema. In the language of science-technology they spoke as experts, fully comfortable with their agency or ability to act, in contrast to the market schema where even very senior personnel were conscious of the limits of their actions.

Quality Control

The implication of the discussion of identity and agency in the previous section is that, within the dominant market schema, interviewees construct their ability to act as significantly constrained, whereas within the science-technology schema they are rhetorically empowered. Strategically, then, it is of benefit to these producers to be able to use the science-technology quality construction of "relevance." They do this, first, by constructing a rhetorical "customer" who does not correspond to the actual customer. This rhetorical customer is satisfied through greater relevance, unlike the actual customer. Second, the slippery subjective concept of relevance is quantified, reified, and discussed in highly technical language, which makes it unavailable to other actors.

The Rhetorical "Customer"

The term customer ordinarily refers to someone who buys products from a company. In that sense, greater customer satisfaction would, in most cases, lead to larger or more frequent purchases and positively impact the revenue of the company. In the discourse of search engine producers, however, "customers" are equated with users. Users—the people who type in queries and click on results—are not customers of search engines in an ordinary sense because they do not purchase products from search engine companies. In fact, the customers of search engines (in the ordinary sense) are the hundreds of thousands of businesses that purchase advertising and other services. The rhetorical customer/user serves the function of creating "customer satisfaction" through greater relevance, and thus making relevance a key benchmark of quality in the market schema as well as the science-technology schema. None of the interviewees mentioned changing the search engine to make it friendlier to advertisers, and indeed many were openly hostile to advertising, as discussed above.

Quantifying Subjectivity

Relevance is the linchpin of producer control of search engine quality. Recall that in the most basic terms, the relevant search engine result provides answers to the users' questions. Yet doing this is not a simple operation, since what the user wants is subjective, as the following interviewee points out:

[I]t is completely subjective based on the customer's frame of mind. So we are of course trying to develop models so we can figure out what that subjectivity is and therefore get our customers the best thing that we can do. But it is, you know, all ... completely whatever they want! And it changes—after they look at the first result, it can completely change. It's a little bit of Heisenberg. Really that is what it is. (Interviewee E)

The user/customer conflation is clear in this quote. The interviewee uses the language of the science-technology schema as he discusses “developing models” to “figure out what that subjectivity is.” What the interviewees are engaged in as they “develop models” is a process of making an objective, causal, and factual experiment out of the uncertain “Heisenberg” process of answering an often not-very-specific question on the part of the user. For example, what is the correct result for the query “new york apple” (a type of apple or a reference to the city)? How about “Napa” (both an auto-parts chain and a wine-growing region in California)? How about “abortion” (medical advice, addresses of clinics, or political issue)? The results generated from this process must be seen to be objective, and indeed Google specifically defends its results from accusations of inappropriateness on this basis: “Our search results are generated completely objectively and are independent of the beliefs and preferences of those who work at Google” (Google, 2004).

In addition to being objective, the results must be replicable and subject to improvement. This process of reifying an intensely subjective choice is a difficult moment for producers. The Chief Scientist of another large search engine was working on creating a new type of search algorithm for specialist queries at the time of the interview. He indicated that one of the most important parts of this work was creating a measurable, improvable model out of subjective preferences: “[I]n the case where something is brand new, we work very hard to understand how we would go about measuring it” (Interviewee B).

However, once the model is completed and the subjective has become objective, most of the day-to-day work begins. According to this chief scientist, about 80% of the work they do is incremental improvements to existing technology, also called “tuning” the search. Interviewee E is particularly enlightening on this topic as he is the Program Manager for search relevance at a major search engine. He goes on to describe in highly technical language, using specialized terms, how quality as relevance is constructed:

We have to use things like precision, recall at n or r is like 15. We have to use all of that good stuff. That's a secondary ... Primarily what we are trying to do is,

we are trying to figure out what is a model. Once we have a model, that is when you start to use things like precision at n . You have a belief that these documents are good for a given query. Now I can actually crunch some numbers and improve on that. (Interviewee E)

This very mathematical language, with the use of n and r (measures of precision in information retrieval) echoes that of Interviewee B, quoted earlier discussing how he can “provedly improve” search quality “because you can do a test and see what is the impact.” Thus, the account of the day-to-day process of “tuning” the search engine is one where the science-technology schema comes to the fore:

We assume that we have got ... we call them relevance judgments, right, or test sets, or whatever you want to call them ... So, you have been given this test set. Tuning is simply a matter of optimizing the test set. Effectively it’s a classification problem. You know, dividing documents from your entire corpus into good and bad. (Interviewee E)

Here again the tuning process is referred to via the specialized language of information science, discussing test sets, optimization, classifications, and the document corpus, which are known, identified elements of information retrieval technologies. Thus, according to this discourse, although the model for relevance is established by interviewing groups of users to see “what really is good and bad” (Interviewee E), the day-to-day management is really “pretty straightforward and relatively boring scientific number crunching” (Interviewee E). This “objective” relevance judgment is then re-incorporated into the world of the market by being used as a competitive measure to judge the quality of other search engines, as the quote from Interviewee B earlier in this section indicated. Here again, relevance as a measure, which works well within the organization, faces some strain from its subjective roots:

We have people whose job is based around relevance. And that is both relevance just looking at the result, it’s also relevance looking at the competition’s result and so on. Because just because it’s different does not mean it’s not relevant. Because sometimes a query, sometimes it’s subjective too—if you don’t know enough from a query, you don’t know what a user’s looking for. (Interviewee H)

In the quote above, relevance is used as a competitive yardstick even while its validity is being questioned. This was the case in more than one interview. One interviewee who formerly worked in a major media conglomerate as head of search was still shaking his head in bemusement years later over why people’s relevance judgments were not exactly in line with their assessment of search engine quality:

We did some studies internally when we were doing some stuff at [Engine 4], comparing how [Engine 4’s] results were compared to [Engine 3’s] and I

remember at one point that this experiment showed that our results were actually better but were being rated by the users as not as good. (Interviewee G)

The difficulty here is squaring results that are “actually” better—from the objective point of view of the experiment—with the subjective ratings of the users. It is vital for the producers that the reified and quantified “relevance” not be moved back to subjectivity, as the following interviewee articulates:

There are two features as far as relevance, right? One is, given a query, produce the most relevant things. That task in and of itself is to find that level, it becomes purely technical. The business side says “give us the most relevant thing,” that is all they care about. The related task, and this is where you start to see other aspects of the company come into play is when you are defining the language, for example you want to allow the customer to say what relevance means and give them some options. That is where various—the business and the marketing side will start to say “ooh, it would be really great to have xyz feature.”

(Interviewee E)

As long as relevance is the agreed “purely technical” measure, the “business side” just wants it to be better. As soon as the definition is put up for grabs, or companies “allow the customer to say what relevance means,” then sales and marketing begin to intervene in the development of the search product. Thus, the elision of relevance, quality, and customer satisfaction (elucidated above and in the section on quality in the market schema) is necessary for the producers to keep control of the search engine code.

The Difficulty of Articulating the Public Good

The last section discussed how the quality metrics of customer satisfaction and particularly relevance serve the strategic ends of producers, helping them to overcome the limits on their agency that the market schema implies. Nonetheless, the focus on relevance constrains the articulation of other quality goals. For example, in journalism, objectivity, fairness, diversity, and representation are typical examples of quality goals. In the course of this research, interviewees mentioned many everyday practices in search engine programming that could be considered censorship of search results and have the potential to lead to biases in search. These included blacklisting, or the exclusion of certain sites or site owners; whitelisting, or the automatic inclusion of certain sites or site owners; weighting content according to whether sources were considered to be authoritative or not; and adjusting results based on pressure from executives to respond, for example, to current news events. None of these practices were considered problematic, because all were linked to obtaining greater relevance in search engine results.

Not all of the interviewees were entirely happy about this state of affairs, reflecting the wider criticisms discussed in the introduction. In two interviews (D and I), there were hints of another minor schema, what we might call a “public service”

schema, emphasizing equal access and fairness. Yet even Interviewee I, who says that the goal of his search engine is to be “more transparent,” has a difficult time discussing an alternative quality judgment such as bias. He introduces the topic into the interview but also rejects it in the same sentence, which suggests that bias is in conflict with other ideas about search engines:

I don't think that major search engines today are horribly biased but they're also not objective, and the more you operate secretively, it makes it harder for people to see where you're subjective. Maybe subjective is a better word than biased, although it pretty much means the same thing. (Interviewee I)

Here the word “bias” seems to be dissonant with the discursive “objectivity” contained within science-technology schema's discussion of the day-to-day operation of search engines.

Another interviewee was asked what he thought about people who said that search results were becoming too commercialized. He refused to believe it was an issue:

This is not an example of the commercialization of search, but of the commercialization of documents available on the Web ... I mean if there are only commercial documents on the Web, then even a noncommercial search engine will come up with a list of commercial documents. (Interviewee C)

However, if there were to be an issue, he goes on to suggest that it would be technical and infrastructural in nature:

I see the bigger problem being centralization. The crucial part of your infrastructure is centralized, that is, you introduce a single point of failure. (Interviewee C)

The idea of a technical solution is echoed by another interviewee, whose utopian vision of the solution to a potential problem of bias or the over-commercialization of search included not only perfect technology but also education and literacy distributed throughout the world:

What we need to do is to create a world of educated, thinking people. Then the ads will have no effect, and we don't need to be concerned with what the corporations do. You need to look at it in a longer timespan—advertising is just a feature of a particular stage in evolution. Technology will get to a point where there is a big space, and so you need to make the profit on the difference and the quality, and yes there will be popular things. But the technology will distribute everything to everyone who needs it, and everyone will be able to find what they like. That's much better than regulation. (Interviewee J)

Perhaps more common is the view of a third engineer who simply suggested that if search engines were censoring their results, they were doing it for the good of the population:

When you get into this logic, unfortunately that's the tragedy of the commons, right, you have this free resource, people with this mindset will actually go and use it all, and destroy the value, so it's the same for the Web. I think search engines are totally justified to not be kind on those spammers. Sorry, I'm getting a little excited here. (Interviewee A)

The "guardian" of the war schema defends the Web in the quote above. It would be fair to say that quality schemas including ideas of full disclosure, representativeness, or diversity operate at a tangent to the way in which producers primarily frame their work.

Conclusion

The research questions that began this article were first, *how do search engine producers conceive of quality?* and second, *what are the implications of these conceptions of quality for the future development of search engines?* The evidence from the interviews examined in this article suggests that search engine producers conceive of quality in two separate but interrelated ways. First, a quality search engine, from the producer's perspective, has high *customer satisfaction*. This definition of search quality is embedded in a larger cultural schema that I have called the "market" schema, in which search engines are primarily conceived of as businesses. Second, a quality search engine produces very *relevant* responses to queries. Again, this definition of quality is related to the cultural schema that I have characterized as "science/technology." Search engines from the science-technology point of view are primarily pieces of engineering.

The implications of these conceptions of quality are far reaching precisely because they are embedded in larger cultural schemas. Structuration theory emphasizes how cultural schemas and their associated norms guide the allocation of resources. This article has shown that in the case of search engines, several schemas are at work simultaneously. The schemas clearly in the ascendant—the dominant market schema and the science-technology schema—provide little scope to raise issues of public welfare, fairness, or bias. Instead, they emphasize profit, in the case of the market schema, or progress and efficiency, in the case of the science-technology schema, or defense, in the case of the war schema.

A key feature of cultural schemas is that they are generalizable or, as Sewell puts it (drawing from Bourdieu) *transposable*; that is, "that they can be applied to a wide and not fully predictable range of cases outside the context where they are initially learned" (1992, p. 17). The market, science, and war are precisely such transposable allegories, whose demands constrain the pursuit of alternate standards of search quality. For, as Giddens (1984) and Sewell (1992) both note, schemas are crucial elements of structure that mobilize resources on behalf of their users. As the search engine producers discovered, the market schema is especially productive when it comes to mobilizing resources within the pre-existing structure of a business.

Producers, however, do not feel that they are experts in the market. They may be engineers, researchers, or designers, and they may have become managers or even business owners over the years, as happened to several of the interviewees. Nevertheless, they do not identify themselves as business people, and when working on search engines, they felt that their ability to affect major features of search was circumscribed as soon as their companies began to make significant amounts of money from the search. They act discursively to reclaim their abilities to control search by acting as experts on an objective measure of search quality, which is in turn linked to customer satisfaction and therefore revenue.

Thus, although search engine producers, like all other cultural actors, have a multiplicity of transposable schemas at their disposal (examples might include art, public service, family, or many others), the implication is that these schemas are not a simple matter of choice; they have real, material consequences for their users.

Structuration theory indicates that technological schemas and associated norms will have an effect on how resources are directed within society, helping to maintain old structures or create new ones. In this case, it is clear that a considerable amount of resources are dedicated to maintaining search quality in the forms of relevance and customer satisfaction. Structuration theory also indicates that pre-existing structures, such as corporate hierarchy or indeed capitalism as a whole, will profoundly influence which kinds of technological schemas and norms become widespread. In these interviews, the minor public service schema provides a potential alternative to the market and science-technology schemas. However, the difficulty its proponents have in articulating it, as well as the opposition of others to hearing it articulated, suggest how difficult it would be to build up an alternative kind of structure within which search engines could operate.

Note

- 1 According to the U.S. Department of Labor, in 2004 women made up 27% of the workforce in computer and mathematical occupations, including 25% of computer software engineers and 26.7% of computer programmers, and 36.7% of the workforce in management occupations, including 31% of computer and information systems managers and 5.9% of engineering managers (U.S. Bureau of Labor Statistics, 2005). It is possible that the number of women in the highly mathematical field of information retrieval would be less, although there is no direct evidence for this.

References

- Amnesty International. (2006). *Undermining Freedom of Expression in China: The Role of Yahoo!, Microsoft and Google*. London: Amnesty International. Retrieved March 4, 2007 from [http://web.amnesty.org/library/pdf/POL300262006ENGLISH/\\$File/POL3002606.pdf](http://web.amnesty.org/library/pdf/POL300262006ENGLISH/$File/POL3002606.pdf)

- Barbaro, M., & Zeller, T. (2006, August 9). A face is exposed for AOL searcher no. 4417749. *The New York Times*. Retrieved January 30, 2007 from <http://www.nytimes.com/2006/08/09/technology/09aol.html> [requires subscription].
- Bergman, M. K. (2001). White paper: The deep Web: Surfacing hidden value. *Journal of Electronic Publishing*, 7 (1). Retrieved March 4, 2007 from <http://www.press.umich.edu/jep/07-01/bergman.html>
- Bijker, W. E. (1995). *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*. Cambridge, MA: MIT Press.
- Cothey, V. (2004). Web-crawling reliability. *Journal of the American Society for Information Science and Technology*, 55(14), 1228–1238.
- Edelman, B. (2006, June). *Measuring the safety of Internet search engines*. Paper presented at the Rising Power of Search Engines Special Conference, Berlin.
- Finkelstein, S. (2004). *Jew Watch, Google, and search engine optimization*. Retrieved March 4, 2007 from <http://sethf.com/anticensorware/google/jew-watch.php>
- Giddens, A. (1984). *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Polity Press.
- Google. (2004). *An explanation of our search results*. Retrieved March 4, 2007 from <http://www.google.com/explanation.html>
- Hargittai, E. (2000). Open portals or closed gates? Channeling content on the World Wide Web. *Poetics*, 27(4), 233–253.
- Introna, L. D., & Nissenbaum, H. (2000). Shaping the Web: Why the politics of search engines matters. *The Information Society*, 16(1), 169–185.
- Johnson, T. (2005, July 17). In China, sophisticated filters keep the Internet near sterile. *Knight Ridder Washington Bureau*. Retrieved May 24, 2006 from Lexis-Nexis Executive database.
- Kleinberg, J. M., & Lawrence, S. (2001). The structure of the Web. *Science*, 294(5548), 1849–1850.
- Lawrence, S., & Giles, C. L. (1999). Accessibility of information on the World Wide Web. *Nature*, 400(6740), 107–109.
- Machill, M., Neuberger, C., & Schindler, F. (2003). Transparency on the Net: Functions and deficiencies of Internet search engines. *Info—The Journal of Policy, Regulation and Strategy for Telecommunications*, 5(1), 52–74.
- Marshall, M. (2005, July 7). Google facing search engine China quandary. *San Jose Mercury News*. Retrieved May 24, 2006 from Lexis-Nexis Executive database.
- Mowshowitz, A., & Kawaguchi, A. (2005). Measuring search engine bias. *Information Processing & Management*, 41(5), 1193–1205.
- Orlikowski, W. J., & Gash, D. C. (1994). Technological frames: Making sense of information technology in organizations. *ACM Transactions on Information Systems*, 12(2), 174–207.
- Potter, J., & Wetherell, M. (1987). *Discourse and Social Psychology: Beyond Attitudes and Behaviour*. London: Sage.
- Sewell, W. H. (1992). A theory of structure: Duality, agency and transformation. *American Journal of Sociology*, 98(1), 1–29.
- Singhal, A. (2001). Modern information retrieval: A brief overview. *IEEE Data Engineering Bulletin*, 24(4), 35–43.
- Sullivan, D. (2004). *Google in controversy over top-ranking for anti-Jewish site*. Retrieved March 4, 2007 from <http://searchenginewatch.com/searchday/article.php/3345071>

- U.S. Bureau of Labor Statistics. (2005). *Women in the Labor Force: A Databook* (No.985). Retrieved March 4, 2007 from <http://www.bls.gov/cps/wlf-databook-2005.pdf>
- Van Couvering, E. (2004, July). *New media? A political economy of search engines*. Paper presented at the International Association of Media and Communications Researchers, Porto Alegre, Brazil.
- Vaughan, L., & Thelwall, M. (2004). Search engine coverage bias: Evidence and possible causes. *Information Processing & Management*, 40(4), 693–707.

About the Author

Elizabeth Van Couvering is a doctoral student in Media & Communications at the London School of Economics. Her doctoral thesis examines bias in search engine results.

Address: Department of Media & Communications, London School of Economics, Houghton Street, London WC2A 2AE, UK