DEMYSTIFYING GOOGLE:

The Book Search Controversy and How to Understand It

by

Grant Patten

April 4, 2011

"In this pervasive atmosphere of privately acquiring, processing, and selling information, the public library system, a long-standing custodian of the idea and practice, of information as a social good, is tottering. Its function is being redefined and stripped of its social character." - Herbert Schiller, Information Inequality (1996)

INTRODUCTION

When it becomes clear that the activities of a private sector organization are quite contrary to the public interest, then that organization's activities should be duly regulated or replaced entirely with a public sector organization. Many legal scholars (Fraser, 2010; Samuelson, 2010; Vaidhyanathan, 2007) have argued that the Google Book Search (GBS) settlement poses a significant threat to the public interest. Some of them, especially Vaidhyanathan, make convincing arguments that the GBS project should be controlled by libraries instead of Google. Rather than explicitly agreeing or disagreeing with those sentiments, though, I will instead use a number of research methodologies to explore the question of how Google has managed to construct such a powerful position for itself. The objective is to "demystify the power of the powerful" (Law, 1992, p. 8) and, in doing so, shed light on some points of intervention that could potentially be used to counteract Google's dominance.

Unlike most commodities, the value of an information good, such as a book, derives from the content rather than the material form of that good. ("Information good," n.d.) The fact that information is "a peculiar commodity" (Bates, 1988, p. 76) that has resisted attempts at definition has no doubt helped Google with its GBS initiative. But by interrogating the question of how an information good should be valued, we can identify a number of concerns with the GBS. When benefits from the use of a good accrue to the user, that is known as private value. When they accrue to the supplier of the good, that is known as ancillary private value. When they accrue to society at large, that is

known as ancillary social value. (Bates, p. 83.) When someone makes use of a digitized

book in the GBS for their own reading pleasure, that is private value. Google will also

acquire ancillary private value here by getting paid for displaying targeted

advertisements to the user. The somewhat less obvious point here is that that person's

use of the GBS can also have ancillary social value to society at large. When people

read, generally speaking, they are educating themselves. And, as Bates says, "the

proper functioning of a democratic society is often said to rely on an educated

population that is able to make informed choices." *(ibid.)* But with the GBS settlement

agreement, Google is clearly privileging ancillary private value over ancillary social

value in that certain individuals and institutions will not be able to afford fully accessing

GBS content for reasons of monopolistic pricing practices. (Fraser, 2010, p. 1) Bates

writes, "there is currently no incentive for private firms or individuals to be concerned

with aspects of ancillary social value." (p. 86) Despite this reality, however, Google has

managed to appear as if it has the public interest in mind better than any private

organization in recent memory. How has Google managed to do this so effectively?

DESCRIPTION OF THE CASE

        In 2005, a class action lawsuit was brought against Google by the Authors Guild

and five publishers. In 2008, Google settled the lawsuit by agreeing to pay $125 million

to the plaintiffs. ("Google Book", 2011) The Department of Justice (DOJ) then

determined that the agreement violated US anti-trust laws and recommended to federal

judge Denny Chin against approving the agreement. One of the concerns the DOJ had

with the agreement was that it would allow Google to essentially determine the prices

for institutional subscriptions to its book search service. These prices would not be

based on market competition; therefore, Google could engage in price gouging without

much difficulty. The problem is amplified by the fact that the settlement stipulates that

only Google will be entitled to digitize all existing orphan works. All other potential

competitors would have to try negotiating digitization on an individual basis, a very

challenging task indeed considering that the rights holders are often nowhere to be

found. Based on the DOJ complaints, an amended settlement agreement was

submitted to judge Chin in November of 2009. The amended settlement agreement was

rejected by judge Chin in March of 2011. Chin reasoned that the agreement "simply

went too far... [it] would grant Google significant rights to exploit entire books, without

permission of the copyright owners." At the time of this writing, the various stakeholders

are planning to hold a conference in April of 2011 to discuss next steps. (Page, 2011)

OVERVIEW OF THE METHODS

    This paper applies two well-known methodologies for studying critical information

issues to the GBS controversy – the social construction of technology (SCOT) and

actor-network theory (ANT). SCOT is a good starting point because it allows us to

organize all of the different stakeholders around a complicated issue like GBS in a clear

and coherent manner. The framework directs us to categorize stakeholders as

producers, advocates, users or bystanders. (Humphreys, 2005, p. 235) Granted, life is

not that clear-cut, and a certain amount of porousness between categories is inevitable.

Advocates can also be users, for instance. A researcher must ultimately make an

informed value judgment on where to place each stakeholder. (Humphreys, p. 234) But

once these judgments are made and the chaos of a controversy is molded into a form of

some kind, opportunities for insightful analysis arise. ANT is the logical next step

because while SCOT is effective at describing the social world, it sometimes fails to pay adequate attention to non-human elements. ANT emphasizes that we must consider not only the human but also the non-human elements that make up a heterogeneous network (Law, 1992, p. 2) if we are to fully understand how power is established. These non-human elements often contain profound meaning and must be analyzed thoroughly if we are to understand Google's enormous success. The Google algorithms exemplify this notion and ANT provides many interesting ideas that can be applied to GBS. Rather than concentrating exclusively on one idea, this paper will instead reveal a number of ideas for future research and potential "points of intervention" against Google. But the paper does not make normative claims. Whether or not these points of intervention should be pursued is ultimately left up to the reader to determine.

POSSIBILITIES FOR STUDY

How has Google managed to generate this tremendous amount of goodwill for itself? Many would answer this question simply by stating, "because they're the best. They created the best algorithms." This technological deterministic perspective is quite widespread and should be countered by applying the GBS to the social construction of technology. (SCOT) Humphreys writes, "one can identify four general categories of social groups which span across technologies: producers, advocates, users, and bystanders." (2005, p. 234) The engineers at Google who code the algorithms and come up with the technical design of the GBS fall into the producers category. Humphreys points out that "designers' ideological positions influence their work." (p. 235) Because of the highly collaborative nature of the Google enterprise, it might be more fruitful to consider that statement in the context of the entire company. Google

engineers build Google's particular brand of ideology into its code – the algorithms are not neutral by any means. Google's algorithms have been known to favour big brand rankings over smaller brands, for instance. (Schwartz, 2009) So even if there is more user activity taking place on Slashdot around a discussion of Microsoft, Microsoft's various official websites will always pop up before Slashdot when "Microsoft" is typed into the Google search engine. Google's search results have also been known to favour their own services. For example, a search for "sore throat" at one point returned Google Health as the most prominent result. (Edelman, 2010) Investors and advertisers also play an important role in "the shaping of social construction and informing the development of the technology." (Humphreys, p. 236) The GBS, just like the traditional Google search, will feature targeted marketing from advertisers. The design of the interface will no doubt be largely governed by this business model. While a student is conducting research using GBS, then, she will have to contend with the potential distraction of advertisements. Presumably, a library or public sector organization would not incorporate an advertising component into its book search interface. A current example is the Internet Archive, (www.archive.org) which does not feature advertising.

Advocates are usually policymakers, lobbyists or academics who work to "convince the public not to buy something or to change a public habit." (Humphreys, p. 237) It is interesting to note how advocates who were ostensibly opposed to Google have actually helped it to establish its formidable position with the GBS. The members of the Authors Guild who initially sued Google were upset over copyright infringements. But it could be argued that Google actually wanted these advocates to sue them. Google was taking a major legal gamble when it deliberately infringed thousands of

copyrights by scanning books without permission. Google no doubt anticipated that various advocates would sue and expected a lawsuit. The advocates took the bait, and ultimately Google was able to negotiate a relatively painless settlement agreement for itself. Google would only have to pay the Authors Guild $125 million, ("Google Book," n.d.) not a particularly significant amount of money for a company worth billions. Samuelson writes, "use of a class action settlement to restructure markets and to reallocate intellectual property rights, particularly when it would give one firm a de facto monopoly to commercialize millions of books, is arguably corrosive of fundamental tenets of our democratic society." (2010, p. 34) Google was able to use the pressure from these advocates, then, to actually shape the development of the GBS technology in its favour. A possible point of intervention here might be to insist that such issues that seem so relevant to the public interest should actually be decided through a public legislature such as the U.S. Congress rather than privately through the court system. (Samuelson, p. 32) Perhaps then Google would not have managed to work itself into such a monopolistic position.

Users "help to construct the meaning of an artifact" through talking about, buying, and using a technology. (Humphreys, p. 238) Google has also benefited from the mostly positive reaction it has had to the GBS from users. The Google Books website features a User Stories page populated with comments like "today I was browsing Google [Book Search] for the first time and immediately found three in-print books on the subject I am researching. That's three books I am about to buy." ("Google Books User," n.d.) This is fairly indicative of the tame discourse around GBS in the press. The critical discourse around GBS remains confined mostly to academia, although Vaidhyanathan has been

making a valiant effort to bring these criticisms into the public eye with his *Googlization*

of Everything (2011) book. No doubt, the universally perceived usefulness of Google

search and its other services have helped Google to sustain its positive image among

the general public. Bystanders do not directly use the technology, but "they help shape

the cultural and social norms" (Humphreys, p. 239) surrounding the technology. A

variety of independent bloggers and websites critical of Google and the GBS can be

identified, such as Google-Watch.org and IHateGoogle.org. But this type of Google-

critical discourse remains on the fringes and lacks a strong form of organization.

Perhaps if the bystanders were to work toward evolving into the advocacy category,

they could make more of an impact. The authors needed the representation of the

Authors Guild. Similarly, it seems that these standalone anti-Google initiatives would

benefit from a collective organization of some sort. Maybe then the critical discourse

around Google would have an easier time finding its way into the public eye.

Closure is when "the interpretive flexibility of an artifact diminishes... [and the]

relevant social groups perceive their problems with an artifact to be solved."

(Humphreys, p. 242) The settlement agreement has been rejected, but because the

stakeholders are still debating about next steps to take, GBS might still be said to be in

the interpretive flexibility stage. Two of the most powerful stakeholders, however –

Google itself and the publishers/authors – have shown that they are capable of reaching

an agreement. Therefore, closure has already occurred between these two prominent

stakeholders. But Humphreys reminds us that, although it is difficult to do, closure

"always has the potential to be reopened." *(ibid.)* In order to see how the GBS closure

might be reopened, it would be useful to look at the historical example of the Google

search engine. Noam writes, "the search engine industry was initially dominated by

Yahoo!, Excite, AltaVista, and Infoseek. Google became popular by using a rank-based

search algorithm." (2009, p. 285) AltaVista's algorithms posed one solution to the

problem of Internet users not being able to find enough relevant websites through

browse-based interfaces. Google redefined that solution with its own unique rank-based

algorithm. The increased specificity that the Google results provided convinced enough

people to stop using other search engines, thus creating closure around the problem.

More recently, we have seen a number of competitors to Google appear with

redefinitions of the original search problem, such as Bing and Overture. Overture has

redefined the problem away from search specificity and more toward advertiser

freedom. Overture allows a lot more room for ad copy than Google's relatively stringent

restrictions. ("Copywriting," n.d.) Continually finding opportunities like this to address

problem areas that Google neglected in its original closure will no doubt allow upstart

search engines like Overture to take away some market share from Google. What can

opponents of the GBS learn from this? Rather than accepting that the GBS has been

closed for good once a legal agreement is approved, stakeholders could identify

potential problem areas that the GBS has neglected. This could potentially provide a

solid foundation for a competitor to create an alternative to the GBS.

Stabilization occurs when "the characteristics of [an] artifact come to be taken for

granted as the essential ingredients of the technology." (Humphreys, p. 243) Google

has been attempting to stabilize the search algorithms for the GBS. The enormous

success of the web search algorithms has no doubt informed the design of the GBS, but

Google engineers have been making tweaks to the code. PageRank, the algorithm used

for the web search, does not work with scanned books because of the lack of links

between pages. (Madrigal, 2010) Humphreys notes that many technologies have failed

in the past because of imposing "strict production and deployment guidelines

prematurely." (p. 243) Google is, admittedly, not the kind of traditional top-down

organization that is prone to make such mistakes. But there does seem to be a real

danger of premature deployment with the GBS. The current availability of the GBS has

revealed numerous technical issues that will likely not disappear once a legal

agreement is officially approved. "[Harvard librarian Robert] Darnton suggests that if

libraries and archives scanned books for a truly digital library —the job would be done

right , with none of the missing pages, botched images, faulty editions, omitted art work,

censoring, and misconceived cataloging that mar Google's enterprise." (Samuelson,

2010, p. 40) It is difficult to say exactly how Google develops its metadata because it is

not open source, but it is certainly fair to presume that the metadata developed during

the stabilization of the web search technology contributed to the development of the

book search technology. This has thus far resulted in a highly problematic technology

for Google, as the GBS is filled with tons of faulty metadata. (Nunberg, 2009) It seems

that a point of intervention here would be to design a book search from scratch with its

very own set of algorithms and metadata, rather than using stabilized web search

algorithms to inform the design of the new technology.

The fragmented consumptive audience (Humphreys, p. 243) of the GBS also

hinders its stabilization. "The most difficult part of making Google Books work, said

James Crawford, the team's engineering director, was determining the intent of the

service's heterogeneous user base." (Madrigal, 2010) Humphreys reminds us, "the

features of the mountain bike are *not stable* because different users have different

needs to which the Post-Fordist market responds." (p. 243) The technical problem of

faulty metadata and the social problem of heterogeneous users will both work to hinder

the stabilization of the GBS technology. Casual readers searching for the latest fiction

bestseller, for instance, might be quite pleased with the GBS. But academics who are

attempting to conduct research might be less satisfied. It seems that an opportunity for

intervention here would be to identify the needs that are left unsatisfied by the GBS. An

alternate technology could then be developed that benefits a particular "niche" audience

more than the GBS.

Flexibility of structure "includes how we think about an artifact's design and

engineering as separate from its use." (Humphreys, p. 244) There are three levels from

the most abstract to the most specific – superordinate, basic and subordinate. We might

consider "web search services" the superordinate level, "textual search services" the

basic level and "Google Web Search" or "Google Book Search" the subordinate level. It

is instructive to consider how it might be inappropriate to neatly group the GBS into this

hierarchy. It helps Google's position when it convinces people to think about the GBS as

merely a form of textual search. After all, the textual Google web search clearly

constitutes fair use because it is merely linking to material that has already been made

available online. But while the GBS may include a textual search component, it is far

more than a search engine. With the GBS, Google has taken material created by other

sources and put it online themselves. This is not a practice carried out by the regular

web search. Furthermore, the textual search technology does not even work particularly

well with GBS. Vaidhyanathan writes,

> Searching the text of books is rarely a better way to search than searching among books. Books are discreet documents that operate with internal cohesion more than external linkages. They are not "small pieces loosely joined," nor should they be. ...Privileging textual searching over more established forms of book indexing is a mistake. (2007, p. 1228)

While Google has been doing its best to fit the GBS comfortably within the textual search hierarchy, perhaps it is a mistake to think about book digitization as a form of textual search in the first place. Vaidhyanathan clearly believes that such book digitization work should be left to librarians trained in the traditional area of indexing. Google's engineers do not necessarily know anything about indexing and, as we have seen, applying textual search algorithms to books simply does not work very well. It seems that a promising point of intervention here would be to help the public understand that Google's construction of book digitization as a form of textual search is not the only option. Libraries could simply digitize a number of the same books as Google and use more traditional forms of bibliographic organization to show to people that the librarians are capable of doing a much better job than Google's engineers.

Actor-network theory (ANT) describes the process of translation, which is about "how actors and organisations mobilise, juxtapose and hold together the bits and pieces out of which they are composed." (Law, 1992, p. 6) It would be fruitful to consider how Google has been embedding its ideology into its technical artifacts. When asked about the possibility of imposing certain regulations on Google, CEO Eric Schmidt replied with another question: "would you prefer to have the government running innovative companies or would you rather have the private sector running them? ... We run Google based on a set of values and principles. And we work very, very hard to make sure people know what they are." (Vaidhyanathan, 2011) Law says, "thoughts are cheap but

they don't last long, and speech lasts very little longer. But when we start to perform

relations -- and in particular when we embody them in inanimate materials such as texts

and buildings -- they may last longer." (p. 6) So how has Google embedded Schmidt's

discourse into its code? Answering this question requires obliterating the notion that

there is anything neutral about Google's search engine. Aside from the aforementioned

example of "health," there are many other interesting instances of Google heavily

favouring its own services over others in its search results. Typing in "news," "video,"

"maps," "traffic," "product search," "blog," "email," "talk" and "calendar" all return Google

services either at or near the very top of the results page. (Cleland, 2008) Schmidt says

Google has to work very hard to publicize its products, but surely it isn't overly difficult

for engineers to tweak the algorithms in Google's favour. It is very easy to say that one's

company is innovative, as Schmidt does. Google ensures that those words are backed

up by constantly reminding searchers about its supposedly innovative products

whenever a word as generic as "video" is typed into its search engine. Google's

algorithms are like immutable mobiles – entities that are movable through space but

don't easily change. (Law, p. 6) One uses Google's algorithms to navigate their way

through different areas of virtual space whenever a new search is conducted. These

algorithms cannot be changed, however, except by Google's engineers because they

are not open source. The algorithms are rarely changed in a substantial way, and when

they are, the occurrence generally incites a substantial amount of press coverage.

(Tartakoff, 2011) So what can this teach us about the GBS? By identifying Google's

algorithms as immutable mobiles, we can realize that the algorithms are an important

means through which Google enacts its strategy. But the algorithms have been known

to break down from time to time, (Taher, 2009) and this is a point of vulnerability that could be exploited by potential competitors. This seems especially true in the case of the GBS, with its problematic metadata and search algorithms.

Law says that "translation is more effective if it anticipates the responses and reactions of the materials to be translated." *(ibid.)* Google's business model is very much about effective anticipation – predicting what its users want based on browsing habits and disseminating that information to advertisers. There is speculation that this model will be much the same once the GBS is fully implemented. Users will likely receive advertisements within the interface based on the books that they have been viewing. Google perhaps has an advantage with regard to organizational calculation, then, because it has become expert at anticipating how different groups of users will respond to online advertising. Presumably, then, Google will be able to effectively anticipate how users will respond to the fully implemented GBS.

But Law also urges us to ask "how is it (if at all) that the heterogeneous bits and pieces that make up [an] organisation generate an asymmetrical relationship between periphery and centre?" Google has traditionally established an asymmetrical relationship between periphery (everyday users) and centre (Google management) in its anticipatory business model. The asymmetrical relationship is not only a Google problem but a broader problem with online businesses in general. These companies predict what users want based on consumer profiling that is sometimes accurate but, more often than not, rather useless and sometimes even harmful. No kind of substantive communication or negotiation has been established between user and company in these business models, making them very asymmetrical. This asymmetrical

relationship affects not only the casual users of the GBS but also the authors of the

scanned books. Samuelson writes, "professional writers fear the consequences of their

loss of control over uses Google will make of their books. An author who has written a

critique of stereotypes of women as sex objects may, for example, be quite unhappy if

Google runs ads next to her text that promote the sale of sex toys or breast

enhancement surgery." (p. 25) It seems that this asymmetrical relationship could be

effectively countered by developing an alternative to the GBS that opens up a better line

of communication between users and advertisers. This is not merely a utopian idea;

there are online identity management (OIM) tools currently being developed to serve

this very purpose, such as Information Cards. ("Windows CardSpace," n.d.) A public

sector organization such as a library could use an OIM tool to develop a viable

alternative to the GBS that creates a more symmetrical, balanced relationship between

the periphery and the centre.

ANT allows us to develop an understanding of power by analyzing "the effect of

interaction between materials and strategies of [an] organisation." (Law, p. 8) This

seems to push us toward analyzing how Google's material practices interact with its

overall business strategy. The "Top 10 Reasons to Work at Google" list from Google's

official website gives us a glimpse of Google's strategy. Point four states, "work and play

are not mutually exclusive. It is possible to code and pass the puck at the same time."

("Top ten," n.d.) We can also discern that Google's work practices tend to coincide with

this encoded strategy. Aaron Swartz's article *how Google keeps employees by treating*

them like kids (2006) provides evidence of this. Orlikowski writes, "to the extent that

knowing 'what the organization is' is enacted in practice, we might usefully begin to think

about identity as an ongoing accomplishment, enacted and reinforced through situated

practices." (2002, p. 270) It seems that Google has been very clever about ensuring that

its written mandate is enacted through employees' practices, which has no doubt

contributed positively toward the development of its power. Google is a unique

organization in that it pays attention to its employees' practices and adapts itself to

those practices rather than attempting to impose any particular model of work from top-

down. It is in this sense that Google values noncanonical work practices (Brown &

Duguid, 1991, p. 40) and encourages its employees to experiment. Google is an

enacting organization in that it "[drives] innovation by allowing the parts of an

organization to step outside the organization's inevitably limited core world view and

simply try something new." (Brown & Duguid, p. 51) This is exemplified by Google's

policy to encourage their engineers to spend 20% of their work week experimenting with

personal projects. (Roque, 2010) The key point here is that Google's innovative

progress with the GBS thus far is due largely to the work environment that the

organization has fostered. A public sector organization such as a library could learn

from Google's encouragement of communities-of-practice, (Brown & Duguid, p. 41)

although it would not necessarily have to treat employees like children as much as

Google does. But in order to develop a book digitization technology within the public

sector that is capable of competing with the GBS, libraries need to foster a somewhat

less traditional and more experimental environment in which information professionals

are free to regularly try out new ideas. In fact, rather than merely emulating Google's

approach, libraries should study Google's practices and determine how they could be

improved. An article on TechCrunch entitled *Why Google Employees Quit* provides a

rich source of data on this topic. (Arrington, 2009) For instance, a common response

from TechCrunch users to the Google employees' complaints is that they all sound like

"spoiled children." *(ibid.)* Indeed, Google makes no attempt to conceal the well-known

facts that it encourages its employees to sit on big bouncy red balls while working and

decorates its Mountain View headquarters with a giant T-Rex skeleton. It seems, then,

that libraries could learn from Google in adopting its sensibility to foster experimentation

and even a certain degree of play amongst its employees. But libraries could also learn

to perhaps not take this sensibility as far as Google has, as if taken to the extreme it

might result in the creation of spoiled employees with a massive sense of entitlement.

The bigger Google gets, the harder it will likely be for it to retain its innovative approach

as an enacting organization. Rather than continuing to encourage spontaneous

innovation at all times, the reality is that the sheer size of Google will lead to employees

having their practices encoded into formal descriptions. Indeed, one sees evidence of

this in a 2008 post from ex-employee Aaron. Explaining why he quit, he writes, "the

bureaucracy and authoritarian 'gods of coding rules and regulations' were crippling for

an experienced developer." *(ibid.)* Libraries do not necessarily have to become as

technically innovative as Google, but it seems that there are weaknesses in the

organization that could be exploited.

Information infrastructure studies encourages us to consider how the GBS fits

within distributions along technical/social and global/local axes. (Bowker *et al.*, 2007, p.

6) If "an infrastructure occurs when the tension between local and global is resolved,"

(Bowker *et al.*, p. 7) then it could be surmised that the GBS is not an infrastructure yet

because this tension seems far from being resolved. The aforementioned problems with

the GBS metadata exemplify this point. The GBS metadata is very local in that it was

developed internally by Google's engineers. The engineers likely had a broader

audience in mind, but the GBS metadata has not transferred as smoothly to the global

sphere as the traditional Google web search has. Perhaps, as aforementioned, this is

because textual search simply does not work for books. From a global/technical

perspective, then, Google is not being very successful at designing a transparent,

universal metadata standard that works for everyone who uses the GBS. Google seems

determined to remain within the local/technical realm in which an artifact is "built on an

installed base." (Bowker *et al.*, p. 6) Indeed, the GBS seems to be built largely on the

installed base of Google web search. From a global/social perspective, it seems that the

reach/scope *(ibid.)* of the GBS is not broad enough to include all of its potential users.

With its current metadata structure, certain people are clearly being excluded and

misrepresented. "Once in place, information procedures and standards become a

general requirement or driver that facilitate development for some, that misrepresent or

perhaps ignore some, and that potentially alienate others." (Bowker *et al*., p. 10) The

aforementioned example of the feminist author who receives ads on sex toys while her

book is being displayed on GBS seems quite exemplary of this misrepresentation. It

seems that there may be an opportunity for libraries here to show greater "sensitivity to

[the] silences and absences that may be organizationally instantiated" by Google. *(ibid.)*

Perhaps if a metadata standard for a book digitization project were developed in the

public sector instead, it would fit more comfortably within the global/technical and

global/social axes. But, admittedly, developing a metadata standard is a highly technical

process that requires much expertise. Is there a false assumption here that the public

sector will be able to do a better job than Google at such a technical task? There is at least one historical example that suggests the public sector can do an excellent job – Dublin Core. This metadata standard was developed by the Online Computer Library Center (OCLC), a library consortium based in Dublin, Ohio and has become standard in the fields of library science and computer science. ("Dublin Core," n.d.)

Savolainen discusses the classificatory struggle between information behaviour and information practice. Scholars are often unsure about where to classify certain information seeking concepts and phenomena. He writes, "classifications... may be interpreted as examples of discursive struggles and ways to use the power to name the phenomena in question." (2007, p. 117) Google has been benefiting massively from a similar struggle over classification – transformativeness vs. derivativeness. Google has been using its legal and financial clout to insist that GBS represents a transformation of the original work and is therefore unproblematic. Arguing that the snippet-based interface of the GBS is transformative, "Google wants to minimize the importance of the original scanning of the book, the very copying that the publishers want the court to consider as operational and significant." (Vaidhyanathan, 2007, p. 1223) If Google is able to convince the courts that the GBS is transformative, then most of its legal problems would fade away because transformativeness constitutes fair use. But, as Vaidhyanathan points out, the distinction between transformativeness and derivativeness is just about as nebulous as the distinction between information behaviour and information practice. He writes, "when considering the composition of distinct creative works, it serves well to have a broad and strong sense of transformation. But... there is nothing close to consistency in the ways courts establish

the transformativeness of a use." *(ibid.)* So what is the lesson that can be learned here?

It seems a little simplistic to say that the legal system must work on establishing a

clearer distinction between transformativeness and derivativeness, but that is perhaps

exactly what has to be done. Savolainen writes, "there is a genuine need to generate a

self-reflexive and critical attitude among researchers toward their familiar concepts in

order to avoid being 'trapped' in their own discursive formations." (2007, p. 127)

Similarly, all stakeholders involved in the book digitization debate must think critically

about this issue and attempt to clearly articulate their opinions regarding what

constitutes transformativeness and what constitutes derivativeness. When a consensus

emerges, perhaps pressure could then be collectively applied to the courts to bring their

definitions more in line with what the majority of stakeholders in the industry consider to

be proper and fair definitions. Without a clear articulation of these definitions, private

organizations like Google will simply carry on implementing questionable services by

taking advantage of vague and ill-defined laws.

REFLECTIONS

Where did I start and how did my ideas evolve over the course of this project? I

started out wondering how commons-based peer production through sites like YouTube

has been affecting the music and film industries. That evolved into a concern for how

the book publishing industry is being affected by online digitization initiatives. Benkler is

optimistic for the music and film industries. He argues that the breakdown of copyright

has been hurting the recording side of the music industry, but the artistic side has been

mostly unharmed. With film, he argues that film markets will contract around commons-

based peer production and directors will find ways of sustaining themselves, just on a

smaller scale. (smokinrobocop, 2009) Is there room to be just as optimistic about the

book publishing industry in the face of technologies such as the GBS? I believe so, but

this research project has revealed to me that the issue in book publishing is less about

the technology itself than how it is allowed to develop. Legal scholars (Fraser, 2010;

Samuelson, 2010) have shown that Google would have claimed a monopoly over book

digitization if the settlement agreement had been approved. Granted, a monopoly might

not be problematic if the organization at least runs the service efficiently. But these

scholars have also argued convincingly that the GBS actually represents a highly

problematic, haphazard form of book digitization that could be improved upon massively

if developed through public sector institutions such as libraries instead. (Vaidhyanathan,

2007) Then again, it is probably not in the public interest to disallow Google from

bringing its unparalleled technical expertise to book digitization initiatives. Whether or

not Google should be allowed to continue with the GBS, then, is a debate that has fair

opinions on both sides. But what seems clear is that any future agreement relating to

the GBS requires a different approach than the stakeholders have taken thus far.

Fortunately, judge Denny Chin seems to have recognized this sentiment. But the

stakeholders are not simply going to leave it at that and, no doubt, a brand new

settlement agreement will be tabled to the courts before long. The concepts and

methodologies for studying critical information issues outlined in this paper could only

be of assistance to these stakeholders in identifying the best course of action to take

with regard to the thorny issue of the Google Book Search.

BIBLIOGRAPHY

Arrington, M. (2009). Why Google Employees Quit. Retrieved on April 1, 2011, from
        http://techcrunch.com/2009/01/18/why-google-employees-quit/

Bates, B. (1988). Information as an economic good: Sources of individual and social
        value. In V. Mosco & J. Wasko (Eds.), *The political economy of information*.
        Madison: University of Wisconsin Press. (pp. 76-94).

Bowker, Geoffrey C., Baker, K.S., Millerand, Florence and Ribes, David. (2009).
        Towards Information Infrastructure Studies: Ways of Knowing in a Networked
        Environment, in J.D. Hunsinger, M. Allen and L. Klastrup (eds), *International
        Handbook of Internet Research*: Springer.

Brown, J. S., and P. Duguid. (1991). Organizational learning and communities-of-
        practice: Toward a unified view of working, learning, and innovation.
        Organization Science: 40-57.

Cleland, S. (2008). New evidence of Google search bias -- Its relevant to DOJ
        investigation of Google-Yahoo ad-deal. Retrieved on April 1, 2011, from http://
        www.precursorblog.com/content/new-evidence-google-search-bias-its-relevant-
        doj-investigation-google-yahoo-ad-deal

Copywriting Differences for Paid Search Ads in Google and Overture. (n.d.). Retrieved
        on April 1, 2011, from http://www.seo-advantage.com/seo-topics/library-
        differences-google-overture.php

Dublin Core. (n.d.). In *Wikipedia*. Retrieved on April 1, 2011, from
        http://en.wikipedia.org/wiki/Dublin_Core

Edelman, B. (2010). Hard-Coding Bias in Google "Algorithmic" Search Results.
        Retrieved on April 1, 2011, from http://www.benedelman.org/hardcoding/

Fraser, E.M. (2010). Antitrust and the Google Books Settlement: The Problem of
        Simultaneity. *Stanford Technology Law Review* 4.

Google Book Search Settlement Agreement. (n.d.). In *Wikipedia*. Retrieved on February
        6, 2011, from
        http://en.wikipedia.org/wiki/Google_Book_Search_Settlement_Agreement

Google Books User Stories. (n.d.). Retrieved on April 1, 2011, from
        http://books.google.com/googlebooks/testimonials.html

Humphreys, L. (2005) Refraining Social Groups, Closure, and Stabilization in the Social
        Construction of Technology. *Social Epistemology*, 19, 231-253.

Information good. (n.d.). In *Wikipedia*. Retrieved on March 13, 2011, from
        http://en.wikipedia.org/wiki/Information_good

Law, J. (1992). Notes on the theory of the actor-network: Ordering, strategy, and heterogenity, published by the Centre for Science Studies, Lancaster University, Lancaster.

Madrigal, A. (2010). Inside the Google Books Algorithm. Retrieved on April 1, 2011, from http://www.theatlantic.com/technology/archive/2010/11/inside-the-google-books-algorithm/65422/

Natural monopoly. (n.d.). In *Wikipedia*. Retrieved on March 13, 2011, from http://en.wikipedia.org/wiki/Natural_monopoly

Noam, E. (2009). Media Ownership and Concentration in America. Oxford: Oxford University.

Nunberg, G. (2009). Google Books: The Metadata Mess. Retrieved on April 1, 2011, from http://people.ischool.berkeley.edu/~nunberg/GBook/GoogBookMetadataSh.pdf

Orlikowski, W. J. (2002). Knowing in Practice: Enacting a Collective Capability in Distributed Organizing. *Organization Science*, 13(3), 249-273.

Page, B. (2011). New York judge rules against Google books settlement. Retrieved on April 1, 2011, from http://www.guardian.co.uk/books/2011/mar/23/google-books-settlement-ruling

Roque, C. (2010). Mix Up the Workweek by Setting Your Own "20-Percent Time". Retrieved on April 1, 2011, from http://gigaom.com/collaboration/mix-up-the-workweek-by-setting-your-own-20-percent-time/

Samuelson, P. (2010). Google Book Search and the Future of Books in Cyberspace. forthcoming in Minn. L. Rev.

Savolainen, R. (2007). Information Behavior and Information Practice: Reviewing the Umbrella Concepts of Information-Seeking Studies. *The Library Quarterly, 77*(2), 109-132.

Schiller, H. (1996). Information Inequality. New York: Routledge.

Schwartz, B. (2009). Google's Vince Update Produces Big Brand Rankings; Google Calls It A Trust "Change". Retrieved on April 1, 2011, from http://searchengineland.com/google-searchs-vince-change-google-says-not-brand-push-16803

smokinrobocop. (2009). Yochai Benkler Conflicts in Cultural Production. Retrieved on April 1, 2011, from http://www.youtube.com/watch?v=0OhdSbQ3Yps

Swartz, A. (2006). The Goog Life: how Google keeps employees by treating them like kids. Retrieved on April 1, 2011, from http://www.aaronsw.com/weblog/googlife

Taher, A. (2009). 'Human error' shuts down Google. Retrieved on April 1, 2011, from
        http://technology.timesonline.co.uk/

Tartakoff, J. (2011). Who Gained From Google's Search Algorithm Change? Retrieved
        on April 1, 2011, from http://paidcontent.org/article/419-who-gained-from-
        googles-search-algorithm-change/

Top ten reasons to work at Google. n.d. Retrieved on March 13, 2011, from http://
        www.google.com/intl/en/jobs/lifeatgoogle/toptenreasons/index.html

Vaidhyanathan, S. (2007). The Googlization of Everything and the Future of Copyright.
        University of California, Davis 40(1207).

Vaidhyanathan, S. (2011). The Googlization of Everything (And Why We Should Worry).
        University of California Press. (Ch. 1).

Windows CardSpace. (n.d.). In *Wikipedia*. Retrieved on April 1, 2011, from http://
        en.wikipedia.org/wiki/Windows_CardSpace