This article explores the implications of a shift from public to private provision of information through focusing on the relationship between Google and public libraries. This relationship has sparked controversy, with concerns expressed about the integrity of search results, the Google Book project, and Google the company. In this paper, these concerns are treated as symptoms of a deeper divide, the fundamentally different conceptions of information that underpin the stated aim of Google and libraries to provide access to information. The paper concludes with some principles necessary for the survival of public libraries and their contribution to a robust democracy in a rapidly expanding Googleverse.

Abstract
This article explores the implications of a shift from public to private provision of information through focusing on the relationship between Google and public libraries. This relationship has sparked controversy, with concerns expressed about the integrity of search results, the Google Book project, and Google the company. In this paper, these concerns are treated as symptoms of a deeper divide, the fundamentally different conceptions of information that underpin the stated aim of Google and libraries to provide access to information. The paper concludes with some principles necessary for the survival of public libraries and their contribution to a robust democracy in a rapidly expanding Googleverse.

Contents
Libraries and Google
The romance phase: ‘We have everything in common’
Cracks appear in the relationship
Reality check: ‘We want different things’
The word ‘information’
Conclusion: Negotiating a healthy relationship

Libraries and Google
'To google' has become a household verb, meaning "to search for information on the Internet." [1] In the month of April 2008, Google handled 5.1 billion queries in the U.S. alone [2]. Its market share is almost 90 percent in the U.K. and Australia [3], 80 percent in Europe [4] and 70 percent in the United States [5]. By 2004 Google had indexed eight billion pages; in 2006 Google claimed to have worked out how to index billions more [6].

In the past, librarians' access to, and skill and familiarity with using databases and catalogues, indexes and reference works was what made them the specialists in information retrieval. Now librarians are told that people expect to find things at the library in the same way as they find things on Amazon, iTunes, and Google (Dempsey, 2006). In response, many libraries are trying to make their catalogue more 'Google–like' with a single search box and relevance–ranked results based on keywords (Calhoun, 2006). Ask a question using an online reference service and there is a good chance that the librarian at the other end will be using Google to answer the query, or demonstrating how to conduct more effective searches using Google. Some librarians advocate welcoming Google into the reference interview, not only to take advantage of Google’s search function, but also to take advantage of the relationship that patrons already have with Google, a relationship which they assume is positive (Cirasella, 2007). In Australia, at least one university library runs 'GoogleSmart' classes to help students learn to use Google effectively. Some librarians enthusiastic about the Google Book project consider that it means that there is no need for libraries to house large collections of physical books, freeing libraries to be 'high–end consulting centres' (Sandler, 2005).

Librarians have written extensively on Google and about the relationship between libraries and Google, much of it sparking heated debate [2]. A widely circulated set of 'provocative statements' about university libraries claims that before 2011 'All information discovery will
begin at Google, including discovery of library resources’ (Taiga Forum Steering Committee, 2006). Some in the library profession are unreserved in their admiration for the search engine; for example, Phipps and Maloney exclaim ‘All we really know is, if we could have, we would have invented it.’ [8] Others are less sure. This paper argues that the concerns about the relationship between libraries and Google are symptoms of a more fundamental difference concealed by the word "information".

For the purposes of exposition, the argument is organized using the metaphor of a romantic relationship. Public libraries, despite being a heterogenous collection of institutions, are treated as the one party in a relationship with the other party, Google. Metaphors structure our perceptions and understanding (Lakoff and Johnson 2003) and although there are no pretences here to any psychological insight about relationships, the relationship metaphor is deliberately chosen to highlight specific aspects of the association between Google and libraries. In particular, as this paper will show, use of this metaphor makes clear how the association between Google and libraries has implications for democracy. As Lakoff and Johnson argue, what is most important about the use of a particular metaphor ‘is not the truth or falsity of a metaphor but the perceptions and inferences that follow from it and the actions that are sanctioned by it.’ [9]

A quick diversion into the pop psychology of relationships is necessary in order to delineate the contours of the relationship metaphor and hence establish the framework for this paper. It is assumed that the general pattern of relationships contains the following stages, not as distinct sequences, but with particular stages being more prominent at various times. Initially, there is a romance phase when it seems that both parties have everything in common. A perceived breach of trust by one party can lead to the realisation that rather than having everything in common, each party wants different things. When one party is at risk of having their identity subsumed by the other, it is crucial for this party to regain a separate sense of self, a process that involves setting appropriate boundaries with the other party. It should be noted that while this paper focuses on the relationship between public libraries and Google, much is equally applicable to the relationship between public libraries and other commercial search engines or Internet companies. Google is here singled out because it is the most well-known, the most discussed in the literature, and as the paper will show, because of the enormity of its reach and ambition.

The romance phase: ‘We have everything in common’

Superficially, it seems that Google and libraries have much in common, namely the desire to provide access to information. Google’s stated mission is to organise the world’s information and make it useful. The mission statement of the American Library Association (ALA) includes the aim of ensuring ‘access to information for all.’ [10] Similarly, objectives of the Australian Library and Information Association relate to information provision and include promoting ‘the free flow of information and ideas.’ [11]

While librarians have been ‘welcoming Google into the reference interview’ (Cirasella, 2007) and trying to emulate Google’s simple interface, Google, in turn, has been ‘actively courting’ libraries and librarians (Williams, 2007). Google has ambitious plans to digitize the millions of books currently held in public libraries and put them online. This project apparently had its genesis in the frustration experienced by Google co-founder, Larry Page when, as a teenager, he was unable to find electronics manuals that would help him reassemble the gadgets he had taken apart (Vise and Malseed, 2005). The Google book project aims to make it easier for people to find relevant books. In its initial incarnation in 2004, as the Google Library Project, this plan led to lawsuits by publishers because of its disregard towards copyright restrictions (Haigh, 2006). Renamed the Google Book Project, the Project has been modified so that access to the complete full text is only possible for those works where the copyright has expired. Those libraries on board have welcomed the initiative as a way of opening up their collections when they could not have afforded to digitize their collections themselves. For example, in preliminary discussions between Oxford Library and Google ‘it was agreed that Google’s long–term mission and Oxford’s aim of digitizing and making its material available over the Web were well–matched.’ (Milne, 2005) Although Google won’t say how many books it has digitized so far, it is estimated that it is many millions of books each year [12]. While the Google Book Project links Google users with library content, as of May 2008, a user who finds an item through Google book search can also link directly to the catalogues of the closest OCLC member libraries which have that item in their collection [13].

The Google Book Project can be viewed as the perfect opportunity for libraries that could never afford to digitise their own collections. However, cracks are beginning to appear in the relationship between libraries and Google. Perhaps the first breach of trust was in relation to

[8] Others are less sure. This paper argues that the concerns about the relationship between libraries and Google are symptoms of a more fundamental difference concealed by the word "information".

[9] A quick diversion into the pop psychology of relationships is necessary in order to delineate the contours of the relationship metaphor and hence establish the framework for this paper. It is assumed that the general pattern of relationships contains the following stages, not as distinct sequences, but with particular stages being more prominent at various times. Initially, there is a romance phase when it seems that both parties have everything in common. A perceived breach of trust by one party can lead to the realisation that rather than having everything in common, each party wants different things. When one party is at risk of having their identity subsumed by the other, it is crucial for this party to regain a separate sense of self, a process that involves setting appropriate boundaries with the other party. It should be noted that while this paper focuses on the relationship between public libraries and Google, much is equally applicable to the relationship between public libraries and other commercial search engines or Internet companies. Google is here singled out because it is the most well–known, the most discussed in the literature, and as the paper will show, because of the enormity of its reach and ambition.

[10] Similarly, objectives of the Australian Library and Information Association relate to information provision and include promoting ‘the free flow of information and ideas.’

[11] While librarians have been ‘welcoming Google into the reference interview’ (Cirasella, 2007) and trying to emulate Google’s simple interface, Google, in turn, has been ‘actively courting’ libraries and librarians (Williams, 2007). Google has ambitious plans to digitize the millions of books currently held in public libraries and put them online. This project apparently had its genesis in the frustration experienced by Google co-founder, Larry Page when, as a teenager, he was unable to find electronics manuals that would help him reassemble the gadgets he had taken apart (Vise and Malseed, 2005). The Google book project aims to make it easier for people to find relevant books. In its initial incarnation in 2004, as the Google Library Project, this plan led to lawsuits by publishers because of its disregard towards copyright restrictions (Haigh, 2006). Renamed the Google Book Project, the Project has been modified so that access to the complete full text is only possible for those works where the copyright has expired. Those libraries on board have welcomed the initiative as a way of opening up their collections when they could not have afforded to digitize their collections themselves. For example, in preliminary discussions between Oxford Library and Google ‘it was agreed that Google’s long–term mission and Oxford’s aim of digitizing and making its material available over the Web were well–matched.’ (Milne, 2005) Although Google won’t say how many books it has digitized so far, it is estimated that it is many millions of books each year. While the Google Book Project links Google users with library content, as of May 2008, a user who finds an item through Google book search can also link directly to the catalogues of the closest OCLC member libraries which have that item in their collection.

[12] The Google Book Project can be viewed as the perfect opportunity for libraries that could never afford to digitise their own collections. However, cracks are beginning to appear in the relationship between libraries and Google. Perhaps the first breach of trust was in relation to
the restrictions that Google put on access to scanned library books in the Google book project. Google requires that it be the only search engine that can link to the digitized content of the Google Book Project [14]. Because of this restriction, some libraries have refused advances from Google to include their collections in the Google Book Project, preferring to work with the non-profit Open Content Alliance or Gutenberg Project [15].

Some of the reluctance on the part of librarians is because Google’s future business plans are unclear. Google has a history of perfecting the technology first and then working out how to make money from it. (Fabos, 2006). So, for example, it would not be unreasonable to imagine that perhaps in the future, while reading a Google scanned version of Mark Twain’s The Adventures of Tom Sawyer, sponsored links for canoes, and canoeing adventures, will appear on the side of the page or as pop-ups.

While some in the library profession are still in the romance phase and defensive about any criticisms of Google (for example, Sandler, 2005), others have started the ‘reality check’. A range of concerns are being voiced about the relationship between libraries and Google. (Many of these concerns are also relevant to other Internet companies; in particular, Microsoft and Yahoo who have both been involved in digitization projects with libraries.) An overview of these concerns is provided as necessary context: firstly concerns related to the integrity of search results, then concerns relating to the Google Book Project, and finally concerns relating to Google the company.

It should be noted that the list of concerns that follows has been compiled from existing literature about the relationship between libraries and Google. It is a list of expressed concerns rather than a coherent position regarding the relationship between Google and libraries. Bear in mind that the focus of this paper is not on the legitimacy or otherwise of these concerns; these concerns are merely presented as symptoms of a more fundamental incompatibility. Although both Google and libraries have the aim of providing access to information, they each want entirely different things. This paper provides a critical analysis of the different conceptions of information underpinning what appears to be the same aim. It will argue that public libraries are concerned with the content of information and aim to provide access to information in order to enable the placement of targeted advertising; the content is only important to Google inasmuch as it facilitates targeted advertising.

Concerns regarding the integrity of Google search results

Relevance of Google search results

As Brabazon (2006) points out, while access to Google gives the illusion of having the world’s information at one’s fingertips, Google is merely an automated search engine. As such, it specializes in information retrieval rather than reference. The effectiveness of a search engine at information retrieval is judged by precision and recall. While recall measures the ability to retrieve all the relevant documents, precision measures the search engine’s ability to retrieve only the relevant documents (Brin and Page, 1998). This raises the question of what constitutes a ‘relevant’ document. Implicit in Brin and Sergey’s claim that we want our notion of “relevant” to only include the very best documents, is the idea that ‘relevance’ is an objective measure of the relationship between the search term and the results. Actually, ‘relevance’ is a subjective judgement by the user. Morville makes this clear in his simple definition: ‘relevant results are those which are interesting and useful to users.’ [16] In other words, just because Google says it has returned the most relevant results, does not mean that it has.

Some influential librarians have been scathing in their criticisms of Google search results (for example, Gorman, 2003; Mann, 2007), with the former President of the Bibliothèque Nationale de France describing the results as ‘massive amounts of disorganized information.’ [17] While a search on Google can easily retrieve millions of pages of search results, almost two-thirds (62 percent) of search engine users do not go past the first page of results and less than 10 percent go past the first three pages of results [18]. This means that for most users, the results of a search are effectively confined to the first page of results. The "relevance" or usefulness of these results that show up in the first page has been questioned from the following angles.

While supposedly neutral, the Google algorithm has been criticized for favoring of the most popular sites. As Jeanneney puts it, ‘success breeds success, at the expense of newcomers, minorities, marginals.’ [19] This is not just an issue with Google. Introna and Nussbaum have analysed how it is intrinsic to the design of algorithm-driven search engines that they systematically favor particular types of site, ‘those with popular appeal and mainstream commercial purpose’ or ‘backed by entrenched economic powers.’ [20] Although the specifics of Introna and Nussbaum’s analysis are out of date, the general argument still holds. The mechanics of search engines can lead to the marginalization of dissent (Fabos, 2007; Grimmelman, 2009).

On the other hand, Google has been criticized for promoting ‘undesirable’ voices. Examples abound of finding inappropriate results using Google. For example, white supremacist sites have appeared as the third organic result on a search for "Martin Luther King" (Haigh, 2006).
Searching for information on ‘Jews’ retrieves links to anti-Semitic sites (Piper, 2005). Globalwarming.org has appeared as the first organic result of a search on “Global warming”. Globalwarming.org is a site devoted to refuting global warming, but funded largely by oil–related industries through the Competitive Enterprise Institute [21].

Another criticism relating to the relevance of Google search results points to the limited coverage. Of course, Google can only retrieve information that is on the Internet. Companies and governments have been shown to engage in Web scrubbing, the practice of editing or removing historical documents on the Web, to avoid financial or political losses [22]. Search engines like Google do not uncover these historical documents. In fact, search engines index only a fraction of online content. It was estimated that in late 2007 search engines covered less than 20 billion of the 900 billion pages of online information, that is, approximately two percent of the total [23].

**Sponsored results**

The results on the first page include those that are there because of the calculations of the algorithm (organic links) and those that are there because a company has paid for them to appear (sponsored links). Although Google claims that its ads are clearly labeled and separate from the search results, it has been found that most people do not distinguish between the sponsored links and the organic links, determining credibility by rank in the results page. This accords with research on advertising and publicity in traditional media. Lord and Putrevu (1993) found that when advertising is presented in the format of news or an editorial, people assume that it is credible because they have been socialized to think that news is objective and accurate. Pew Research found that almost two–thirds of people who used a search engine (62 percent) were not even aware that some search results were there because they had been paid for. Furthermore, most searchers (more than 80 percent) considered that they could not always distinguish between paid and unpaid search results (Fallows, 2005).

**Search engine bias**

Closely related to concerns about the inclusion of advertising in search engine results are concerns about search engine bias. As the founders of Google admitted in their famous paper on the structure of Google, ‘we expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers.’ (Brin and Page, 1998) Furthermore, Brin and Sergey considered such bias as ‘particularly insidious’ because of the extreme difficulty in evaluating a search engine or detecting search engine bias [24]. It has been noted that there is a certain objectivity and authority attributed to mathematical calculations, even though these calculations are based on particular assumptions (Roszak, 1986; Ruhleder, 1995). While it is notoriously difficult to get information about Google’s algorithm, a whole industry of search engine optimization has sprung up to try to engineer the appearance of particular Web pages in the first page of results for particular search terms [25]. Despite the susceptibility of the search engine algorithm to manipulation and ‘Google–bombing’ (Grimmelmann, 2009), users of search engines are increasingly placing trust in the authority of the information that appears at the top of the search results. For example, a 2006 study [26] showed that one–third of search engine users believed that the companies that appeared in the top results of a query were the industry leaders.

**Censorship of search results**

In late 2005, both Yahoo and Google came under fire for complying with the Chinese government’s request to censor particular sites from the results of searches conducted in China. Google also conducts its own form of censorship and is able to ban a site from appearing in search engine results because it has violated one of Google’s conditions of use (Jeanneney, 2007) [27]. There is no way that a user can know what results have been omitted.

**Confusion of information retrieval with knowledge**

A range of librarians, academics and social commentators blame Google for confounding information retrieval with knowledge, resulting in what they call ‘dumbing down’ or ‘the Google effect’ (Brabazon, 2006; Gorman, 2003; Haigh, 2006; Herring, 2005; Mann, 2007). Making a clear distinction between search and research, Brabazon argues that librarians and educators have a responsibility for ‘stressing that Google is the start — not the entirety — of a search.’ [28] Without entering the thorny debate about what constitutes knowledge, it can be appreciated that keying a few terms into a search box and cutting and pasting from a quick skim of the results is not the same as traditional methods of researching a topic.

**Concerns regarding the digitization project**

It has already been mentioned that some major libraries have refused to be part of the Google Book Project because of the restrictions in access to the digitised works. Below are some further reasons for skepticism regarding the Google Book Project. These relate to perceptions of bias in the selections of work for digitization, as well as concerns about digitization itself.

It is considered impossible to digitize all of the world’s printed resources, and the Google
approach has seemed somewhat ad hoc while also biased towards the English language (Jeanneney, 2007; Roush, 2005). The concern is that printed works which are considered to be part of a culture’s wealth will disappear. The works which will be digitized will be those ‘which are best suited to satisfy the demands of advertisers.’ [29]

In his examination of the quality of the scanning in the Google Book Project, Duguid (2007) assesses the quality of the scanning as ‘at times, completely inadequate’. His conclusion is that even if the quality of scanning is improved, there are intrinsic aspects of existing physical books that cannot be turned into digital information. Concerns have been expressed that the relationship between the reader and a digital text will be superficial in comparison with the intimate relationship that can develop between a reader or scholar and a physical text (Ruhleder, 1995). Haigh is not the only one concerned by the possibility that ‘searching a book will become an increasingly adequate substitute for reading it.’ [30] In addition, while the longevity of the physical book is proven, there are questions about the longevity of digitized information (Jeanneney, 2007; Roush, 2005).

Concerns regarding Google Inc.

In 1980 when the UNESCO General Conference passed a resolution for a New International Information Order, it was in official recognition that control of information is linked to the control of social and material realities (Pendakur, 1983). Whereas in 1980, there were concerns that particular countries were in control of information, now there are concerns that a single company could have a monopoly on information (Posner, 2007). With regard to the Google search engine, James Grimmelmann has particularly sober words about the political decisions inherent in what comes up first in the Google search results:

‘Whoever controls search engines has enormous influence on us all. They can shape what we read, who we listen to, who gets heard. Whoever controls the search engines, perhaps, controls the Internet itself. Today, no one comes closer to controlling search than Google does.’ (Grimmelmann, 2009).

With regard to the Google Book Project, Roush puts it succinctly:

‘Letting a for–profit organization like Google mediate access to library books, after all, could either open up long–hidden reserves of human wisdom or constitute the first step toward the privatization of the world’s literary heritage.’ (Roush, 2005).

On the other hand, however, rather than controlling information, Google could go out of business, rendering the digitized collections inaccessible (Jeanneney, 2007). Although Google may appear infallible and continues to post increased revenues [31], there is criticism that Google’s stock is overvalued (Ignatius, 2006) and predictions that people will turn away from algorithm–driven search. In South Korea, one of the biggest users of the Internet, a different type of model of search predominates, and Google’s market share is estimated to be as low as 1.7 percent [32] (Freedman, 2007). There are speculations that Google will not be able to survive the transition to mobile computing [33]. Some predict that Google will go the way of IBM, which went from owning two-thirds of the information technology business to being ‘just another company.’ [34] Whether one thinks that Google is more likely to end up having a monopoly on information or is more likely to go broke, the issue is one of the lack of public control over a private company.

Cracks appear in the relationship

It is clear that all is not rosy in the relationship between Google and libraries. Many, including prominent members of the library profession, have voiced concerns about the integrity of search results, their relevance, the impact of advertising, and censorship. Some libraries have refused to get involved with the Google Book Project, voicing concerns about the restrictions on access to content, bias in the selection of works, the quality of digitization and the limitations inherent in digitisation. There are fears that people will search rather than read books, and Google is blamed for ‘dumbing down’, conflating information retrieval with knowledge. Finally, there are concerns about the lack of public control over a private company. Wildly different scenarios are envisaged; Google may end up with a monopoly on information or is more likely to go broke, the issue is one of the lack of public control over a private company.

Google has tried to respond to some of this criticism. For example, since the Google Book Project began, Google has tried to shake perceptions that the Project is limited to English–speaking libraries. As of May 2008, in addition to partnerships with the library of Oxford University and libraries in America, Google had also established partnerships with university libraries in Switzerland, Japan, Belgium and Spain as well as a German State Library and the National Library of Catalonia [35]. With regard to the censorship of results in China, Google
co–founder Sergey Brin has publicly admitted that it was a mistake [36]. When Google–bombing leads to offensive results, Google now includes a disclaimer "Offensive search results. We’re disturbed about these results as well". To return to the relationship metaphor, however, even if Google could respond to each of the concerns outlined above, this would not, in itself, cover the cracks that have appeared in the relationship between Google and libraries. It is proposed in this paper that the various concerns outlined above are all symptoms of a deeper problem with the relationship between Google and libraries. The core of this problem is that, far from having a common aim, Google and libraries want completely different things. Although Google and libraries each aim to provide access to information, the context and interpretation of this similar aim is very different.

Reality check: ‘We want different things’

**What libraries want**

The rationale for the public library mission to provide access to information for all is made explicit in the following principle of the Australian Library and Information Association: ‘Freedom can be protected in a democratic society only if its citizens have unrestricted access to information and ideas’. [37]

Historically, public libraries have not always been associated with unrestricted access to information and ideas. At times public libraries have had the express purpose of preserving and extending hegemony (McNeely and Wolverton, 2008), extreme examples being libraries in Nazi Germany (Dalton, 1992) and the Soviet Union (Abramov and Skvortsov, 1978; Kasinec, 2001). Even now, public libraries in Western democracies can be said to practice covert if not outright censorship (Moody, 2004).

The “grand tradition” of public libraries was articulated most famously by McColvin in the 1950s. Public libraries were to have a key role in advancing democracy, promoting knowledge, and empowering citizens through the possibility for self–education (McColvin, 1956). This ‘grand tradition’ of public libraries has been frequently criticized for being exclusively white and middle–class (Bundy, 1978; Dawes, 1978; Newman, 2007). In Britain and the U.S., the post–war open access movement marked the beginning of a conscious effort on the part of public librarians to shed an elitist, gatekeeper heritage and become accessible to everybody, rather than just a minority of middle–class library users (Gerard, 1978; Newman, 2007). The idea that libraries have an important role in assisting democracy through making information accessible to everybody has continued into the twenty–first century.

Democracy, however, requires a balanced flow of information (Pendakur, 1983). As Dervin (1994) and Lievrouw (1994) point out, citizens need unrestricted access to alternative voices as well as that information which serves to legitimise existing institutions and power dynamics.

It should be noted that a balanced flow of information is an ideal to aspire to. In reality, it may be that the status of authority tends to be conferred on information that supports the existing power structures (Dervin, 1994; Lievrouw, 1994; Webster, 2002). For example, Dervin argues that in the area of information about health, authority is conferred on allopathic medicine while “alternative health approaches” are excluded. Lievrouw is even less optimistic about the achievement of the ideal of balanced information, suggesting that most of the information we receive is “subject to institutional gatekeeping processes involving agenda–setting or "spin control".” [38]

Drawing on Kierkegaard and Dreyfus (2001), one could argue that, in addition to balanced information, democracy requires some sort of priority given to more significant information. In the context of the rise of the popular press, Kierkegaard argued that too much information gives rise to nihilism. Dreyfus revisits Kierkegaard’s argument to suggest that the free flow of information on the Internet can similarly undermine democracy. There is an enormous amount of information on the Internet and the structure of the Internet means that the information is not ranked in terms of importance. Dreyfus concludes that it is too difficult for citizens to make sense of this overabundance of information with the inevitable result of citizen indifference and passivity. He advocates some sort of hierarchical ordering of information in terms of its significance.

Given that for public libraries, the aim of providing access to information is bound up with ideals of a healthy democracy, it would seem that libraries have a role in providing balanced information and giving some sort of priority to more significant information.

**What Google wants**

As already mentioned, Google’s stated mission is to organise the world’s information and make it universally accessible and useful. When looking at how Google tries to achieve its stated mission, attention is usually focused on Google’s patented mechanism for information
retrieval, the Google search engine. In this paper, it will now be argued that the search engine has become just one of the many tools that Google uses to organise the world’s information and make it universally accessible and useful. Actually the Google search engine doesn’t organize information in any conventional sense. It indexes Web pages so that they can be retrieved using a free text search and the only organizing of Web pages is to order search results using its algorithm. However, it does organise information in a way that is useful for advertisers. Through its pay–per–click operations, search engines like Google enable advertisers to target advertising according to user intentions and interests as revealed through what users search for on Google and where they go on the Web. Google collects and retains a vast amount of personal information on each person who uses its services. Each Google search a user undertakes is associated with the computer she is using and is recorded (via a cookie on the hard drive). A user who installs Google toolbar on their personal computer is effectively sending back to Google a record of every Web site they visit, regardless of whether they use the toolbar (Piper, 2005). This enables personalised search, whereby the search engine ‘remembers’ a user’s previous searches and activities on the Internet, so is better able to provide a relevant response to the searcher. For example, if the search is for Casablanca and previous searches have been for Hollywood movies, then the search engine will only show results related to the movie Casablanca, rather than the place Casablanca. While some librarians have welcomed the prospect of such personalized search as a way of improving search for the user (Eisenberg, 2008), others are more wary of it as a way for advertisers to increase the impact of their advertising (Röhle, 2007).

Google has much in common with a commercial television network. Both make their revenue from selling advertising and hence are driven by the need to attract advertisers. TV attracts advertisers through ratings, which may have nothing to do with quality of content (Hassan, 2004). Similarly, Google attracts advertisers through the sheer volume of traffic, not necessarily the quality of the content it delivers. For both, the quality of content is only important insofar as it attracts viewers in the case of TV, or users in the case of Google. However, no matter how this search engine is perfected, it can only retrieve information that is ‘on the Internet’, in other words, digital information that is accessible via the Internet. In the words of the founders:

“Sometimes you don’t get a good answer to a search because the information simply isn’t available on the Web. So we are working hard to encourage ecosystems that can generate more content from more authors and creators.” [39]

The ecosystem they are referring to here is Knol, a question and answer service with user–generated content. However, Google is doing much more than putting what is conventionally understood as information on to the Web. The next section argues that all becomes information as Google encourages us to live on the Web.

**All becomes information as Google encourages us to live on the Web**

Just as TV networks have a financial incentive to extend hours of broadcasting and thereby increase advertising revenue, so Google has an incentive to increase its advertising revenue by adding to the amount of information on the Web. Google initially did this through extending its reach into the ‘invisible Web’, extending its coverage of online content and through digitization of books. However, at the same time, it has been steadily expanding what we normally think of as information. Writing in 2004, Morville says, in reference to Google’s mission, ‘the tricky thing about Google is (they) keep changing their definition of what we normally think of as information. First, it was just Web sites. Then images, online discussions, blogs, news … Google’s mission, ‘the tricky thing about Google is (they) keep changing their definition of information. First, it was just Web sites. Then images, online discussions, blogs, news … then, the contents of your desktop computer. Then, the contents of the worlds’ largest research libraries … ’ [40]

As well as trying to ensure that information is accessible to all, Google is involved in trying to make sure that people are accessing more and more information via the Web. Google has done this by pioneering a brilliant new model of business expansion, introduced here as infogration. Infogration is radically different from the traditional model of horizontal integration which involves buying up competition, and vertical integration, which involves buying upstream and downstream industries. Infogration involves capturing different aspects of physical and social reality and representing them with digital information. In other words, infogration involves the integration of aspects of the world in to the medium of information into which targeted ads can then be placed.

To be successful, infogration requires that we live more of our lives on the Web. Hence, Google has been actively encouraging us to live more of our lives in the “Googlevverse”. Google wants us to not only use its search engine to search for information, and to read using Google Books but also to use Google products (a shopping site) to buy things, to plan to buy things using Google Shopping List (a Web–based list of intended purchases), to let friends know what we want for our birthday or Christmas using Google WishList (a public list of what we wish we could buy), to find out what is happening using Google News, to share tagged photos using Google Picasa, and to hang out with friends using Orkut (Google’s version of a social networking site, very popular in Asia and India), or connect with people with similar interests using Google Groups. The list continues and expands with each new Google product. Google encourages us to communicate using Gmail, to plan trips using
Google Maps, to bring Google along for the ride with Google Maps integrated into our cars and giving us directions [41], to virtually visit other places using Google StreetView, to describe ourselves in blogs (using Blogger) and be updated with the blogs we are interested in using FeedBurner. Google suggests that we manage our financial information using Google Finance and that we manage our health using Google Health (a repository for all of our medical records and a way of keeping our doctors up-to-date about our health).

Lastly, and perhaps most bizarrely, Google is working towards being able to encourage us to describe ourselves in terms of our genetic make-up. Google aims to build up a genetic database, analyzing the genetic information of individuals and populations (Vise and Malseed, 2005). Google is a Series A investor in the personal genome service 23andMe (cofounded by Anna Wojicki, partner of Brin Sergey, and former Google product manager). The somewhat disingenous mission of 23andMe is to be the world’s trusted source of personal genetic information, enabling ‘individuals to access and learn about their own genetic information’ [42].

All becomes information as Google encourages us to live on the Web. Once infograted, these aspects of our physical and social reality (the world and our lives) are, in principle, accessible from anywhere, at anytime. Of course, they come complete with relevant ads. In addition, the business model of infogration leads to the production of new information which are particularly useful for advertisers, enabling profiling of consumers to an unprecedented extent.

From a database of intentions to a live feed of practices

In his book Search, John Battelle (2005) coined the phrase ‘the database of intentions’, to refer to the aggregate of all online searches. Writing in 2004, he correctly predicted some future trends with regard to what he saw as the goal of perfect search. He correctly predicted a trend towards ubiquity, the ability to search on mobile devices, and the ability to find real-world objects embedded with digital information. Still in its early stages, an example of the latter is the ability to type a tracking number into Google and track the progress of RFID-tagged courier items. Battelle correctly predicted a trend towards personalization, and anticipated large-scale digitization of analogue recordings, such as books, films, television, and music.

Google, however, has gone much further than Battelle envisaged in Search. Search terms alone are a blunt instrument. Rather than relying on a database of presumed intentions, Google is beginning to leverage from a live feed of actual practices, as it tries to infograte the very substance of our world and lives. In the mid-1980s, Webster and Robins wrote that new technologies for the organisation of information are crucial for the system of mass consumption, controlling consumer behaviour through ‘the stimulation of needs, the recording of tastes, the surveillance of consumption’. [42] The technologies they were writing about then were things like supermarket check–out scanning. The capacity of these technologies to collect and organize consumption information is completely eclipsed by twenty-first century technologies (see, for example, Andrejevic, 2006). The suite of Google Products enable personalized online advertising, designed according to consumer preferences revealed by online and offline activity.

The business model of infogration enables tracking of the following aspects of our lives. What we search for, what we buy, what we intend to buy, what we read, who we hang out with, what we talk about, where we go (in the virtual world and the real world) and who we are, both metaphysically (through blogs and communications), and literally (through health and genetic information). Google’s business is all about tracking consumption habits and putting targeted advertising in front of us as often as possible. To illustrate: if you are looking at your genetic information online and find that you have a gene for obesity, expect ads for weight loss programs, weight loss drugs, and appetite suppressants to appear whenever you are online.

The word ‘information’

In discussing Google’s aim of organising the world’s information and making it universally accessible and useful, we have deviated far from public library concern with information necessary for civil society and a functioning democracy. How is it possible for a similar aim of ‘access to information’ to lead down such divergent paths? Google is assisted in their enterprise by the ‘ideology of information’, the political work that the word ‘information’ does in its connotations of neutrality (Agre, 1995; Roszak, 1986; Schiller, 1994). As Roszak observes ‘A fact, a judgement, a shallow cliché, a deep teaching, a sublime truth, or a nasty obscenity. All are “information”’. [44] The word ‘information’ becomes emptied of meaning and like the word ‘community’, it becomes generally acknowledged as ‘a good thing’, regardless of its nature. It becomes taken for granted that it is something that citizens need access to in order for democracy to function well. However, as Roszak observes, ‘countless bits and bytes of data sizzling across the global power grid’ [45] have little to do with
democracy.

We can go further in tracing the origins of the ambiguity inherent in the word ‘information’. The word ‘information’, is doubly articulated (Silverstone, et al., 1994), referring both to form and content. Roszak describes how in the 1940s, mathematical engineer Claude Shannon appropriated the word ‘information’ to refer to signals (ones or zeros) that can be transmitted electronically. Shannon was concerned with the quantity of bits transmitted rather than their meaning. A meaningless jumble of letters may have the same ‘information’ as a message that secures the future of the planet. Whereas Shannon’s definition of information referred to form, another standard understanding of information is that it is data which has meaning ascribed to it, or more precisely, ‘some pattern of organization of matter and energy that has been given meaning by a living being’ (Bates, 2005). This definition of information refers to the content of information. These two different meanings of the word information are reflected in the different concerns of Google and public libraries. For the company Google, information is a medium for targeting advertising. Hence, Google is concerned with the free flow of digital information, information that is accessible anywhere anytime. In other words, Google is concerned with the form of the information. In contrast, public libraries aim to provide access to information in order to strengthen democracy. This requires a balanced flow of information and some sort of ordering of significance. In other words, libraries are concerned with the content of information. Google is only concerned with the content inasmuch as it is enables targeted advertising.

Despite the connotations of neutrality in the word ‘information’, information is not neutral. The criticisms relating to the integrity of search results provided clear evidence of this. While there have been breathless advocates of the superiority of the digital over the analogue (for example, Mitchell, 1995; Negroponte, 1995), infogration, or the translation of aspects of our world into information is not a neutral process either. The concerns expressed regarding Google’s book project point to this, especially the concerns relating to the form of the information. In contrast, public libraries aim to provide access to information in order to strengthen democracy. This requires a balanced flow of information and some sort of ordering of significance. In other words, libraries are concerned with the content of information. Google is only concerned with the content inasmuch as it is enables targeted advertising.

The criticisms relating to the integrity of search results provided clear evidence of this. While there have been breathless advocates of the superiority of the digital over the analogue (for example, Mitchell, 1995; Negroponte, 1995), infogration, or the translation of aspects of our world into information is not a neutral process either. The concerns expressed regarding Google’s book project point to this, especially the concerns relating to the form of the information. In contrast, public libraries aim to provide access to information in order to strengthen democracy. This requires a balanced flow of information and some sort of ordering of significance. In other words, libraries are concerned with the content of information. Google is only concerned with the content inasmuch as it is enables targeted advertising.

Conclusion: Negotiating a healthy relationship

This paper has characterised the association between Google and libraries as a relationship. It has described the romance phase when it seemed that both Google and libraries had the common aim of providing access to information. It has provided evidence of cracks appearing in this relationship and has argued that Google and libraries want very different things with regard to access to information. The conclusion will now outline the implications of the preceding analysis for libraries, and in doing so, makes explicit the consequences of conceptualising the association between Google and libraries as a romantic relationship. In the language of this metaphor, the conclusion of this paper is that a healthy relationship between libraries and Google requires libraries to regain a separate sense of self and set appropriate boundaries in their relationship with Google.

In negotiating their relationship with Google, whether it be deciding whether or not to continue to provide reference services, in a reference interview, in teaching people how to use Google, or in negotiating the digitization of the library’s physical collection, those in the library profession need to bear in mind that the word ‘information’ conceals irreconcilable differences between libraries and Google. Of course, it is not the suggestion of this paper that libraries sever ties with Google. That would be absurd, if not impossible. The suggestion here is that libraries should not rely on Google for information provision. Libraries should resist the Taiga Forum’s claim that all information discovery will begin with Google. They should teach library users, through example, about the difference between freely flowing information and balanced information. They should not be afraid of giving priority to more significant information. They should also be discussing the losses involved in representing an aspect of an analogue world with ones and zeros. As philosopher of technology Don Ihde says, it is “what is revealed is what excites; what is concealed may be forgotten.” Libraries need to pay attention to that which is concealed by Google’s search results and by digitised information. What is concealed includes vital aspects of human knowledge and culture and it is part of the task of the public library to preserve these things.

As publicly funded institutions, public libraries need to serve some conception of the public good. This requires that they be clear-sighted about the differences between what they seek to do and what the company Google does. Libraries need to reassert their identity as providers of balanced and significant information. Of course, this involves making subjective, value-laden judgements, the type of judgements librarians have always made in the past when selecting items for the collection. There should be endless and difficult debates about what is significant information for libraries, and what constitutes a balance. Without these debates, libraries will lose their bearings, swallowed up by Google (or Google’s successor) and spat out into a sea of too much information. They will take hopes of a healthy democracy with them.
About the author

Vivienne Waller is a Research Fellow at the Institute for Social Research of the Swinburne University of Technology.
E-mail: vwaller [at] swin [dot] edu [dot] au

Acknowledgements

The author would like to acknowledge the generous support of the State Library of Victoria and colleagues at the Institute for Social Research. This research was funded by Australian Research Council grant LP0775215. The author takes all responsibility for the views expressed in this article. The publishing of this article does not constitute an endorsement or any other expression of opinion by the State Library of Victoria.

Notes


9. Lakoff and Johnson, 2003, p. 158. Using a different metaphor, for example, characterising the association between Google and libraries as a competition, would highlight different aspects of the association and lead to an emphasis on different courses of action for libraries.


14. Microsoft had a similar restriction before it pulled out of its digitization project.

15. See also Albanese (2008) and Fialkoff (2008), for a range of responses from librarians to Google’s agreement with book publishers and authors regarding access to books scanned in the Google Book Project.


21. A search by the author in April 2008 brought this up as the second result.


24. As many commentators have noted, this paper was written in 1997, while Google was still a non–for—profit venture at Stanford University.


27. James Grimmelmann, 2009, has an excellent discussion of this.


32. In South Korea, human volunteers, rather than an algorithm, provide answers to user queries.


42. See https://www.23andme.com/, accessed 4 June 2008.


References


**Editorial history**

Paper received 7 April 2009; accepted 20 August 2009.

"The relationship between public libraries and Google: Too much information" by Vivienne Waller is licensed under a [Creative Commons Attribution–Non–Commercial–No Derivative Works 2.5 Australia License](http://creativecommons.org/licenses/by-nc-nd/2.5/au/).

The relationship between public libraries and Google: Too much information by Vivienne Waller.

*First Monday*, Volume 14, Number 9 - 7 September 2009

http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2477/2279