# Transparency and the Ubiquity of Information Filtration?

Tama Leaver, Michele Willson & Mark Balnaves

In past decades, the notion of information filtering was primarily associated with censorship and repressive, non-democratic countries and regimes. However, in the twenty-first century, filtration has become a widespread and increasingly normalised part of daily life. From email filters—designating some messages important, some less important, and others not worth reading at all (spam)—to social networks—with Facebook and Twitter harnessing social ties to curate, sort and share media—through to the biggest filtering agents, the search engines—whose self-professed aims include sorting, and thus implicitly filtering, *all* our information—filters are inescapable in a digital culture. However, as filtering becomes ubiquitous and normalised, are citizens *en masse* becoming too accepting or, worse, largely ignorant, of the power these filters hold?

Filters are widely accepted as a necessary part of managing the supposed information overload of a digital economy, but do we need to be more critical and more aware of filtering both in general and in specific instances? This paper explores some of the issues that are raised by the increasing prevalence and indeed necessity of filtering and organising information in digital arenas where human action and technological affordances are increasingly conflated. While there are vast numbers of filtering processes that could be analysed, by limiting the examination below to certain sites where informatics and social filtering intersect—around the issues of Net Neutrality, the whistle-blower website *WikiLeaks*, the operation of search engines and the increasing role of social search facilitated by online social networks, and social gaming—we emphasise the need for a more complex and nuanced situating of the processes of information filtering and the social, political and cultural impact of various filtering approaches. Moreover, rather than vilifying or celebrating information filtering *per se*, we argue for more transparency in filtration processes and practices, in an effort to make filtration both visible and understandable to everyday users.

## *Understandings of Filtering*

Filters in common parlance are understood as a mechanism or technique for managing, blocking or controlling the behaviours or properties of something. Filters allow some properties or things through and block others. They are variously seen as ensuring the purity of something by blocking out something undesirable. Coffee filters allow water to pass through ground coffee without allowing the grinds themselves to pass through—thus ensuring the coffee is of a higher quality and more easily consumed. Air or water filters likewise prevent ‘contaminants’ from entering the filtered environment, ensuring a degree of purity or a desired state of being. Thus filtering can be seen simultaneously as a repressive or discriminating function that precludes access, as a mechanism that requires identification of a desired state, and as a useful technique for enhancing desirable outcomes.

Information filtering itself is often positioned as a subset of information management and retrieval practices (Hanani, et al., ‘Information Filtering’; Belkin and Croft, ‘Information Filtering’). It is understood as a way of regularly managing information flows according to the desires of users. Thus email filters can be set to block email that the system has been programmed to recognise as spam or undesirable. Likewise many systems can be customised to allow individual users to finesse the filtering process; to block out some users, for example, from particular ‘locations’.

There is a vast literature that situates information filtering practices using technological means and, increasingly, digital media platforms. Hanani, et al. differentiate several types of filtering according to four different parameters:

> i) *where filtering is initiated*, being either active, with information being sought out, or passive, omitting or blocking information received;
>
> ii) *the location of the operation*, for example, the source of information, via filtering servers, at the user’s site, and so forth;
>
> iii) the *approach to filtering*, distinguishing between cognitive and social filtering; and
>
> iv) the *method of acquiring knowledge about users*, distinguishing between explicit, implicit and combined methods. (‘Information Filtering’, 206)

This type of definitional framework is useful in terms of taking what might otherwise be identified as monolithic processes of filtration and breaking them into discrete, variable components or dimensions. However this is only one step in understanding the political and social import of filtering. For example, this type of framework does not interrogate why information is being filtered, what the premises or intent of filtering exercises are, or who benefits and who loses (if such a clear distinction can be made) from specific systems of filtration, let alone from the inevitability of filtration as such. Instead, it assumes that information is a tangible material item that can be manipulated, distributed, contained and controlled according to various degrees of efficiency for the end user. We suggest, by contrast, that given the ubiquity, normalisation and often invisibility of information filtration, close discussion is required, with the aim of making filtration practices, and the decisions driving these practices, far more visible than is currently the case.

## *The Traffic-Ramp Control*

While the Internet is global in reach and infrastructure, its US origins often mean that US regulations and practices set the stage for international information practices. For example, in the US, Comcast and other network carriers sometimes delay specific types of traffic; the most notable being video, music and other files shared using peer-to-peer software like BitTorrent. Internet service providers compare these practices to a traffic-ramp control mechanism that regulates the entry of additional vehicles on to a freeway at peak times. The argument from network carriers is that this is not preventing people from using the Internet but temporarily delaying their access to particular services and particular types of information that are perhaps deemed as less desirable.

However, the common term for this practice—throttling—is indicative of the impact such practices can have. One of the problematic presumptions, for example, is that a significant proportion of peer-to-peer file-sharing involves the unauthorised distribution of material under copyright. Even if this is true, there are also perfectly legitimate and authorised uses of peer-to-peer file-sharing; and unless proven in a court of law, Internet service providers cannot determine the legal status of particular files, or types of traffic, on their infrastructure. Moreover, at present the application of throttling could be applied to any type of file or data from any website or service; hence concerns about commercial or political preferences and deals being struck.

These types of actions are a part of the discussion in what is now called the Net Neutrality debate but they also concern the issue of filtration. They raise questions about how companies that run the services on the Internet might act not only to restrict Internet traffic, but thereby to change the nature of the communicative spaces themselves. What does it mean for the Internet to be ‘open’? The US Federal Communications Commission (FCC) proposed a set of principles in 2005 that it thought should govern the Internet, in order to ‘encourage broadband deployment and preserve and promote the open and interconnected nature of the public Internet’:

- consumers are entitled to access the lawful Internet content of their choice.
- consumers are entitled to run applications and use services of their choice, subject to the needs of law enforcement.
- consumers are entitled to connect their choice of legal devices that do not harm the network.
- consumers are entitled to competition among network providers, application and service providers, and content providers. (Policy Statement)

The FCC said that these guidelines were not ‘enforceable’ but, in 2007, issued a Notice of Inquiry seeking examples of the beneficial or harmful actions of network platform providers: whether they had biased particular content and traffic, how users might be affected by these biases, and whether user choice of network providers was sufficient to ensure overall equity (Inquiry into Broadband). Comcast’s restriction of BitTorrent transfers was one of the obvious targets of this inquiry. By 2008 a three-judge panel in Washington rejected an FCC August 2008 cease and desist order against Comcast. The judges decided that the FCC failed to tie its actions to an actual law enacted by Congress; that is to say, it did not have the power to regulate network management practices (McCulagh, ‘Court’).

At first glance, large Internet companies supported the principles of Net Neutrality and an FCC role in policing it. The US Congress in 2006 rejected five bills backed by Google, Amazon.com, Free Press, and Public Knowledge that gave the FCC enforcement powers. But a closer examination of company actions reveals a somewhat different picture. Google is a good example of the twists and turns in the Net Neutrality saga.

In 2007 the FCC licensed wireless spectrum with binding rules that forced any wireless carrier that had won a spectrum auction to let users employ whatever handsets, services and applications they wanted to connect to it. Verizon, which had been bidding against Google for Android and eventually won, filed a lawsuit against the FCC rules. Google opposed Verizon’s rejection of the FCC rules. In 2010, however, Google and Verizon, as part of their bilateral Net Neutrality trade agreement, argued that Congress should ratify the notion that open wireless rules are unnecessary (Singel, 2010). Google defended its seeming reversal:

> We have taken a backseat to no one in our support for an open Internet. We offered this proposal in the spirit of compromise. Others might have done it differently, but we think locking in key enforceable protections for consumers is progress and preferable to no protection. (Ctd in Singel, ‘Google’)

However, if one compares this statement to that from a post on Google's official blog in 2007, at which time the Internet giant supported Net Neutrality for wireless, then a telling contrast is visible:

> The nation's spectrum airwaves are not the birthright of any one company. They are a unique and valuable public resource that belong to all Americans. The FCC's auction rules are designed to allow U.S. consumers—for the first time—to use their handsets with any network they desire, and use the lawful software applications of their choice. (Sacca, 'Consumer Choice')

Whatever the regulatory outcome in the Net Neutrality arena, it is clear that the key players are fully aware that important distortions can occur within Internet 'open space'. Indeed, both sides of the debate tend to position themselves as champions of openness, with pro-Net Neutrality supporters arguing that legal protections are needed to ensure that all Internet traffic remains, while those against enshrining Net Neutrality in law often argue that the free market is the best guarantor of open access (Sridhar and Goldman, 'Net Neutrality').

The Net Neutrality debate thus raises important issues about the openness of the Internet, and about the definitional slipperiness of 'openness' as a term. Regardless of the terminology, though, Comcast's restriction of certain types of Internet traffic at certain times is a clear and visible example of information filtering. Moreover, the Net Neutrality debate highlights one of the nuances of filtering: even delaying rather than just blocking information can be as effective as preventing information flow altogether. Here social desires (the need to achieve close to instantaneous information exchange) and the algorithmic logic of information control (as exemplified by the blocking or throttling software) have a deep but strained relationship, something which is indicative of information filtering today. The issues around Net Neutrality highlight the important point that information filtration is not simply a case of blocked versus unblocked, but issues of speed, quantity and quality of information flow are similarly significant.

## *Beyond Agenda Setting*

While news media have long provided particular perspectives, information filtration needs to be distinguished from agenda setting, where a set group summarises information or events and returns them to readers. Journalists, for example, have an impact on issues as agenda-setters. Journalists can focus on news stories that attract people's attention. 'Framing' and 'priming' are an important part of agenda setting. Frames call attention to some aspects of reality rather than others. They make interpretation possible through a journalist's use of syntax, themes, script and rhetoric. Priming in conjunction with framing is a psychological process where news emphasis on specific issues or events activates other memories associated with those issues or events. For instance, media references to 'terrorists', depending on how the story is framed, might elicit in people's minds images of Guantanamo Bay or the Taliban. Cognitive priming linked with agenda-setting research enables us to look at how the media influence a person's perception of the importance of issues or events. As Jospeh N. Capella and Kathleen Hall Jamieson argue:

> News stories, even those strategically framed, often carry substantive information about issues, albeit set in the context of self-interested manipulation. Attentive exposure can alter political knowledge by increasing the accessibility of information, changing the associations among the constructs and cuing and strengthening existing localized networks of concepts. (*Spiral*, 60)

It is not that there is an over-arching conspiracy in the process of framing and priming, although of course in areas like perception management and psychological operations (the modern versions of grey and black propaganda), there may well be. In any democratic society there will be a series of 'labellings' and discourses that become well-known and commonplace. What makes a democracy a democracy and the fourth estate the fourth estate is the capacity to recognise and to correct serious wrongs that emerge or have emerged over time (Bennett and Paletz, *Taken*; Iyengar and Simon, 'News Coverage').

What is different with filtration as opposed to agenda-setting is the involvement and role of machinic choices in the aggregation of opinion, and in the sheer diversity of actual opinion from a massive array of sources, many or most of whom are not journalists. While most viewers are aware on some level that journalists bring a perspective to the news they report, often there is a presumption that algorithmic information filtration is objective since immediate human action might not be required, although this myth entirely ignores the human values at play in the construction and programming of the algorithms themselves. Rather than journalists consciously or subconsciously framing and priming, and a broadcast system that creates and transmits the stories and images to individuals, the Internet can produce the impression of a large-scale following of agendas in real time, something never possible under traditional broadcasting. Determining whether or not that following is genuine is a problem because fake traffic aggregators can easily inflate figures on the number of followers of a story. But there are additional layers to this, ranging from attribution of the source—astro-turfing—through to the role of 'super-aggregators'.

## *Leaking Filters?*

While journalists may once have been acknowledged and celebrated as the fourth estate, the profession as a whole has lost some credibility over the past decade, not least due to the centralisation of media ownership, the massive increase in sensationalism and the dominance of commercially-driven stories. In this setting, many alternative approaches to news reporting and significant information sharing have emerged, the most controversial of these being *WikiLeaks*. At the most basic level, *WikiLeaks* positions itself as an anti-filter, a team and a website that will publish and host information that has been censored, banned or otherwise filtered elsewhere. For example, when the Australian federal government was planning to implement a mandatory national-level internet filter, *WikiLeaks* revealed the initial 'blacklist' of websites that would have been blocked, clearly against the wishes (and, at face value, the laws) of the Australian government (Lynch, 'CRACK THE WORLD').[1] The enigmatic and controversial head of *WikiLeaks*, Julian Assange, has a background in information technology and activism; so it comes as no surprise that the whistle-blower website employs state-of-the-art encryption technologies to ensure the anonymity of online sources and informants (ibid.). Indeed, one of the main safeguards for *WikiLeaks*' very existence lies in the organisation's employment of the best possible measures of information protection .

The relationship between journalists and *WikiLeaks* has been a tense one: on one hand, *WikiLeaks* is a fantastic source of material for stories; on the other hand, because the service is hosted in Sweden, it circumvents many of the regulations of journalistic ethics and media laws to which journalists operating outside Sweden are, ostensibly, required to adhere. In this sense, journalism may appear at face value to be part of the mechanisms of government control and filtration, while *WikiLeaks* valiantly resists the logic of the filter. However, while the mythology around *WikiLeaks*—a mythology driven, at least partly, by Assange—champions its independence and uniqueness, two major leaks, containing hundreds of thousands of military documents about the military operations in Afghanistan and Iraq have changed the way even *WikiLeaks* must filter information. After the first leak of some 70,000 documents about military operations in Afghanistan, *WikiLeaks* was widely criticised for releasing information that could not only compromise military activities, but which could also lead directly to identification of informants and other significant personnel, inviting Taliban retribution. *WikiLeaks* argued, however, that the documents had only been released in full after they were already available due to a journalist accidentally publishing the password to the encrypted but publicly available file (see Parry, 'Guardian'). In contrast, when the subsequent 400,000 documents relating to the Iraqi occupation were released, a significant number of names and other pieces of information had been removed by *WikiLeaks* in an attempt to minimise any harm the documents could cause for people who would otherwise have been identified (Shaughnessy, ‘WikiLeaks’). Despite *WikiLeaks* being in ardent opposition to filtering *per se*, the necessity of applying filters to these military documents highlights the social good arguably served by some forms of filtering.

One of the other significant changes in light of the Afghanistan and Iraq document leaks in 2010 was that *WikiLeaks* had to rely on some of the traditional news filters, the journalists, in order to make sense of the massive amounts of information they had acquired. In both instances, a select group of newspapers, including *The Guardian* in the UK and *The New York Times* in the US, were provided with the leaked documents in advance of the public release. While this pre-release may seem odd, the sheer size and number of documents leaked to *WikiLeaks* made this a necessity—there was simply no way a small organisation like *WikiLeaks* could make sense of the material in their hands without the assistance of trained professionals. Predictably, amongst the many revelations coming from the Iraq documents, the news media also took the time to highlight their own continuing importance. In an article entitled ‘A Renegade Site, Now Working With the News Media’, by Noam Cohen, *The New York Times* celebrated the fact that *WikiLeaks* had seen the wisdom of making the most of the filtering abilities of professional news organisations. Equally, there were news stories asking whether *WikiLeaks* had ‘sold out’, now that they were ostensibly respecting some of the filtering directives of the US government (Bercovici, ‘Growing Up’). Thus, even for an organisation which sees itself as antithetical to filtering, *WikiLeaks*, too, relies on a complex combination of informatic and social filtering systems to effectively make sense of the deluge of information available in a digital culture.

## *Don’t Filter Evil?*

While *WikiLeaks* is a visibly, politically influential website, no online service or platform matches the ubiquity of the corporations behind Internet search engines. Google is the most popular search engine, and the company is notable for its oft-quoted corporate motto, ‘Don’t Be Evil’. However, a global company servicing the search needs of potentially billions of users has many different notions of evil to contend with. In this context, the question of what Google will filter is an important one—and an issue that dramatically raised its head when it was discovered that the first link returned by Google for the search ‘Jew’ was an anti-Semitic hate site. Understandably, in response, there were widespread calls for Google to remove the offensive result (Daniels, *Cyber Racisim*, 170). Google argued that ‘they’ were not responsible for the results returned, but rather an automatic software algorithm was responsible. Rather than going down the ‘slippery slope of interfering with the results of a given search query’ (Halavais, *Search Engine*, 122), no matter how offensive the results, Google stood firm, refusing to modify its algorithm. Only after concerted protests did Google add a warning that was returned with the offensive results, pointing to Google’s own explanation of how the results were determined and why this was not a sign of Google’s support in any way. Nevertheless, Google held fast to its decision not to filter these specific results.

Google’s decision may appear at first glance clear and even noble in its attempt to objectively allow a software algorithm to accurately gauge the popularity of specific results. However, as Alexander Halavais has noted in his *Search Engine Society*, while the US results were unchanged, Google removed the offensive results delivered by its searches inside France and Germany. In these countries, hate speech is illegal and thus returning the anti-Semitic result would have broken national laws and opened up Google to legal action. What is notable here is not the different national legal systems—a continual source of confusion for online companies operating globally—but the speed and silence with which Google did actually filter these same results in certain countries. Apparently ‘Don’t be Evil’ has an invisible extension: ‘Don’t Be Evil According to the Laws of Whatever Country Requests Results’. Similarly, while no longer the case for the most part, for more than a year Google agreed to abide by China’s censorship laws in order to operate inside their national boundaries, on the inside of the Great Firewall. During that period, searching for the Dali Lama, for example, or for the Tiananmen Square massacre, returned no results at all, or a sanitised list, empty of the politically sensitive unfiltered material (Hinman, ‘Searching Ethics’). Google has since altered its approach to China, and to filtering results within China, but these nation-specific examples highlight the ease with which Google can and will filter search results if legally required. For Google, too, filtering is not an absolute evil, but a relative one. While these high-profile examples stand out, they do beg the uncomfortable question about what other results may be filtered if national governments, or others, were to demand it.

## *Knowledge Aggregators*

If the case of Google's geographically variable search results highlights the extent to which search requests on politically sensitive issues may be compromised by acts of 'censorship', information filtration is in fact central to the very processes of knowledge aggregation that underpin even the most banal online searches. Knowledge aggregation takes many forms on the Internet. The term *news aggregator*, for example, describes websites or search engines that select, retrieve and link to news items from anywhere on the Internet. *Google News* is an aggregator of this kind and has had its own troubles from news providers who think that their content is being taken without due financial recompense. Moreover, *Google News* is a personalisable aggregator, in the sense that each individual can configure the service, adding a layer of social preferences to the software routines filtering the news. An RSS reader is also an aggregator as it pulls together threads relevant to the person who has subscribed to it.

But while news aggregation is easily seen (objections from traditional media content producers notwithstanding) as a social or public 'good' by virtue of its accommodation of Internet users as seekers of knowledge about current affairs, there are other forms of knowledge aggregation that are far less visible in terms both of the processes and the targets of data collection. User tracking, for instance, is a process in which knowledge about users *themselves* is aggregated, and routinely without any knowledge or consent on the part of the individual user—as *Telegraph* journalist Rowena Mason discovered when she visited the office of data brokering firm Acxiom and spoke to the corporation's chief executive at that time:

> 'Oh we do have you on our database. I guarantee you,' Mr Meyer assures me. 'Your name address, phone number. You have a cat. You're right handed. That sort of thing.'
>
> This is true. I'm not sure if it's a lucky guess, but I'm impressed.
>
> Mr Meyer, a brash, confident chief executive, explains that while the company has been nervous of promoting its activities in the past, he has no fear of a higher profile.
>
> 'We're the biggest company you've never heard of,' he grins, with a hint of Southern drawl. 'In the past we were afraid of people knowing us, but I'm trying to get business awareness and if consumers have privacy concerns I want to know.' (Mason, 'Acxiom')

It is not surprising that Mason had no idea how much Acxiom could know about her, since many people have never heard of the corporation. Most people, though, have flash cookies hidden on their computers. These data files—which are mostly undetectable, except now for Firefox browsers—collect all kinds of user data for aggregators like Acxiom, from personal identification details to media use, browsing history, and other online movements. Working behind the scenes for Google and many other major Internet companies, Acxiom had accumulated, according to Mason, approximately 1,500 facts about half a billion people worldwide as of 2009.

Giant aggregators like Acxiom argue that they are neutral in terms of their effects on the relative 'openness' of the Internet, where 'openness' is understood in terms of user access to services and information. Others nevertheless disagree, primarily because of the ways in which mass opinion is re-presented back to the Internet and traditional media:

> the opinions crafted by individuals (presumably after or through discourse in a small community) can be aggregated and passed to other users and communities for further discussion and subsequent aggregation. Such sites cut out the human mediation traditionally required in a social network, allowing for a seemingly direct representation of public opinion in the blogosphere. Through a system of uncoordinated coordination, collective action has become possible on a previously unimaginable scale, due to the small amount of effort required by each human in order to bring about the so-called 'wisdom of crowds'. (Geiger, 'Habermas')

How representation of public opinion might be distorted by the way data is owned, collected and distributed by aggregators is the flip-side, in other words, of the concerns expressed above about Comcast playing with traffic flow. In this sense, the personal customization that can occur on the basis of information collected by Acxiom is simultaneously a form of filtering, whether explicitly authorised by individuals or invisibly enacted by websites and platforms.

But it is precisely this question of the visibility or otherwise of such data collection and filtration processes that broadens—or ought to broaden—the scope of investigations into the 'openness' of the Net to include, alongside concerns about 'censorship' (whether politically or technologically 'motivated'), the question of transparency. For instance, one dimension of Google's information filtering operation that is often banal enough to go unnoticed is the pervasive practice of search personalisation. Since December 2009, regardless of whether a user is logged into Google or not, the search engine uses a range of available informatic cues—from the type of browser and operating system being used, to the user's physical location—to tailor each and every search result (Pariser, *Filter*). While the customisation of each search result may be subtle, at times the power of 'personalisation' can lead to wildly different results for the exact same search, questioning the whole notion of algorithmic objectivity which Google often uses to defend its search results. Eli Pariser argues in *The Filter Bubble* that search personalisation is the tip of the iceberg, and that user-specific filtered results are increasingly becoming the norm for many online corporations and services.

What is most concerning, though, is not personalisation in and of itself, but that it is inescapable—it is not possible to get non-personalised results from Google, for example—and that it is often invisible. Invisibility matters, argues Pariser, since a lack of awareness of the way search results are filtered often means that insufficient critical attention is paid to the sort of results Google provides. Pariser goes on to argue that 'democracy requires citizens to see things from one another's point of view, but instead we're more and more enclosed in our own bubbles. Democracy requires a reliance on shared facts; instead we're being offered parallel but separate universes' (*Filter*, 5). While Pariser's perspective might seem alarmist, two of the products that Google has announced it is working on in 2012—*Google Now* and *Project Glass*—illustrate more extreme scenarios where this personalisation may become pernicious. *Google Now*, a service built into Google's Android mobile operating system, utilises everything Google knows about a user to provide results without any search activity at all; Google predicts exactly what a user will want to know given everything Google already knows about that person. In parallel, *Project Glass* is Google's attempt to build Google-powered glasses, giving users, effectively, a display only visible to that user, built into a pair of wearable, mobile computing, glasses. Looking at these two services in light of Pariser's work on personalisation and filtering, Google may soon be returning results that only the specific user can see, with no searching required, based entirely on the algorithmic filtration of the information generated by past search and communication history. The name *Project Glass* thus appears deeply ironic, for while the glasses themselves may appear transparent, the information filtration driving the material displayed on them will be largely invisible.

## *Social Filtering*

Where popular conceptions of search engine filtration are often dominated by notions of 'algorithmic objectivity' at the expense of any scrutiny of its social conditions, the reverse is perhaps more often the case with regard to social networking and peer-recommendation systems. Social network sharing of information amounts to a form of filtering to the extent that access both to information and to people is shaped by recommendations from peers. Social filtering was recognised as important long before the advent of Facebook and other social network sites—see, for example, Mark S. Granovetter's work on weak social ties ('Strength'). Recommender systems such as those used by Amazon, reputations systems such as those used by eBay and other types of social network filtering simply instantiate, therefore, extistent social practices in a technological form through technical means, such that the action of social filtering is 'delegated' to a combination of machinic and human processes (Latour, 'Missing Masses').

However, the point where online sharing might diverge from past, 'pre-digital' social filtering practices lies in the coincidence of 'other' functions or affordances with this (heightened) attention to collaborative filtering. By building social filtering principles into a technical construct—in the form of a recommender algorithm, say, or a social networking site like Facebook—a different relationship is orchestrated between the data entered and the actions of the various 'participants' (who include the people, the site, the algorithms that drive the filtering, etc). This is to say that, in the context of online systems:

> i) the filtering or recommendation becomes more explicit, formalised and easily accessible than it may have been previously;
>
> ii) it increases personal relevance in information retrieval practices inasmuch as people are able to inform their choices through trust in the system of recommendations—and, it has been noted in the literature, that there is a strong degree of trust placed in peer recommendations (see Walther, et al., 'Interaction', and below);
>
> iii) it relies on 'black-boxed' code or algorithms, which is to say that users are provided with little knowledge of the information that is filtered out, let alone any insight into the logic informing either the code's production or the organisational or commercial objectives that are achieved through the management of data; and
>
> iv) it enables other invisible, unrequested actions to take place simultaneously, most notably the aforementioned work of data mining.

This last point is related to the third inasmuch as users may well be completely unaware of the other ends for which their information, their choices and so forth are being captured, manipulated and repurposed—a point summed up by the 'exploitative' dimension of social networks:

> Social networks register a 'refusal of work'. But our net-time, after all, is another kind of labour. Herein lies the perversity of social networks: however radical they may be, they will always be data-mined. They are designed to be exploited. Refusal of work becomes just another form of making a buck that you never see. (Ippolita, et al., 'Digital Given')

Perhaps one of the clearest instances where data mining and information filtering are applied invisibly in a context which occludes these practices is the area of social gaming. Social games, which are games played via social networking platforms such as Facebook, are increasingly popular, with hundreds of millions of cumulative players. Within the world of social gaming, the most well-known company is Zynga, producer of *Farmville* and *Mafia Wars*, among other well-known games. During a

2010 presentation, Vice President of Analytics and Platform Technologies Ken Rudin argued that, at its core, Zynga is a company that succeeds by carefully filtering, mining and analysing user data, with the aim of establishing ways of predicting and influencing user behaviour, social interaction between users, and the likelihood of users buying virtual goods for use within the gameworlds ('Analytics'). Rudin stressed that information filtering and data mining were not specialist areas within Zynga, but rather deeply integrated into all product and platform development, and that the resulting algorithms were the core business of Zynga and other successful social gaming companies. Through careful analysis of user data, Rudin argued, Zynga had managed to strategically cultivate user acceptance of virtual goods as central to the game-playing experience, with the result that, despite advertising being the standard revenue-raiser for most online platforms, Zynga made 95% of its profit from the sales of virtual goods. By the end of 2011, Zynga reported that user activity from its range of games was creating over 15 Terabytes of mineable data a day (Takahashi, 'Zynga'). In this regard, it is worth noting that Rudin left Zynga in 2012 to take up the position of Head of Analytics at Facebook, a fact emphasising the extent to which the so-called 'social graphs'[2] underpinning the world's largest social networking service and the world's largest social gaming company, are less informatic representations of pre-existing personal relations than performative products of the incredibly refined data filtration and analytic processes used to drive the sales of advertising and virtual goods, usually in the guise of improving users' social experiences. As Rudin bluntly stated, Zynga sees itself as 'an analytics company masquerading as a games company' ('Analytics'). And it is a very successful masquerade, given that few *Farmville* players, for instance, are likely to consider how their every movement and mouse-click is being algorithmically studied and filtered to increase the odds that they will part with real cash for virtual goods in future gameplay.

## *Conclusion*

> We are addicted to ghettoes, and in so doing refuse the antagonism of 'the political'. Where is the enemy? Not on Facebook, where you can only have 'friends'. What Web 2.0 lacks is the technique of antagonistic linkage. Instead, we are confronted with the Tyranny of Positive Energy. Life only consists of uplifting experiences. Depression is not a design principle. Wikipedia's reliance on 'good faith' and its policing of protocols quite frequently make for a depressing experience in the face of an absence of singular style. There ain't no 'neutral point of view'. This software design principle merely reproduces the One Belief System. Formats need to be transformed if they are going to accommodate the plurality of expression of networked life. Templates function as zones of exclusion. (Ippolita, et al., 'Digital Given')

Filtering overlaps with censorship—a politically loaded term—but in many ways is more complex than the notion of censorship suggests. As Ippolita, et al. suggest, filtering processes of the type addressed here ultimately enforce a particular type of world view, reducing the scope for disagreement, discussion, conflict and diversity of opinion. This is not straightforwardly 'a bad thing', since there are clearly some ideas, images and practices that societies reject as unconscionable and whose ready accessibility would cause anxiety even for the most 'liberal' or 'tolerant' of citizens. However, questions of what to filter, when, how, and by whom are clearly contentious. And it is precisely the lack of transparency, we argue—in terms both of which information filtration practices are enacted and of how these practices influence everyday communication and knowledge—that is of most concern. The continually expanding blacklist of sites banned for identifying blacklisted sites looks like a potentially slippery slope to navigate. Blocking a URL nationally presents at one level, that is, a particular understanding of the nation-citizen relationship, which may well require further negotiation. At another level, however, such filtration may function to block access only for certain types of people —the technologically 'unsavvy' perhaps, those people who do not have the technical know-how to circumvent the filter. And in that event, what is ostensibly implemented as a system for filtering *content* turns out to operate instead as a system for filtering *users*.

While information filtering is undoubtedly an essential practice—there being so much information ‘out there’ that effective and meaningful access to it would be largely impossible in the absence of certain strategies and mechanisms for mapping, assessing and navigating through the sea of data—the specific forms and practices of filtering ought themselves to be able to withstand critique, scrutiny and evaluation. The challenges facing any such critique are to be found in the increasing ubiquity of information filtering practices and in the increasing *opacity* of online filtration systems. Such systems are not reducible to simple regimes of censorship but rather constitute complex aggregates of individual actions, organisational objectives and technological mechanisms—most significantly, ostensibly ‘neutral’ algorithms whose workings are both invisible and unfathomable to many. What is filtered out by such systems and mechanisms can tell us as much about the particular moral, political or social priorities underpinning them as what is ‘allowed’ through the filter. Whether users are in a position to tell what has and has not been filtered out is thus a key consideration. In the twenty-first century media landscape, information filtration is already ubiquitous. In the face of this predominance, the challenge is to ensure that, in the further development of such filtration systems, the principle of transparency attains ubiquity as well.

## Notes

Research on this paper has been supported by the Australian Research Council (ARC) linkage grant ‘Online Money and Fantasy Games—an applied ethnographic study into the new entrepreneurial communities and their underlying designs’ (LP110200026). An earlier version of this article was presented at the 2011 Australian and New Zealand Communication Association (ANZCA) conference (Leaver, et al., ‘Ubiquity’). The authors wish to thank the reviewers of that earlier paper, acknowledge useful comments received during the conference itself, and also thank the *Ctrl-Z* reviewers for their useful insights in improving this paper.

1. 2009 saw an extensive debate about censorship and filtering in Australia. This debate centred on the Australian Government’s proposal to introduce compulsory filtering of Internet content in Australia through ISPs. ISPs were to be required to ‘block’ the transmission of particular URLs noted on a blacklist drawn up by the Australian Communications and Media Authority (ACMA). Under the proposal, individuals and organisations would also be fined for adding hyperlinks to banned sites. Several pages of *WikiLeaks* were added to this blacklist, according to Morris (2009), after *WikiLeaks* leaked a list of Danish banned websites (ironically, the ACMA blacklist was itself leaked to the media, circumventing its own internal filtering processes). Questions were raised about how the decisions were made as to what was included on the list and what was excluded, who made these decisions and implicitly, what world-view and moral and political premises would inform these decisions. #back

2. Broadly speaking, a social graph refers to a visualisation of an individual’s social network. However, the term is now primarily associated with the social networking company Facebook who have deployed the term to mean the digitally linked social networks revealed by online social network connections. More to the point, Facebook commercially exploits this social graph, selling customers access to its Open Graph software, a means of targeting advertising material and other applications at individuals with specific types of connections in the Facebook social graphs. #back

## References

Andrejevic, Mark. ‘Social Network Exploitation’, in *A Networked Self: Identity, Community, and Culture on Social Network Sites*, ed. Zizi Papacharissi. New York and Abingdon, UK: Routledge: 2011, pp. 82-102.

Belkin, Nicholas J. and W. Bruce Croft. ‘Information Filtering and Information Retrieval: Two Sides of the Same Coin?’ *Communications of the ACM* 35, 12 (1992): 29-38.

Bennett, W. Lance and David L. Paletz, eds *Taken By Storm: The Media, Public Opinion, and U.S. Foreign Policy in the Gulf War*. Chicago: University of Chicago Press, 1994.

Bercovici, Jeff. 'In Growing Up, Did Wikileaks Also Sell Out?' *Forbes*, 25 October 2010.

Capella, Joseph N. and Kathleen Hall Jamieson. *Spiral of Cynicism: The Press and The Public Good*. Oxford: Oxford University Press, 1997.

Cohen, Noam. 'A Renegade Site, Now Working With the News Media'. *The New York Times*, 1 August 2010.

Daniels, Jessie. *Cyber Racism: White Supremacy Online and the New Attack on Civil Rights*. Lanham, Md: Rowman & Littlefield, 2009.

Ellison, Nicole B.; Cliff Lampe, Charles Steinfeld and Jessica Vitak. 'With a Little Help From My Friends: How Social Network Sites Affect Social Capital Processes', in *A Networked Self: Identity, Community, and Culture on Social Network Sites*, ed. Zizi Papacharissi. New York and Abingdon, UK: Routledge: 2011, pp.124-45.

Federal Communications Commission (US). FCC Releases Policy Statement on Broadband Internet Access (FCC 07-151) (2005).

Federal Communications Commission (US). FCC Launches Inquiry into Broadband Market Practices (FCC 07-31) (2007).

Galloway, Alexander R. 'What is New Media?: Ten Years after the Language of New Media'. *Criticism* 53, 3 (2011): 377-84.

Geiger, Stuart R. (2009). 'Does Habermas Understand the Internet? The Algorithmic Construction of the Blogo/Public Sphere'. *Gnovis* 10, 1 (2009).

Granovetter, Mark S. 'The Strength of Weak Ties'. *American Journal of Sociology* 78, 6 (1973): 1360-80.

Halavais, Alexander. *Search Engine Society*. Cambridge: Polity Press, 2008.

Hanani, Uri.; Bracha Shapira and Peretz Shoval. 'Information Filtering: Overview of Issues, Research and Systems'. *User Modeling and User-Adapted Interaction* 11 (2001): 203-59.

Hinman, L. M. 'Searching Ethics: The Role of Search Engines in the Construction and Distribution of Knowledge'. *Web Search*, *Information Science and Knowledge Management* 14 (2008): 67-76.

Ippolita; Geert Lovink and Ned Rossiter. 'The Digital Given: 10 Web 2.0 Theses'. *Fibreculture* 14 (2009).

Iyengar, Shanto and Adam Simon. 'News Coverage of the Gulf Crisis and Public Opinion; a Study of Agenda-Setting, Priming, and Framing'. *Communication Research* 20, 3 (1993): 365-83.

Latour, Bruno. 'Where Are the Missing Masses?', in *Shaping Technology/Building Society: Studies in Sociotechnical Change*, eds Wiebe E. Bijker and John Law. Cambridge, MA: MIT Press, 1992, pp. 225-58

Leaver, Tama; Mark Balnaves and Michele Willson. 'The Ubiquity of Information Filtration', in Refereed Proceedings of the Australian and New Zealand Communication Association Conference: Communication On The Edge, ed. Alison Henderson. Hamilton, NZ: ANZCA, 2011.

Lynch, Lisa. '"WE'RE GOING TO CRACK THE WORLD OPEN": Wikileaks and the Future of Investigative Reporting'. *Journalism Practice* 4, 3 (2010): 309-18.

Manovich, Lev. *The Language of New Media*. Cambridge, MA and London, UK: MIT Press, 2001.

Mason, Rowena. 'Acxiom: The Company That Knows If You Own a Cat Or If You're Right-Handed'. *The Telegraph*, 27 April 2009.

McCullagh, Declan. 'Court: FCC Has No power to Regulate Net Neutrality'. *C/Net*, 6 April 2010.

Moses, Asher. 'Banned Hyperlinks Could Cost You $11,000 a Day'. *The Sydney Morning Herald*, 17 March 2009.

Pariser, Eli. *The Filter Bubble: What the Internet Is Hiding from You*. London: Viking, 2011.

Parry, Nigel. (2011, August 31). 'Guardian Investigative Editor David Leigh Publishes Top Secret Cablegate Password Revealing Names of U.S. Collaborators and Informants … in His Book'. *NigelParry.com: The Website Less Traveled*, 31 August 2011.

Rudin, Ken. 'Actionable Analytics at Zynga: Leveraging Big Data to Make Online Games More Fun and Social'. Presented at the 2010 TDWI BI Executive Summit, San Diego, CA, 2010.

Sacca, Chris. 'Consumer Choice is Always the Right Answer'. *Google Public Policy Blog*, 13 September 2007.

Shaughnessy, Larry. 'WikiLeaks Redacted More Information in Latest Documents Release'. *CNN*, 23 October 2010.

Singel, Ryan. 'Why Google Became A Carrier-Humping, Net Neutrality Surrender Monkey'. *Wired*, 10 August 2010.

Sridhar, Aparna and Jon Goldman. 'Give Us Net Neutrality or Give Us Death'. *Bloomberg Businessweek*, 16 September 2010.

Takahashi, Dean. 'Zynga Reports a Profit for Third Quarter on Eve of IPO'. *VentureBeat*, 4 November 2011.

Walther, Joseph B.; Caleb T. Carr, Choi, S., Scott Seung W. Choi, David C. Deandrea, Jinsuk Kim, Stephanie Tom Tong and Brandon Van der Heide. 'Interaction of Interpersonal, Peer, and Media Influence Sources Online: A Research Agenda for Technology Convergence', in *A Networked Self: Identity, Community, and Culture on Social Network Sites*, ed. Zizi Papacharissi. New York and Abingdon, UK: Routledge: 2011, 17-38.

Ctrl-Z: New Media Philosophy
ISSN 2200-8616

< Contents

< Close Issue