Bias in algorithmic filtering and personalization

Article *in* Ethics and Information Technology · September 2013

DOI: 10.1007/s10676-013-9321-6

Engin Bozdag
Philips, Amsterdam

Ethics Inf Technol (2013) 15:209–227
DOI 10.1007/s10676-013-9321-6

ORIGINAL PAPER

Bias in algorithmic filtering and personalization

Engin Bozdag

Published online: 23 June 2013

Abstract Online information intermediaries such as Facebook and Google are slowly replacing traditional media channels thereby partly becoming the gatekeepers of our society. To deal with the growing amount of information on the social web and the burden it brings on the average user, these gatekeepers recently started to introduce personalization features, algorithms that filter information per individual. In this paper we show that these online services that filter information are not merely algorithms. Humans not only affect the design of the algorithms, but they also can manually influence the filtering process even when the algorithm is operational. We further analyze filtering processes in detail, show how personalization connects to other filtering techniques, and show that both human and technical biases are present in today's emergent gatekeepers. We use the existing literature on gatekeeping and search engine bias and provide a model of algorithmic gatekeeping.

Keywords Information politics · Bias · Social filtering · Algorithmic gatekeeping

Introduction

Information load is a growing problem in today's digitalized world. As the networked media environment increasingly permeates private and public life, users create their own enormous trails of data by for instance communicating, buying, sharing or searching. The rapid and extensive travelling of news, information and commentary makes it very difficult for an average user to select the relevant information. This creates serious risk to everything from personal and financial health to vital information that is needed for fundamental democratic processes. In order to deal with the increasing amounts of (social) information produced on the web, information intermediaries such as Facebook and Google started to introduce personalization features: algorithms that tailor information based on what the user needs, wants and who he knows on the social web. The consequence of such personalization is that results in a search engine differ per user and two people with the same friends in a social network might see different updates and information, based on their past interaction with the system. This might create a monoculture, in which users get trapped in their "filter bubble" or "echo chambers" (Sunstein 2002, 2006; Pariser 2011b). Social media platforms, search and recommendation engines affect what a daily user sees and does not see. As knowledge, commerce, politics and communication move online, these information intermediaries are becoming emergent gatekeepers of our society, a role which once was limited to the journalists of the traditional media.

The gatekeeping process is studied extensively by multiple disciplines, including media studies, sociology and management. Gatekeeping theory addresses traditional media bias: how certain events are being treated more newsworthy than others and how institutions or influential individuals determine which information passes to the receivers (Smith et al. 2001). Gatekeeping theory does address the rising power of online information intermediaries, but it focuses on two things: (a) the increasing role of the audience in which users can determine what is newsworthy through social networks (b) the changing role of the journalist, from a gatekeeper to a gatewatcher (Bruns 2008;

E. Bozdag (✉)
Delft University of Technology, P.O. Box 5015, 2600 GA Delft, The Netherlands
e-mail: v.e.bozdag@tudelft.nl

Shoemaker and Vos 2009). The existing theory often considers the online information intermediaries themselves as neutral or treats a web service only as an algorithm, operating without human bias (Hermida 2012; Lasorsa et al. 2012; Bruns 2011). Because these information intermediaries automate their core operations, often, mistakenly, they are treated as objective and credible. Machines, not humans, appear to make the crucial decisions, creating the impression the algorithms avoid selection and description biases inherent in any human-edited media.

Several authors have shown that computer systems can also contain biases. Friedman and Nissenbaum (1996) show that software can systematically and unfairly discriminate against certain individuals or groups of individuals in favor of others. Bias can manifest itself in a computer system in different ways; pre-existing bias in society can affect the system design, technical bias can occur due to technical limitations, emergent bias can arise sometime after software implementation is completed and released (Friedman and Nissenbaum 1996). Several authors have shown how search engines can contain technical biases, especially in coverage, indexing and ranking (Van Couvering 2007; Diaz 2008; Mowshowitz and Kawaguchi 2002; Vaughan and Thelwall 2004; Witten 2007). However, these works are only focusing on the popularity bias. As we will show, many other factors can cause bias in online services.

In this paper we show that online services that process (social) data are not merely algorithms; they are complex systems composed of human operators and technology. Contrary to popular belief, humans do not only take part in developing them, but they also affect the way they work once implemented. Most of the factors that cause human bias in traditional media still play a role in online social media. Finally, even though personalization is seen as a solution by some to prevent technical biases that exist in non-personalized online services (Goldman 2005), we show that personalization not only introduces new biases, but it also does not eliminate all of the existing ones. Others have already pointed to the dangers of implicit and explicit personalization in online services and traditional media (Katz 1996; Van der Hof and Prins 2008; Sunstein 2002; Pariser 2011b). However, they do not identify the potential sources of bias, processes and factors that might cause particular biases. They also do not connect this debate to existing literature in gatekeeping and search engine bias. Our descriptive model of algorithmic gatekeeping aims to achieve this. As Goldman (2011) has recently written about search engine bias: "competitive jostling has overtaken much of the discussion. It has become almost impossible to distinguish legitimate discourse from economic rent-seeking". This overview of bias will hopefully serve as a reference point and contribute to further rational discussion.

Friedman and Nissenbaum (1996) argue that technical bias places the demand on a designer to look beyond the features internal to a system and envision it in a context of use. Minimizing bias asks designers to envision not only a system's intended situation of use, but to account for increasingly diverse social contexts of use. Designers should then reasonably anticipate probable contexts of use and design for these. If it is not possible to design for extended contexts of use, designers should attempt to articulate constraints on the appropriate contexts of a system's use. We believe that our detailed model will help designers and policy makers to anticipate these probable contexts of use and formulate scenarios where bias can occur.

The paper is structured as follows: In "Information overload and the rise of the filters", section we give background information to the problem. In "Personalization: a technical overview", section we give a summary of personalization and how it poses unique problems. In "A model of Filtering for Online Web Services", section we introduce a model of algorithmic and human filtering for online web services including personalization. In "Discussion", section we discuss implications for ethical analysis, social network analysis and design. "Conclusion" section concludes this paper and lists several questions for future research.

Information overload and the rise of the filters

According to Cisco, in 2015, the amount of consumer generated data on the Internet will be four times as large as it was in 2010 (Cisco 2011). McKinkey's research shows that "big data" is a growing torrent. In 2010, 30 billion pieces of content were shared every month with 5 billion mobile phones contributing to it (Manyika et al. 2011). An IBM study reports that every 2 days we create as much digital data as all the data (digital or non-digital) that was created before 2003 and 90 % of the information in the world today has been created in the last 2 years alone (IBM 2011). In online (social) services, users actively contribute explicit data such as information about themselves, their friends, or about the items they purchased. These data go far beyond the click-and-search data that characterized the first decade of the web. Today, thanks to the advent of cloud computing, users can outsource their computing needs to third parties and online services can offer software as a service by storing and processing data cheaply. This shifts the online world to a model of collaboration and continuous data creation, creating so-called "big data", data which cannot be processed and stored in traditional computing models (Manyika et al. 2011).

Even though the amount of generated data on the social web has increased exponentially, our capabilities for absorbing of this information have not increased. Because the mind's information processing capacity is biologically limited (for example, we possess neither infinite nor photographic memory), we get the feeling of being overwhelmed by the number of choices and end up with "bounded rationality" (Hilbert 2012). Researchers across various disciplines have found that the performance (i.e., the quality of decisions or reasoning in general) of an individual correlates positively with the amount of information he or she receives, up to a certain point. If further information is provided beyond this point, the performance of the individual will rapidly decline (Eppler and Mengis 2004).

One means of managing information overload is through accessing value-added information—information that has been collected, processed, filtered, and personalized for each individual user in some way (Lu 2007). Lu argues that people rely on social networks for a sense of belonging and interpersonal sources are recognized as more credible and reliable, more applicable, and can add value through intermediate processing and evaluation to reduce information overload. The general public prefers personal contacts for information acquisition (Lu 2007). As most of the data is produced and stored in the cloud, users delegate the filtering authority to cloud services. Cloud services are trying to extract value and insight from the vast amount of data available, and fine-tune it in order to show what is relevant to their users, often using the users' interpersonal contacts and social networks.

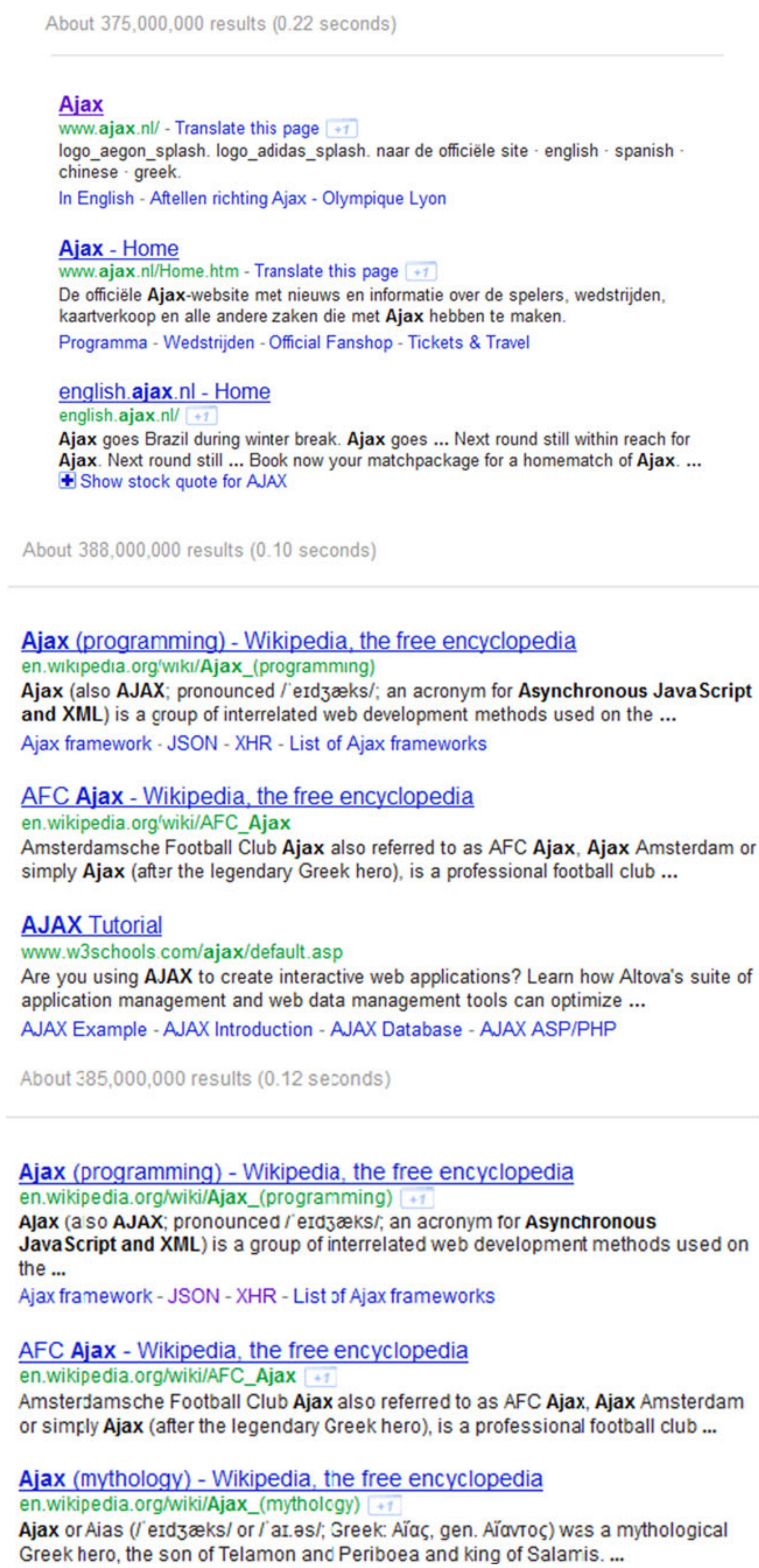
For instance, a search engine returns a list of resources depending on the submitted user query. When the same query was submitted by different users, traditional search engines used to return the same results regardless of who submitted the query. In general, each user has different information needs for their query. The user then had to browse through the results in order to find what is relevant for him. In order to decrease this "cognitive overstimulation" on the user side, many cloud services are exploring the use of personalized applications that tailor the information presented to individual users based upon their needs, desires, and recently on who they know in online social networks. Personalized systems address the overstimulation problem by building, managing, and representing information customized for individual users. Online services achieve this by building a user model that captures the beliefs and knowledge that the system has about the user (Gauch et al. 2007). In this way the system can predict what will be relevant for the user, filtering out the irrelevant information, increasing relevance and importance to an individual user.

Google uses various "signals" in order to personalize searches including location, previous search keywords and recently contacts in a user's social network (Google 2012). As Fig. 1 shows, different users receive different results based on the same keyword search. Facebook on the other hand registers the user's interactions with other users, the so-called "social gestures". These gestures include like, share, subscribe and comment (Upbin 2011). When the user interacts with the system by consuming a set of information, the system registers this user interaction history. Later, on the basis of this interaction history, certain information is filtered out. For instance content produced by certain friends might be hidden from the user, because the user did not interact with those friends over a period of time. Further, photos and videos receive a higher ranking than regular status posts and some posts receive a higher ranking than others (Techcrunch 2011). Personalization algorithms thus control the incoming information (user does not see everything available), but also determine the outgoing information and who the user can reach (not everything shared by the user will be visible to others).

Personalization is a kind of information filtering. However, filtering is not a new concept. During our daily lives we filter information ourselves or delegate the filtering authority to experts, who are called gatekeepers (Priestley 1999). This is because it would require an unreasonable effort and time for any individual to audit all the available information. The gatekeeper controls whether information passes through the channel and what its final outcome is, which in turn determines the way we define our lives and the world around us, affecting the social reality of every person. Traditional media is used to perform this "gatekeeping" role for news, determining what is newsworthy and important for its audience. However, as information technology and cloud computing are gaining importance, online web services that we use every day are slowly taking over the gatekeeping process that used to be performed by the traditional media.

According to Hoven and Rooksby (2008), information is a Rawlsian "primary good", a good that everybody requires as a condition for well-being. Information objects are means to the acquisition of knowledge and in order to be an autonomous person to plan a rational life, we need information (Pariser 2011b). The more (relevant) data individuals can access in their planning, the more rational their life plan will be. Access to information is, then, a value because it may be instrumental in adding alternatives to one's choice set, or in ruling out alternatives as unavailable. As a requirement of justice, in high-technology information societies, people should be educated in the use of information technologies, and have affordable access to information media sufficient for them to be able to participate in their society's common life. Bagdikian (2004) similarly argues that media power is political power and the power to control the flow of information is a major

Fig. 1 Effects of personalization on Google. *First screenshot* is with a logged in user from Netherlands. *Second screenshot* is from an anonymous user from Netherlands. *Last screenshot* is from a logged in user from the US



factor in the control of society. Giving citizens a choice in ideas and information is as important as giving them choice in politics.

In 2005, the Pew Internet and American Life Project reported on the rise of search engines, and surveyed users' knowledge of how they worked. It concluded that "search engines are attaining the status of other institutions—legal, medical, educational, governmental, journalistic—whose performance the public judges by unusually high standards, because the public is unusually reliant on them for principled performance" (Fallows 2005). Personalization and other forms of algorithmic filtering are thus "replacing the traditional repositories that individuals and organizations turn to for the information needed to solve problems and make decisions" (Mowshowitz and Kawaguchi 2002). The services that employ such algorithms are gateways that act as intermediaries between information sources and information seekers. They play a vital role in how people plan and live their lives. Since access to information is a value, and online filters allow or block access to information, building these algorithms is not only a technical matter, but a political one as well. Before discussing how bias can manifest itself in personalization, it is important to first understand how personalization works.

Personalization: a technical overview

Most personalization systems are based on some type of user profile, a data instance of a user model that is applied to adaptive interactive systems. User profiles may include demographic information, (e.g., name, age, country, education level), and may also represent the interests or preferences of either a group of users or a single person. In general, the goal of user profiling is to collect information about the subjects in which a user is interested, and the length of time over which they have exhibited this interest, in order to improve the quality of information access and infer the user's intentions. As shown in Fig. 2, the user profiling process generally consists of three main phases. First, an information collection process is used to gather raw information about the user. Depending on the information collection process selected, different types of user data can be extracted. The

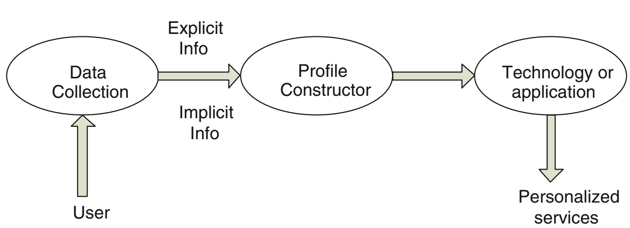


Fig. 2 User profile construction for personalization (adapted from Gauch et al. 2007)

second phase focuses on the construction of a user profile on basis of the user data. Here the collected and stored data are analyzed and processed. In the final phase, the compiled user profile is used in the actual web service, for instance a customized newsfeed in a social networking site, personalized results in a search engine query, or recommended products in an e-commerce site.

A system can build a user profile in two ways:

- Explicitly: the user customizes the information source himself. The user can register his interests or demographic information before the personalization starts. The user can also rate topics of interest.
- Implicitly: the system determines what the user is interested in through various factors, including web usage mining (i.e., previous interaction with the system such as clickthroughs, browsing history, previous queries, time spend reading information about a product), IP address, cookies, session id's, etc.

Explicit user information collection will allow the user to know that the personalization is taking place and he can tailor it to his needs. However, one problem with explicit feedback is that it places an additional burden on the user. Because of this, or because of privacy concerns, the user may not choose to participate. It is also known that users may not accurately report their own interests or demographic data, or, since the profile remains static whereas the user's interests may change over time (Gauch et al. 2007). Implicit user information collection, on the other hand, does not require any additional intervention by the user during the process of constructing profiles. It also automatically updates as the user interacts with the system. One drawback of implicit feedback techniques is that they can typically only capture positive feedback. When a user clicks on an item or views a page, it seems reasonable to assume that this indicates some user interest in the item. However, it is not clear that when a user fails to examine some data item it is an indication of disinterest (Gauch et al. 2007).

Different techniques can be used to make suggestions to users on which information is relevant for them. *Recommendation systems* try to analyze how a user values certain products or services and then predict what the user will be interested in next. A recommendation mechanism typically does not use an explicit query but rather analyses the user context (e.g., what the user has recently purchased or read, and, if available, a user profile (e.g., the user likes mystery novels). Then the recommendation mechanism presents to the user one or more descriptions of objects (e.g., books, people, movies) that may be of interest (Adomavicius et al. 2005; Garcia-Molina et al. 2011).

If this recommendation is done solely by analyzing the associations between the user's past choices and the descriptions of new objects, then it is called “content-based filtering”. Due to increasing user collaboration and user-generated content, personalization can also be done socially. The so-called *social information filtering* (Shardanand and Maes 1995) or *collaborative filtering* (Garcia-Molina et al. 2011) automates the process of “word-of-mouth” recommendations: items are recommended to a user based upon values assigned by other people with similar taste. The system determines which users have similar taste via standard formulas for computing statistical correlations (Shardanand and Maes 1995). For instance, Facebook uses a collaborative filtering called Edgerank, which adds a weight to produced user stories (i.e. links, images, comments) and relationships between people (Techcrunch 2011). Depending on interaction among people, the site determines whether or not the produced story is displayed in a particular user’s newsfeed. This way, a produced story by a user will not be seen by everyone in that user’s contact list. All stories produced by user X can be completely hidden in user Y’s newsfeed, without the knowledge of both users.

According to Chatman (1987) and Lu (2007), people’s information needs are highly diversified and individualized, making applicable and value-laden information most desirable, and yet the hardest to obtain. Interpersonal sources can, to a great extent, minimize these difficulties and maximize the utility of information. Even though personalization technologies such as Grouplens (Resnick et al. 1994) have existed for a while, the rise of social networks and the exponential increase in produced and shared information in online services are changing the impact this technology has. According to Garcia-Molina et al. (2011), information providing mechanisms (e.g. search engines) and personalization systems have developed separately from each other. Personalization systems like recommendation engines were restricted to a single homogenous domain that allowed no keyword search. Search engines on the other hand were geared toward satisfying keyword search with little or no emphasis on personalization or identification of intent. These two systems were separated partly due to a lack of infrastructure. Today, due to a combination of a powerful and cheap back-end infrastructure such as cloud computing and better algorithms, search engines return results extremely fast, and there is now the potential for a further improvement in the relevancy of search results. So, we now see a trend where personalization and information providing mechanisms are blending.

A model of filtering for online web services

Existing work on gatekeeping theory often points out the changing role of the journalist from a gatekeeper to a gatewatcher (Shoemaker and Vos 2009; Bruns 2008). With the increasing popularity of the online media and social networks, every user can share information depending on what he thinks is important. Scholars thus argue that by using online services, the audience can exert a greater control over news selection and can focus on issues that they consider more relevant, which in turn empowers audiences and erodes the degree of editorial influence over the public’s issue agenda (Althaus and Tewksbury 2002). Some even argue that the gatekeeping role performed by the traditional media becomes irrelevant; gates are disappearing (Levinson 1999). Information may diffuse through social networks next to mass media channels; therefore any audience member can be a gatekeeper for others. Journalists now become a “gatewatcher”, providing a critical analysis of existing topics that are chosen by the community (Bruns 2008).

Some also claim that the platforms the new “gatewatchers” operate are neutral. According to Bruns (2011), tools such as Twitter are neutral spaces for collaborative news coverage and curation operated by third parties outside the journalism industry. As a result, the information curated through collaborative action on such social media platforms should be expected to be drawn from a diverse, multiperspectival range of sources. Also Lasorsa et al. (2012) claim that platforms such as Twitter are neutral communication spaces, and offer a unique environment in which journalists are free to communicate virtually anything to anyone, beyond many of the natural constraints posed by organizational norms that are existing in traditional media.

However, as we shall show, the gatekeeping process in online information services is more than a simple transition from editor selection to audience selection or from biased human decisions to neutral computerized selections. We argue that human factors play a role not only in the development of algorithms, but in their use as well. We show that factors that caused bias in mass media news selection still play a role in information selection in online web services. Online information intermediaries, similar to the traditional media, can control the diffusion of information for millions of people, a fact that gives them extraordinary political and social power. They do not provide equal channels for every user and they are prone to biases. Just as any computer system, they can unfairly discriminate against certain individuals or groups of individuals in favor of others (Friedman and Nissenbaum 1996).

Source selection algorithm

At the stage of “Collection and Selection” (Fig. 3), the online service starts to collect its information from various

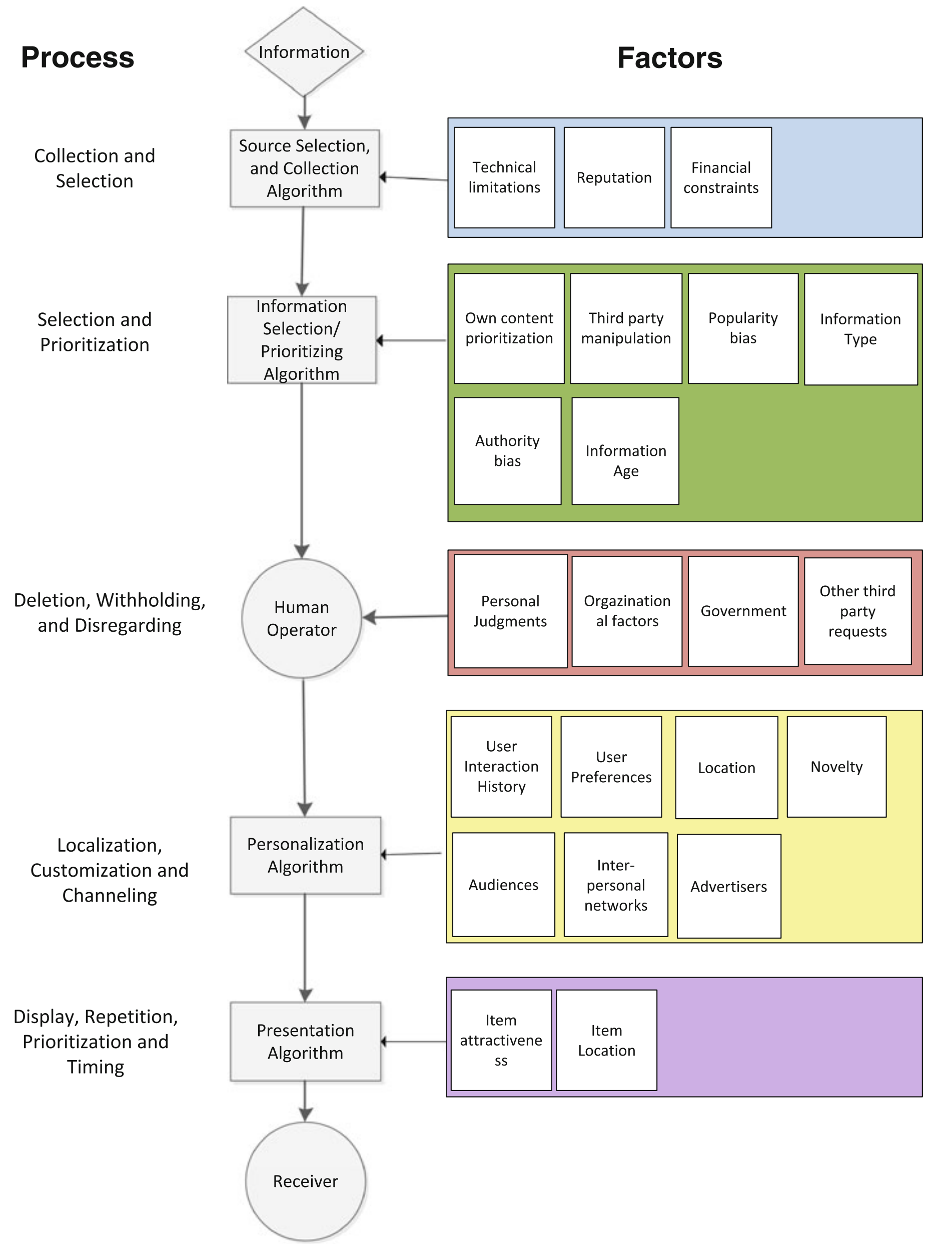


Fig. 3 A model of filtering for online web services including personalization

sources. For instance a search engine will automatically crawl the web, while the social network site will collect information produced by its users. However, similar to the traditional media where gatekeeping starts with journalists (Chibnall 1975; Shoemaker et al. 2008), algorithmic gatekeeping already starts at source selection. First of all, not all information is digital, thus all non-digital information will be absent from online information intermediaries. Further, not all digitally available information will be available to each service, for instance search engines do not index all the data available on the Internet, leading to coverage bias (Goldman 2005; Vaughan and Thelwall 2004). Google admits that the company does not index every one of the trillion pages on the web, because they are similar to each other or because Google considers some of them not useful to the searcher (Google 2008). Technical reasons can also prevent a search engine to crawl a site. The design of the website might make the source collection and indexing process difficult or the site itself might be explicitly blocking the crawling process (Barzilai-Nahon 2008). Further, if a resource has a bad reputation, for instance if it is suspected as an illegal site, it might be left out of the whole collection process. It is also possible that the source does not want to be included in the index due to various reasons. For instance not every page in Facebook or Twitter is indexable by Google. (Sullivan 2012).

Information selection and prioritization algorithm

In traditional media, newspaper editors select some of the messages produced by journalist to make news (Barzilai-Nahon 2009). Algorithms used in web services (such as ranking algorithm in a search engine, or news feed algorithm in a social network) make similar decisions. The design of these algorithms is affected by choices made by designers, i.e., which factors to include in the algorithm, and how to weigh them.[1] To serve majority interests, information intermediaries often include popularity metric in their ranking algorithm. A search algorithm for instance can give more weight to information coming from popular websites, to support majority interests and values. As a result, seekers will have trouble finding the less popular and smaller sites (Nissenbaum and Introna 2000).

Because the information filtering is automated, it might be manipulated by activities from third parties. This happens with the so-called "black-hat" search engine optimization techniques. This is a method of raising the profile of a Web site with methods that Google considers tantamount to cheating (Segal 2011). Another factor is own product/service prioritization. The EU recently received a complaint from a shopping search site that claimed it and other similar sites saw their traffic drop after Google began promoting its own services above conventional search results (Foundem 2009; Efrati 2010; Albanesius 2011; Edelman 2011). Google also integrates content from its social networking platform Google Plus into Google search results, causing protest by the social networking platform Twitter (Searchenginewatch 2012). Studies also showed that Google and Bing search engines both reference their own content in its first results position when no other engine does (Wright 2011; Edelman 2011). Facebook is criticized for favoring the products of its partners (Fong 2011). The algorithm can also prioritize certain types of information over others. For instance, it is claimed that Facebook treats video and pictures as more important than links and status updates (Taylor 2011). Similarly, comments on an item are four times more valuable than "likes" (Wittman 2011).

In traditional media, regardless of the size of an event such as a public protest, the likelihood that the event will be reported in the media will depend on the current agenda. This is because both print and electronic media regularly focus upon selected issues over a sequence of days, creating the phenomena of "issue attention cycles" (Smith et al. 2001). We can observe similar behavior in social media. Twitter has a feature called trending topic (TT), in which most popular topics Twitter users are talking about in a particular location are highlighted. However Twitter does not solely check popularity of an item while determining TT's, it favors novelty over popularity. Twitter checks if the user updates on a specific topic is increasing quickly enough. Even if a topic is large volume wise, if the increase rate is small or if it is not novel, it won't make it to the "trending topics" (Twitter 2010). This means that it is much easier for a term never seen before to become a Twitter trend and the longer a term stays in the TT list, the higher velocity required to keep it there (Lotan 2011). This novelty factor caused the hashtag "IcantRespectYouIf" to be a TT in the US, while #OccupyWallStreet not making it to the list. This is because when #OccupyWallStreet was a TT throughout the world, it had *previously* trended in the US, and now there were no more *new* people in the US talking about it.

According to Gillespie (2012), this choice fosters a public more attuned to the "new" than to the discussion of persistent problems, to viral memes more than to slow-building political movements. The exact algorithm that determines the trending topics is unknown and this opacity makes the TT, and their criteria, deeply and fundamentally open to interpretation and suspicion (Gillespie 2012).

Trending topic differs in important ways from those employed in personalization, as it presents itself as a measure of popularity.[2] However, since algorithms such as TT can differ per country, region or city, they might be used to customize content, as an important signal. Popularity can thus be an input to customize items for a group of users. This is still tailored content, but not for an individual, but for a group of individuals.

Finally, the age of an information source or the age of the information item can also matter. In Google search engine, the number of years a domain name is registered has an impact on search ranking; domain names that exist for a period of time are preferred over newly registered ones (Jacobs 2010). In Facebook, the longer a status update has been out there, the less weight it carries. A news item is prioritized over an old item (Techcrunch 2011). This might for instance lead companies to post updates when their audience is most likely to be online and using Facebook.

Human operator

In traditional media, individual factors such as personal judgment can play a role during the selection of news items for a newspaper. An editor's decisions can be highly

[1] For instance, Facebook uses an algorithm called Edgerank to determine how a newsfeed of a user is constructed. It is believed that several factors are used to select/prioritize user updates, such as affinity between the receiver and sender, and the date of the published update. However, the exact formula is unknown. See Techcrunch (2011).

[2] We would like to thank the anonymous reviewers to point out this fact.

subjective and can be based on the gatekeeper's own set of experiences, attitudes and expectations, leading to a selection bias (Gans 2005). Online web services such as search engines frequently claim that such human bias do not exist in their systems. They claim that their core operations are completely automated, but this is false. Humans in online services also make editorial judgments about what data to collect delete or disregard. According to Goldman, online services manually inspect their index and make adjustments (Goldman 2005). For instance search engines make manual adjustments of a web publisher's overall rating or modify search results presented in response to particular keyword searches (Goldman 2005). The Dutch newspaper Trouw's entire domain name and all hosted pages were removed from Google index because of a violation of the company policy (Groot 2004; Dekker 2006). Google itself has admitted that the company manually demotes websites (Metz 2011a). Similar to blacklisting, search engines can also perform whitelisting. For instance Google recently mentioned that it uses whitelists to manually override its search algorithms (Metz 2011b).

Information deletion or withholding is not specific to search engines. Facebook a photo of two men kissing from a user's Wall due to a violation of the site's terms of service (Zimmer 2011). There are also claims that Facebook denies and removes advertisements designed for gay audience with no nudity or sexual content, labeling it "inappropriate" (Accuracast 2010). Others claimed that Facebook labeled their posts containing links to a political activism site as spam and prevented the users disseminating this information (Badash 2011). Facebook has also removed pages because of offensive content, but later reinstated them (Kincaid 2010; Ingram 2011). Facebook spokesman blamed the human reviewer in some of the cases, but did not reveal the criteria the company uses on what makes content offensive or in violation with the company's terms of use. Twitter similarly removes certain 'trending topics' if it considers it as "offensive" (Costolo 2011).

Scholars in media studies argued that organizational factors in traditional media play a more important role than individual judgments. In the uncertainty of what tomorrow's news will be, journalists use so-called routines, patterned, repeated practices and forms, to view and judge in order to define news as predictable events (Fishman 1988). Similarly, online web services employ operators to delete, withhold or disregard information, to enforce company guidelines. Even though these operators have to obey a set of rules to apply, they have, just like journalists, their own values and can pass personal judgments. This might give the image that the operator is bound to strict rules, and acts merely as an enforcer. However people do not always execute rules in the same way and individual-level characteristics are still important (Shoemaker and Vos 2009).

Human operators of online services have to evaluate removal requests coming from governments. For instance, recently, A Delhi Court ordered 22 social networking sites (including Facebook, Google, Yahoo and Microsoft) to remove all "anti-religious" or "anti-social" content and file compliance reports. Google has a list of content removal requests from governments all around the world (Google 2011). Operators also have to deal with requests coming from third parties. For example, Google regularly removes content due to copyright claims coming under the Digital Millennium Copyright Act, Section 512(c). This act gives providers immunity from liability for their users' copyright infringement, if they remove material when a complaint is received (Chilling effects 2005).

Personalization algorithm

According to Goldman (2005), personalized ranking algorithms reduce the effects of technical bias introduced by algorithms in online intermediaries. Goldman argues that personalization algorithms increase relevancy and produce a different output per individual user. This in turn diminishes the weight given to popularity-based metrics and reduces the structural biases due to popularity. Personalization might increase relevance, however as we show in this subsection, designing only for this value will introduce problems.

User interaction history and user preferences

As we have argued in "Personalization: a technical overview", section users could personalize the information they receive by giving their preferences explicitly. In this way they can receive personalized information on the criteria they know. However, if the user's interests change over the time and if the user does not update their filter, they might miss some information that might be of interest to her. Lavie et al. (2009) found that people might be interested in things that they did not know they were interested in, due to the formulation of the topic. Some users have asserted that they were not interested in politics, but later it was shown that their perception of "politics" was limited to local politics. They later have shown interest in international politics (Lavie et al. 2009). Lavie et al. argue that, overall, users cannot accurately assess their interests in news topics. Similarly Tewksbury (2003) reports that user's declared and actual interests may differ.

In his book Republic.com, Sunstein (2002) developed his concern that explicit personalization will assist us to avoid facts and opinions with which we disagree, leading people to join online groups that conform with their existing beliefs. Since democracy is most effective when citizens have accurate beliefs and to form such beliefs, individuals must encounter information that will sometimes contradict their preexisting views. Sunstein argues that explicit personalization will undermine deliberative democracy by limiting contradictory information.

Implicit personalization using user interaction history has its own concerns. Pariser (2011b) argues that online services can cause citizens to be ill-informed about current events and may have increasingly idiosyncratic perceptions about the importance of current events and political issues. This might occur because online services are trying to improve accuracy at the expense of serendipity, leading to what Pariser calls “filter bubble”. Even if users wanted to diversify their network explicitly, information intermediaries silently filter out what they assume the user does not want to see, hiding information posted by opposite end of political spectrum. For Sunstein, explicit excessive personalization leads to never seeing the other side of an argument and thus fostering an ill-informed political discourse. For Pariser, excessive implicit personalization leads to an unhealthy distaste for the unfamiliar. The problem is thus an automatic cyberbalkanization, not an “opt-in” one. It happens behind the scenes and we do not know what we are not seeing. We may miss the views and voices that challenge our own thinking.

Pariser argues that online personalization algorithms are designed to amplify confirmation bias, Consuming information that conforms to our beliefs is easy and pleasurable; consuming information that challenges us to think differently or question our assumptions is difficult. Pariser notes that we all have internal battles between our aspirational selves (who want greater diversity) and our current selves (who often want something easy to consume). Pariser argues that the filter bubbles edit out our aspirational selves when we need a mix of both. Pariser believes that the algorithmic gatekeepers need to show us things that are not only easy to consume but also things that are challenging, important and uncomfortable and present competing points of view. Pariser states that filter bubbles disconnect us from our “ideal selves”, that version of ourselves that we want to be in the long-run, but that we struggle to act on quickly when making impulse decisions.

Location

As we have shown in “Personalization: a technical overview”, section content can also be personalized based on location. Large web-search engines have been “personalizing” search to some extent for years. Users in the UK will get different results searching for certain terms, especially commercial ones, than users in the US Results can change between different cities as well (Garcia-Molina et al. 2011). The idea is that the user will be more interested in local content. However, this will depend on context of information. For instance, if I am looking for a restaurant, I would want my search engine to personalize results based on location, the system should show me pizzerias in Rotterdam, but not in New York. However, if I am looking for some technical information in a forum to solve a PC problem, then I do not necessarily care about the location (if I can speak multiple languages). Currently, most personalization systems filter information based on location without taking the context into the account. This might always favor local content, even if the quality or the relevance of the local content is inferior to a non-local content.

Audiences

While traditional news media outlets want to satisfy their readers and viewers, it is much more difficult for them to modify their selection criteria in real time, than it is for online gatekeepers. Online gatekeepers have immediate feedback about what queries are issued, what content is selected and what sites are accessed. For instance online services can observe user behavior through entered queries or clicked links to modify its algorithms accordingly. However, online services can also capture user’s intent by using social gestures. Examples of these social gestures include the “like” and “subscribe” buttons in Facebook and the “+1” button in Google search. By clicking on these buttons users express their interests and see what item is popular. Google currently does not use these (anonymous) votes to personalize search results, but such approaches are well known in computer science literature. Search behavior of communities of like-minded users can be harnessed and shared to adapt the results of a conventional search engine according to the needs and preferences of a particular community (Smyth 2007). Because similarities will exist among community members’ search patterns and web search is a repetitive and regular activity, a collaborative search engine can be devised. This human PageRank or “social-graph”, using +1 results to give context to the popularity of a page, can be a supplement (or alternative) to the link graph Google is currently using.

Some claim that the community is wiser than the individual. However, community driven filtering has its own problems. For example, in social news aggregator Reddit, where anonymous users submit links to items, comment on them, vote on the submitted items and comments, the community determines what is newsworthy, for every topic. Users can personalize their news feed by explicitly subscribing to certain subtopics, but the popularity metric is used in every subtopic. In Reddit, the timing of the story submission is important. If a good news item is submitted outside of Internet prime-times, it will not receive enough votes to make it to the front page. The result is that most submissions that originate in the US end up being dominated by US comments, since new comments posted several hours after the first will go straight to the middle of the pile, which most viewers will never get to. Submission time has a big impact on the ranking and the algorithm will rank newer stories higher than older. In Reddit, first votes also score higher than the rest. The first 10 upvotes count as high as the next 100, e.g. a story that has 10 upvotes and a story that has 50 upvotes will have a similar ranking. Controversial stories that get similar amounts of upvotes and downvotes will get a low ranking compared to stories that mainly get upvotes (Salihefendic 2010). Further, the user will receive positive or negative points on the story he submitted. The individual might remove the story due to decreasing points in his reputation.

It is also known that in such vote-based social news sites, the amount of contacts or followers one has can also determine whether his story will make it to the front page. Having a large number of contacts will make it easier to reach the front page (more friends, more votes). Also, some social news aggregators divide the stories into topics. If a topic has a small number of subscribers, the chance that it will make it to front page is small (Klein 2011). Even the items that do not make it to the front page will bring traffic to the submitted site. Therefore social news aggregators like Reddit are being used and manipulated by online marketing professionals, in order to draw more traffic to their products or services. Similarly, Facebook's like button can also be gamed. Digital marketing companies can create fake users and buy "friends" and "likes" (Tynan 2012). These companies use software to automate clicking the "Like" button for a certain page. Such software can bypass Facebook's security system. If popularity is devised by only the number of likes and used as an input for users in a certain region, it can also cause bias in personalization.

Interpersonal networks

According to Chen and Hernon (1982), the general population tends to obtain information through interpersonal networks, rather than formal means. Durrance (1984) found that more than 64 % of her research participants used interpersonal sources. Sturges maintains that there is a "fundamental preference for information mediated by human interaction" and that "there is evidence of this from all parts of the world and from most important aspects of human life" (Sturges 2001). Katz and Lazarsfeld (2005) argue that we live in communities and we are inherently tied to different social connections. We interact in formal or informal social groupings, in so-called "primary groups" such as families, friends, work teams, clubs or organizations. These primary groups delineate major life boundaries for each one of us in society, our routine activities mainly occur in these primary groups.

Since our lives are mainly contained in primary groups, our attitudes and opinions tend to derive from them as well as our sources of information. Primary groups provide us with "social reality" to validate our actions. As we encounter unknown situations and difficult decisions, we turn to and consult our social contacts, including both strong (e.g., family and friends) and weak ties (e.g., colleagues, acquaintances) to help us form opinions and find solutions (Granovetter 1981). Lu (2007) argues that, through interactions concerning a particular issue, a primary group tends to develop a common view and collective approach, hence, provides a social reality that helps and validates decision making by its members. Because members of a primary group share the community language and background information, their communication is made effortless. Information so transmitted becomes easily accessible and digestible (Lu 2007).

Because of these reasons, instead of relying on user's explicit preferences, or using an anonymous popularity metric, personalization services started to use interpersonal relationships to filter information. For instance Facebook launched a program called "instant personalization" with an exclusive set of partners, including the restaurant aggregator site Yelp, Microsoft online document management site docs.com, customizable Internet radio sites Pandora and Spotify. These partners have been given access to public information on Facebook (e.g., names, friend lists, and interests and other information users have shared on their Facebook profiles) to personalize a user's experience on the partner's site. As an example, online music service Spotify requires a Facebook account, and using the friends list in Facebook, it shows the user what her friends have listened to. The idea here is, since these contacts are part of our primary group, we can trust their judgment on which information is newsworthy. If our primary groups are available in every web service we use, then our experience using that web service can be customized.

Similarly Google introduced social search in 2009, personalizing search results based on people you know in Facebook and Twitter, rather than your personal behavior. As a latest move, in 2012, Google introduced a feature called "Search plus your world". This feature personalizes the results using user connections in Google Plus, Google's social networking platform. This means you might see a picture of a friend's car when you search for a new automobile, or a restaurant recommended by a friend when you search for a place to eat. Even if you aren't a Google+user, Google search results will show content posted publicly on the social network that it judges to be relevant—profile pages and pages dedicated to particular topics (Knight 2012).

Advertisers

Traditional mass media is primarily supported by commercial sponsorship. This can cause the newspapers to delete, change or prioritize news items due to advertising pressure (Soley 2002). Same pressure applies to online services; the majority of online service revenues come from advertising (O'Dell 2011; Schroeder 2011; US Securities and Exchange Commission 2009). Personalization is a very attractive tool for advertisers, as user data collected for information filtering can be used for behavioral targeting. This sort of online targeting provides more relevant online advertising to potential upcoming purchases. Using the built up user profile in online services, advertising networks can closely match advertising to potential customers. According to Guha et al. (2010), Facebook uses various profile elements to display targeted advertisement including age, gender, marital status, and education. A Facebook advertiser can target users who live within 50 miles of San Francisco, are male, between 24 and 30 years old, single, interested in women, like skiing, have graduated from Harvard and work at Apple (Korolova 2010). Google allows advertisers to target ads based not just on keywords and demographics, but on user interests as well (Opsahl 2009). Companies have recognized that providing advertisements along with their recommendations (suitably distinguished from the recommendation results) can be extremely profitable. For instance, the auction site Ebay provides a "deal of the day" for all visitors to the site, in addition to "buy it now", special items directly sold from a provider for a fixed price—both of these are essentially advertisements (Garcia-Molina et al. 2011).

Presentation algorithm

Once information is chosen through the information selection algorithm and personalized for the user, it does not mean that it will be seen and consumed. The placement of the information might determine if it makes it out of the filter. Joachims and Radlinski (2007) show that the way a search engine presents results to the user has a strong influence on how users act. In their study, for all results below the third rank, users did not even look at the result for more than half of the queries. Bar-Ilan et al. (2009) report similar findings. Yue et al. (2010) report that the attractiveness of information can also cause presentation bias if the title and abstract of a resource is bolded, it generates more clicks. They also show that people tend to click on the top and bottom results. These findings show that what the user will consume can be affected by the algorithm, even after source selection and personalization.

Discussion

Implications for an ethical analysis

Personalization is the latest step in this algorithmic filtering process. As we have argued, even though personalization algorithms have existed since the 1990s, information providing services such as search engines did not contain such algorithms until recently. This is mainly due to the recent availability of cheap and powerful backend infrastructure and the increasing popularity of social networking sites. Today information seeking services can use interpersonal contacts of users in order to tailor information and to increase relevancy. This not only introduces bias as our model shows, but it also has serious implications for other human values, including user autonomy, transparency, objectivity, serendipity, privacy and trust. These values introduce ethical questions. Do private companies that are offering information services have a social responsibility, and should they be regulated? Should they aim to promote values that the traditional media was adhering to, such as transparency, accountability and answerability? How can a value such as transparency be promoted in an algorithm? How should we balance between autonomy and serendipity and between explicit and implicit personalization? How should we define serendipity? Should relevancy be defined as what is popular in a given location or by what our primary groups find interesting? Can algorithms truly replace human filterers?

A relevant value to bias is information diversity. For instance if a search engine is exercising bias toward an advertiser, it will be limiting the diversity and democracy inherent to the information (Granka 2010). Information diversity is a rich and complex value that can be conceptualized in many different ways, and its interpretation differs significantly per discipline. In media studies, it might be translated as "minority voices having equal access in the media" or "the degree which the media relates to the society in such a way to reflect the distribution of opinion as it appears in the population" (Van Cuilenburg 1999). In Computer Science literature, it can be defined as "variety in the products offered by the system", "helping user find items he cannot easily find himself" (Zhang and Hurley 2008) or "identifying a list of items that are dissimilar with each other, but nonetheless relevant to the user's interests" (Yu et al. 2009). While media studies are analysing this ethical value in detail, almost all scholars of search engine diversity seem to be limiting their understanding of "bias" and "diversity" to popularity bias (Granka 2010). As our model shows, popularity is only one of the many factors that cause bias. We need a normative conceptualization of the value information diversity that borrows notions from media studies, such as media ownership, content diversity, viewpoint diversity, reflection and open-access (Cuilenburg 1999). Only then can we translate this complex value into design requirements of information intermediaries and move towards a solution.

We believe that normative arguments based on our model will be stronger, more concrete and constructive. As an example, take the value user autonomy. Autonomy is centrally concerned with self-determination, making one's own decisions, even if those decisions are sometimes wrong (Friedman and Nissenbaum 1997). Autonomy is thus the individual's ability to govern herself, be one's own person, to be directed by considerations, desires, conditions, and characteristics that are not simply imposed externally upon one, but are part of what can somehow be considered one's authentic self (Christman 2011). It is this aspect of decision-making that allows us to be responsible for the consequences of our actions. While designing technology, one can thus assume that designers should maximize user autonomy by following the simple dictum that more control leads to more user autonomy. After all, if autonomous individuals need to have freedom to choose ends and means, then it could be said that wherever possible and at all levels, designers should provide users the greatest possible control over computing power. Considering this notion of autonomy, one could argue that personalization algorithms should always be fully customized and should be based on explicit personalization. However, as the model shows, explicit personalization based on user preferences is also prone to bias. People might be interested in things that they did not know they were interested in, due to the formulation of the topic. Further, users might not accurately assess their interests in certain information items. As we have mentioned, user's declared and actual interests may differ.

This seems to suggest that autonomy in this context should not be understood as "full user control". User autonomy seems to have less to do with simply the degree of control and more to do with what aspects of the algorithm are controllable, and the user's conception and knowledge of the algorithm. As Friedman and Nissenbaum (1997) notes, achieving higher order desires and goals will enhance autonomy, whereas excessive control may actually interfere with user autonomy by obstructing a user's ability to achieve desired goals. This means that, implicit personalization must be combined with explicit personalization to decrease excessive control. For instance a personalized search engine might be implemented in such a way that, the system enters a dialogue with the user, explicitly stating that a certain query is personalized, explaining why and due to which reasons it is personalized. The system can thus make assumptions to predict what the user might like, but it should refine itself by asking simple questions to the user to confirm if those assumptions were correct. While the user might not control the full algorithm, the system might receive feedbacks and show the user under which conditions it is making certain recommendations.

As we have argued, information should be accepted as a primary good, a vital good for people to plan their lives rationally and to participate adequately in the common life of their societies (Hoven and Rooksby 2008). Thus, having access to information affects the value of liberty perceived by an individual. We therefore argue that personalizing algorithms affect the moral value of information as they facilitate an individual's access to information. Contrary to earlier stages of the Internet-era, when the problem of information access boiled down to having access to hardware, today the problem of access to information concerns the ability to intentionally find the right information, or the likeliness of unintentionally stumbling upon the relevant information.

Some argue that users should sabotage the personalization system by deliberately clicking on links that make it hard for the personalization engines, erasing cookies, unlocking everyone on a social network, posting something and then ask the Facebook friends to click the "Like" button and comment, or simply switch to a service that does not use personalization (Pariser 2011a; Elgan 2011). However, these tactics are tedious, not always possible to perform and their effect depends on the implementation of the current system. Further, personalization might actually have a positive effect on the ecology of the cyberspace: the incentives to game the system and invest in practices like "search engine optimization" can become weaker (Morozov 2011; Goldman 2005). We should come with design suggestions to minimize the bad effects and improve the good effects of this technology instead of trying to get rid of it all together.

The question is then not whether to have personalization or not, but how to design morally good personalization technology. 'Having too much information with no real way of separating the wheat from the chaff' is what Benkler (2006) calls the Babel objection: individuals must have access to some mechanism that sifts through the universe of information, knowledge, and cultural moves in order to whittle them down into manageable and usable scope'. The question then arises whether the service providers currently active on the Internet are able to fulfill the 'human need for filtration'. Although the fulfillment does not hinge on proprietary services alone as there are cooperative peer-production alternatives that operate as filters as well, the filtering market is dominated by commercial services such as Google and Facebook (Hitwise 2010). Having an option to turn it on or off is not really a choice for the users, as they will be too dependent on it in the existence of information overload.

Implications for design

In order to anticipate different contexts of use in personalization, a value based study such as Value Sensitive Design (Flanagan et al. 2008; Friedman et al. 2006) seems to be the right direction. Value sensitive design consists of an empirical investigation accompanied by a philosophical analysis and a technical study. Friedman and Nissenbaum (1996) argue that designers should not only envision a system's intended situation of use, but to account for increasingly diverse social contexts of use. Designers should then reasonably anticipate probable contexts of use and design for these. If it is not possible to design for extended contexts of use, designers should attempt to articulate constraints on the appropriate contexts of a system's use. Bias can manifest itself when the system is used by a population with different values than those assumed in the design. This is especially true for the design of most online information intermediaries, where users from the whole world will be served instead of only local ones.

Another issue that is relevant to the design of personalization algorithms and other filtering mechanisms is exposure diversity. Even if an information intermediary provides a balanced information diet, this does not guarantee that the user will actually consume this information (Napoli 1999; Helberger 2011; Munson and Resnick 2010). Content diversity is not equal to exposure diversity. We need to devise methods to increase the consumption of challenging content by users. Munson and Resnick (2010) distinguished two types of users: challenge averse (those who ignore diverse content) and diversity seeking. They tried to show more diverse content to those who were challenge averse, for instance by highlighting agreeable items or showing agreeable items first. However, this did not increase users' consumption habits, they still ignored challenging items. This requires us to research further how challenging items can be made attractive to users so that they actually consume the incoming information.

Implications for the design of social filtering

Media scholars often argue our interpersonal contacts have become our gatekeepers (Shoemaker and Vos 2009). However, if this approach becomes ubiquitous in design, it can lead to problems. First, this obviously raises concerns for privacy. An item a user has consumed can be shared with others without their notice. The Electronic Privacy Information Center, American Civil Liberties Union and American Library Association claim the changes have made sharing information on Facebook a passive rather than active activity. In this way, users might reveal more than they intend (Nagesh 2011). Even if sharing process was more active, it can still cause issues. For instance, an item a user has shared in a social network in certain context and has forgotten can reappear in a Google search result in a different context. Further, an implicit user profile built for personalization leads to epistemological problems. Does the knowledge about the user (gathered by user's interaction with the system) represent the reality? Does the user interact with its primary group the same way he interacts in the offline world? How much does a user have a say in this built profile and to what degree can he control the dissemination of this representation of himself?

Second, not everyone in our online social networks will be part of our primary group; not every online "friend" is our real friend and we might share different things with our online friends. We sometimes add people to our network because of courtesy, as it otherwise might cause relationship problems in the offline world ("Why did you not answer my friend request?"). To remedy this, we can arrange the level of our relationship with others in a social network; we can divide them into lists or groups. We can then choose what we want to share with which group. However, our contact list in a social network can be connected with a different service, for personalization. When we use our social network in another service, lists we have created can suddenly disappear. For instance, Spotify uses Facebook contact list to provide recommendations per individual user. However, it ignores all the lists that have been created and shows what all friends have listened to regardless of the relationship between the user and the friend. The categorization the user has set in the Facebook platform in order to define and control his relationships are gone when the Facebook data is used elsewhere. Next to increasing information overload, this can also cause privacy issues. Even if I choose to share things with some people in Facebook context, everything I listen to in Spotify will be shown to all my Facebook users. This context loss will be more common as more services integrate with each other.

Third, not everyone has competence on every subject. Scholars in various disciplines have found that there are strategic points for the transmission of information in every group (Agada 1999; Chatman 1987; Lu 2007). Even though it is possible that people can interact randomly with anyone who has available information, information transmission is never a simple aggregation (Slater 1955; Katz and Lazarsfeld 2005). Some individuals, who are more information-savvy, will automatically occupy strategic positions to facilitate access to information to others. Depending on the subject matter, not everyone in a group is equally important or qualified in providing information. Those who have more knowledge will act as gatekeepers. I might trust John's competence in football, and use him as my gatekeeper in this subject, but not in the area of international politics. However, in most online services, we get to see everything published by a user, or nothing at all. We need mechanisms to assess the competency of the information sharer and determine the needed gatekeeper for a given context.

Fourth, online services are trying to capture user's intent by using social gestures. Examples of these social gestures include the "like" and "subscribe" buttons in Facebook and the "+1" button in Google search. By clicking on these buttons users express their interest and communicate to their peers. However, this sort of expression seems somehow limiting (Pariser 2011b). The reason of the expression and the emotion behind the expression is not captured by the button. There is a difference between liking a film, liking a director, liking a genre or liking films of a certain period. I might like a film for various reasons: to recommend to friends, to express my identity, to receive further film recommendations or to add it into my collection for later use. Such buttons are simplifying complex human actions and emotions into a single dimension. As Friedman and Nissenbaum (1996) have argued, attempting to formalize human constructs such as discourse, judgments, or intuitions and trying to quantify the qualitative, discretizing the continuous will lead to biases.

Fifth, online services assume that users want to have an online experience where consuming any sort of information is done socially and collaboratively. This is why Google is making social search the default type of search and Facebook persuades users to share more information or leave a trace of a completed activity, by its "frictionless sharing". These approaches aim to make sharing an effortless activity, in which everything is shared and hopefully some things will be found interesting by the users. However by promoting ease, they are undermining not only privacy, but also autonomy. In a frictionless sharing environment, user now cannot actively reflect on things he consumes and choose on what to share.

Finally, if we know the information we consume is being shared and read by our primary groups, we might change our behavior on what to share, and even choose what to consume if this is shared automatically. According to Sunstein (2008), group members may fail to disclose what they know out of respect for the information publicly announced by others. That is, even if we have big doubts about claims made by the majority of a group, we might think they are not errors at all; not so many people can be wrong. Individuals can also silence themselves to avoid the disapproval of peers and supervisors. As a result of these two forces, information cascades might occur; individual errors might amplify instead of being corrected, leading to widespread mistakes. Information held by all or most will be prioritized over held by a few or one.

Implications for social network analysis

While bias might manifest itself in the social platform, users themselves might be biased in information sharing. Therefore we need to determine whether bias occurs naturally in social networks, as personalization algorithms use more and more social data. Do users tend to follow like-minded users? Do they do this intentionally? Do they only share things that they agree with? Do they receive diverse information directly or indirectly? Do they only want to follow popular items coming from major news sources as the current services, or does the minority receive a chance to contribute to the debate? Is the sharing behaviour of the user changing with what he is receiving? Does culture have an affect in diverse information seeking behaviour? To answer such questions, we need to perform more empirical studies.

Facebook performed one of the few studies that actually studies bias in social networks (Bakshy 2012). The empirical study suggests that online social networks may actually increase the spread of novel information and diverse viewpoints. According to Bakshy (2012), even though people are more likely to consume and share information that comes from close contacts that they interact with frequently (like discussing a photo from last night's party), the vast majority of information comes from contacts that they interact with infrequently. These so-called "weak-ties" (Granovetter 1981) are also more likely to share novel information.

Even though this is one of the first empirical studies that aims to measure information diffusion, there are some concerns with it: First of all, the study is not repeatable and the results are not reproducible. Facebook scientists simply manipulated newsfeed of 253 million users, which only Facebook can perform. Second, our weak ties give us access to new stories that we wouldn't otherwise have seen, but these stories might not be different ideologically from our own general worldview. They might be new information, but not particularly diverse. The research does not indicate whether we encounter and engage with news that opposes our own beliefs through links sent by "weak links". It could very well be that we comment on and re-share links to cat videos sent by our previous neighbour, or read a cooking recipe posted by our vegetarian friend, ignore anything political or challenging/contradictory to our world view. The study measures the amount of different information one gets, not different world-views. Third, the users might refrain from novel information if they consider it to be offensive or distasteful to their (strong or weak) ties. Fourth, even if users are shown novel information, this does not mean they will be exposed to it. They might simply choose to ignore challenging items. Fifth, the information intermediary might filter out the novel content provided by our weak ties. If, for instance, Facebook decides which updates you see on your wall based on the frequency of an interaction, weak ties might as well disappear, as the user will not interact very often with a weak tie. At the moment the only way to prevent this is to manually click on each and every user and choose "show me all updates from this user". Otherwise Facebook will make a decision on what is important based on some unknown criteria.

Conclusion

Gatekeeping theory acknowledges the increasing popularity of social networking, online information seeking and information sharing services. It is often claimed that since users can select and share information online, they can be gatekeepers for each other. This then diminishes the power of media professionals. However, in this paper we have shown that even though the traditional gatekeepers might become less important, users are not becoming the sole gatekeepers. The gates are certainly not disappearing. Platforms on which users operate have an influence; they are one of the new gatekeepers. Online gatekeeping services are not just algorithms running on machines; they are a mix of human editors and machine code designed by humans. People affect the design of the algorithms, but they also can also manually influence the filtering process after the algorithm has been designed. Therefore, switching from human editing to algorithmic gatekeeping does not remove all human biases. Technical biases such as third party manipulation or popularity will exist due to the computerized form of gatekeeping. Also, individual factors such as personal judgments, organizational factors such as company policies, external factors such as government or advertiser requests will still be present due to the role of humans in providing these services.

In this paper, we introduced a model of algorithmic gatekeeping based on traditional gatekeeping model and focused on particular filtering processes including personalization. We show that factors that caused bias in mass media news selection still play a role in information selection in online web services. We have shown that search results in Google can differ, but an extensive empirical research is needed to determine the extent of so-called "echo chambers" in social networks. What percentage of information do users miss or feel like they are missing if they turn on a personal filter or an inter-personal filter? Is there enough variety in their choice of friends? Are users aware of these algorithms? Do they modify their filter periodically or switch to other forms of information sources? Are there routines that are used in the design of personalization algorithms, just like routines used in traditional gatekeeping? How does the introduction of implicit and explicit filtering algorithms affect user trust in systems and user autonomy? More research is needed in order to answer these questions.

Acknowledgments The author would like to thank Martijn Warnier and Ibo van de Poel for their valuable comments. This research is supported by the Netherlands Organization for Scientific Research (NWO) Mozaiek grant, file number 017.007.111.

References

Accuracast. (2010). *Facebook advertising policies homophobic*. May. http://searchdailynews.blogspot.com/2010/05/facebook-advertising-policies.html.

Adomavicius, G., Sankaranarayanan, R., Sen, S., & Tuzhilin, A. (2005). Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems (TOIS), 23*(1), 103–145.

Agada, J. (1999). Inner-city gatekeepers: An exploratory survey of their information use environment. *Journal of the American Society for Information Science, 50*(1), 74–85. http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ582286.

Albanesius, C. (2011). Schmidt, yelp clash over google's search tactics. *PCMAG*. http://www.pcmag.com/article2/0,2817,2393369,00.asp.

Althaus, S. L., & Tewksbury, D. (2002). Agenda setting and the 'new' news. *Communication Research, 29*(2), 180.

Badash, D. (2011). Has facebook censorship gone too far? *The New Civil Rights Movement*. http://thenewcivilrightsmovement.com/has-facebook-censorship-gone-too-far/politics/2011/11/07/29714.

Bagdikian, B. H. (2004). *The New media monopoly: A completely revised and updated edition with seven new chapters*. Beacon Press, May. http://www.amazon.com/dp/0807061875.

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web (WWW '12)* (pp. 519–528). New York, NY, USA: ACM. doi:10.1145/2187836.2187907. http://doi.acm.org/10.1145/2187836.2187907.

Bar-Ilan, J., Keenoy, K., Levene, M., & Yaari, E. (2009). Presentation bias is significant in determining user preference for search results—A user study. *Journal of the American Society for Information Science and Technology, 60*(1), 135–149.

Barzilai-Nahon, K. (2008). Toward a theory of network gatekeeping: A framework for exploring information control. *Journal of the American Society for Information Science and Technology, 59*(9), 1493–1512.

Barzilai-Nahon, K. (2009). Gatekeeping: A critical review. *Annual Review of Information Science and Technology, 43*(1), 1–79.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven: Yale University Press.

Bruns, A. (2008). Gatewatching, gatecrashing: Futures for tactical news media. In M. Boler (Ed.), *Digital media and democracy: Tactics in hard times* (pp. 247–271). MIT Press. http://www.amazon.com/dp/0262026422.

Bruns, A. (2011). Gatekeeping, gatewatching, real-time feedback. *Brazilian Journalism Research, 7*, 117–136

Chatman, E. A. (1987). Opinion leadership, poverty, and information sharing. *RQ*, *26*(3), 53–341. http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ354348.

Chen, C. C., & Hernon, P. (1982). Information seeking: Assessing and anticipating user needs. Neal-Schuman Publishers. http://books.google.nl/books?id=_6fgAAAAMAAJ.

Chibnall S. (1975}. The crime reporter: A study in the production of commercial knowledge. *Sociology, 9*(1), 49–66.

Chilling effects. (2005). Scientology complains that advanced technology appears in Google groups. http://www.chillingeffects.org/dmca512/notice.cgi?NoticeID=2355.

Christman, J. (2011). Autonomy in moral and political philosophy. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford, CA: CSLI, Stanford University.

Cisco. (2011). *Cisco visual networking index: Forecast and methodology, whitepaper*. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html.

Costolo, D. (2011). *The trends are algorithmic, not chosen by us but we edit out any w/obscenities*. July. https://twitter.com/#!/dickc/status/97686216681594880.

Cuilenburg, V. (1999). On competition, access and diversity in media, old and new some remarks for communications policy in the information age. *New Media & Society, 1*(2), 183–207.

Dekker, V. (2006). Google: Een zwijgzame rechter en politieagent. *Trouw*.

Diaz, A. (2008). Through the Google goggles: Sociopolitical bias in search engine design. In S. Amanda., & Z. Michael (Eds.), *Information science and knowledge management* (Vol. 14, pp. 11–34). Berlin Heidelberg: Springer.

Durrance, J. C. (1984). *Armed for action library response to citizen information needs*. New York, NY: Neal Schuman.

Edelman, B. (2011). Bias in search results: Diagnosis and response. *Indian Journal of Law and Technology, 7*, 16.

Efrati, A. (2010). Rivals say Google plays favourites. *Wall Street Journal*, December. http://online.wsj.com/article/SB10001424052748704058704576015630188568972.html.

Elgan, M. (2011). How to pop your Internet 'filter bubble'. *Computerworld*. http://www.computerworld.com/s/article/9216484/Elgan_How_to_pop_your_Internet_filter_bubble_.

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, mis, and related disciplines. *The Information Society, 20*(5), 325–344.

Fallows, D. (2005). *Search engine users*. http://www.pewinternet.org/Reports/2005/Search-Engine-Users/8-Conclusions/Conclusions.aspx.

Fishman, M. (1988). *Manufacturing the news*. Austin: University of Texas Press.

Flanagan, M., Howe, D., & Nissenbaum, H. (2008) Embodying values in technology: Theory and practice. In J. van den Hoven & J. Weckert (Eds.), Information technology and moral philosophy (pp. 322–353). Cambridge: Cambridge University Press.

Fong, J. (2011). Facebook's bias against 3rd party apps. http://www.jenfongspeaks.com/facebooks-bias-against-3rd-party-apps/.

Foundem. 2009. Foundem's Google story.

Friedman, B., Kahn, P. H., & Alan, B. (2006). Value sensitive design and information systems. *Human-Computer Interaction in Management Information Systems: Foundations, 4*, 348–372.

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems, 14*(3), 330–347.

Friedman, B., & Nissenbaum, H. (1997). Software agents and user autonomy. In *Proceedings of the first international conference on autonomous agents—AGENTS'97*, pp. 466–469.

Gans, H. J. (2005). *Deciding what's news: A study of CBS evening news, NBC nightly news, newsweek, and time* (2nd ed.). Evanston: Northwestern University Press.

Garcia-Molina, H., Koutrika, G., & Parameswaran, A. (2011). Information seeking. *Communications of the ACM, 54*(11), 121. doi:10.1145/2018396.2018423.

Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007). *User profiles for personalized information access. The adaptive web* (pp. 54–89). Berlin Heidelberg: Springer.

Gillespie, T. (2012). Can an algorithm be wrong? *Limn* (2). http://limn.it/can-an-algorithm-be-wrong/.

Goldman, E. (2005). Search engine bias and the demise of search engine utopianism. *Yale JL & Technology, 8*, 188.

Goldman, E. (2011). Revisiting search engine bias chapter in (Contemporary Issues in Cyberlaw), *William Mitchell Law Review, 38*, 96–110.

Google. (2008). *We knew the web was big*. http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html.

Google. (2011). *Transparency report*. http://www.google.com/transparencyreport/governmentrequests/.

Google. (2012). *Search plus your world: Personal results*. http://support.google.com/websearch/bin/answer.py?hl=en&answer=1710607.

Granka, L. A. (2010). The politics of search: A decade retrospective. *The Information Society, 26*(5), 364–374. doi:10.1080/01972243.2010.511560.

Granovetter, M. S. (1981). *The strength of weak ties: a network theory revisited*. State University of New York, Department of Sociology.

Groot, J. (2004). Trouw wekenlang niet te vinden op Google. *Webwereld*.

Guha, S., Cheng, B., & Francis, P. (2010). Challenges in measuring online advertising systems. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement (IMC '10)* (pp. 81–87). New York, NY, USA: ACM. doi:10.1145/1879141.1879152. http://doi.acm.org/10.1145/1879141.1879152.

Helberger, N. (2011). Diversity by design. *Journal of Information Policy, 1*, 441–469.

Hermida, A. (2012). Tweets and truth: Journalism as a discipline of collaborative verification. *Journalism Practice, 6*(5-6), 659–668.

Hilbert, M. (2012). Toward a synthesis of cognitive biases: How noisy information processing can bias human decision making. *Psychological Bulletin, 138*(2), 211–237.

Hitwise. (2010). Social networks now more popular than search engines in the UK.

Hoven, J. V., & Rooksby, E. (2008). *Distributive justice and the value of information: A (broadly) Rawlsian approach*. England: Cambridge University Press.

IBM. (2011). Bringing smarter computing to big data.

Ingram, M. (2011). *The downside of facebook as a public space: Censorship*. June. http://gigaom.com/2011/06/21/the-downside-of-facebook-as-a-public-space-censorship/.

Jacobs, G. (2010). Techradar: How to optimise your site for Google Caffeine. *Techradar.com*, April. http://www.techradar.com/news/internet/how-to-optimise-your-site-for-google-caffeine-685436.

Joachims, T., & Radlinski, F. (2007). Search engines that learn from implicit feedback. *Computer, 40*(8), 34–40.

Katz, E. (1996). And deliver us from segmentation. *Annals of the American Academy of Political and Social Science, 546*, 22–33.

Katz, E., & Lazarsfeld, P. (2005). *Personal influence: The part played by people in the flow of mass communications*. New Jersey: Transaction Publishers.

Kincaid, J. (2010). Techcrunch/today's lesson: Make facebook angry, and they'll censor you into oblivion. *TechCrunch*. http://techcrunch.com/2010/11/22/facebook-censorship/.

Klein, J. (2011). A web marketer's guide to reddit. December. http://www.distilled.net/blog/social-media/a-web-marketers-guide-to-reddit/.

Knight, W. (2012). Google hopes to make friends with a more social search: technology review. *Technology Review*. http://www.technologyreview.com/computing/39444/.

Korolova, A. (2010). Privacy violations using microtargeted ads: A case study. In *Proceedings of the IEEE international conference on data mining workshops (ICDMW '10)* (pp. 474–482). Washington, DC, USA: IEEE Computer Society. doi:10.1109/ICDMW.2010.137. http://dx.doi.org/10.1109/ICDMW.2010.137.

Lasorsa, D. L., Lewis, S. C., & Holton, A. (2012). Normalizing Twitter-Journalism practice in an emerging communication space. *Journalism Studies, 13*(1), 19–36.

Lavie, T., Sela, M., Oppenheim, I., Inbar, O., & Meyer, J. (2009). User attitudes towards news content personalization. *International Journal of Human-Computer Studies, 68*(8), 483–495.

Levinson, P. (1999). *Digital McLuhan: A guide to the information millennium* (1st ed.). London: Routledge.

Lotan, G. (2011). Data reveals that "occupying" twitter trending topics is harder than it looks! http://blog.socialflow.com/post/7120244374/data-reveals-that-occupying-twitter-trending-topics-is-harder-than-it-looks.

Lu, Y. (2007). The human in human information acquisition: Understanding gatekeeping and proposing new directions in scholarship. *Library & Information Science Research, 29*(1), 103–123.

Manyika, J., Chui, M., Brown, B., Buighin, J., Dobbs, R., & Roxburgh, C. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute report*. Whitereport. Downloadable at http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

Metz, C. (2011a). Google opens curtain on 'manual' search penalties. *The register*. http://www.theregister.co.uk/2011/02/18/google_on_manual_search_penalties/.

Metz, C. (2011b). Google contradicts own counsel in face of antitrust probe, admits existence of search algorithm whitelists. http://www.theregister.co.uk/2011/03/11/google_admits_search_algorithm_whitelists/.

Morozov, E. (2011). Your own facts. Book review of 'the filter bubble, what the internet is hiding from you'. *The New York Times*.

Mowshowitz, A., & Kawaguchi, A. (2002). Bias on the web. *Communications of the ACM, 45*(9), 56–60.

Munson, S. Z., & Resnick, P. (2010). Presenting diverse political opinions: How and how much CHI'10. In *Proceedings of the SIGCHI conference on human factors in computing systems.*

Nagesh, G. (2011). Privacy advocates want facebook probed on recent changes.

Napoli, P. (1999). Deconstructing the diversity principle. *Journal of Communication, 49*(4), 7–34.

Nissenbaum, H., & Introna, L. D. (2000). Shaping the web: Why the politics of search engines matters. *The Information Society, 16*(3), 169–185.

O'Dell, J. (2011). Facebook's ad revenue hit $1.86B for 2010. *Mashable*.

Opsahl, K. (2009). Google begins behavioural targeting ad program. https://www.eff.org/deeplinks/2009/03/google-begins-behavioral-targeting-ad-program.

Pariser, E. (2011a). 10 ways to pop your filter bubble. http://www.thefilterbubble.com/10-things-you-can-do.

Pariser, E. (2011b). *The filter bubble: What the internet is hiding from you*. London: Penguin Press.

Priestley, M. (1999). Honest news in the slashdot decade. *First Monday, 4*, 2–8.

Resnick, P., Lacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on computer supported cooperative work (CSCW '94)* (pp. 175–186). New York, NY, USA: ACM. doi:10.1145/192844.192905. http://doi.acm.org/10.1145/192844.192905.

Salihefendic, A. (2010). How reddit ranking algorithms work. http://amix.dk/blog/post/19588.

Schroeder, S. (2011). Twitter ad revenue may reach $150 million this year. *Mashable*.

Searchenginewatch. (2012). Twitter: Google search plus your world bad for web users. *Search Engine Watch*. http://searchenginewatch.com/article/2136873/Twitter-Google-Search-Plus-Your-World-Bad-for-Web-Users.

Segal, D. (2011). Search optimization and its dirty little secrets. *The New York Times*.

Shardanand, U., & Maes, P. (1995). Social information filtering: algorithms for automating "word of mouth". In I. R. Katz, R. Mack, L. Marks, M. B Rosson & J. Nielsen (Eds.), *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '95)* (pp. 210–217). New York, NY, USA: ACM Press/Addison-Wesley Publishing Co. doi:10.1145/223904.223931. http://dx.doi.org/10.1145/223904.223931.

Shoemaker, P. J, Vos, T., & Reese, P. (2008). Journalists as gatekeepers. In K. W. Jorgensen., & T. Hanitzsch (Eds.). *The handbook of journalism studies* (pp. 73–87). New York: Routledge

Shoemaker, P. J., & Vos, T. (2009). *Gatekeeping theory* (1st ed.). London: Routledge.

Slater, P. E. (1955). Role differentiation in small groups. *American Sociological Review, 20*(3), 300–310.

Smith, J., McCarthy, J. D., McPhail, C., & Augustyn, B. (2001). From protest to agenda building: Description bias in media coverage of protest events in Washington, D.C. *Social Forces, 79*(4), 1397–1423.

Smyth, B. (2007). A community-based approach to personalizing web search. *Computer, 40*(8), 42–50. doi:10.1109/MC.2007.259.

Soley, L. C. (2002). *Censorship Inc: The corporate threat to free speech in the United States*. USA: Monthly Review Press.

Sturges, P. (2001). Gatekeepers and other intermediaries. *Aslib Proceedings, 53*(2), 62–67.

Sullivan, D. (2012). Google's results get more personal with 'search plus your world'. *Search engine land*. http://goo.gl/xYoRV.

Sunstein, C. R. (2002). *Republic.com*. USA: Princeton University Press.

Sunstein, C. (2006). Preferences, paternalism, and liberty. *Royal Institute of Philosophy Supplements, 59*, 233–264.

Sunstein, C. R. (2008). *Infotopia: How many minds produce knowledge*. USA: Oxford University Press.

Taylor, D. (2011). Everything you need to know about facebook's edgerank. *The Next Web*. http://thenextweb.com/socialmedia/2011/05/09/everything-you-need-to-know-about-facebooks-edgerank/.

Techcrunch. (2011). Edgerank: The secret sauce that makes facebook's news feed tick.

Tewksbury, D. (2003). What do Americans really want to know? Tracking the behavior of news readers on the internet. *Journal of Communication, 53*(4), 694–710.

Twitter. (2010). *To trend or not to trend*. http://blog.twitter.com/2010/12/to-trend-or-not-to-trend.html.

Tynan, D. (2012). How companies buy facebook friends, likes, and buzz. *PCWorld*.

Upbin, B. (2011). Facebook ushers in era of new social gestures—Forbes. *Forbes*.

US Securities and Exchange Commission. (2009). Google Inc., Consolidated Balance Sheets. http://www.sec.gov/Archives/edgar/data/1288776/000119312509150129/dex992.htm.

Van Couvering, E. (2007). Is relevance relevant? Market, science, and war: discourses of search engine quality. *Journal of Computer-Mediated Communication, 12*(3), 866.

Van der Hof, S., & Prins, C. (2008). *Personalisation and its influence on identities, behaviour and social values. Profiling the European citizen: Cross-disciplinary perspectives* (pp. 111–127). Netherlands: Springer.

Vaughan, L., & Thelwall, M. (2004). Search engine coverage bias: Evidence and possible causes. *Information Processing and Management, 40*(4), 693–707.

Witten, I. A. (2007). Bias, privacy, and personalization on the web. In M. Sankara Reddy & H. Kumar (Eds.), *E-libraries: Problems and perspectives*. New Delhi: Allied.

Wittman, C. (2011). Comments 4x more valuable than likes. http://goo.gl/wnSES.

Wright, J. D. (2011) Defining and measuring search bias: Some preliminary evidence. International center for law & economics, November 2011; George Mason Law & Economics Research Paper No. 12–14. Available at SSRN: http://ssrn.com/abstract=2004649.

Yu, C., Lakshmanan, L., & Amer-Yahia, S. (2009). It takes variety to make a world: Diversification in recommender systems. In *Proceedings of the 12th international conference on extending database technology: Advances in database technology* (pp. 368–378). http://dl.acm.org/citation.cfm?id=1516404.

Yue, Y., Patel, R., & Roehrig, H. (2010). Beyond position bias: Examining result attractiveness as a source of presentation bias in click through data. In *Proceedings of the 19th international conference on World wide web* (pp. 1011–1018). http://dl.acm.org/citation.cfm?id=1772793.

Zhang, M., & Hurley, N. (2008). *Avoiding monotony: Improving the diversity of recommendation lists, 2008 ACM international conference on recommender systems (ACM Recsys'08)* (pp. 123–130). Switzerland: Lausanne.

Zimmer, M. (2011). Facebook's censorship problem. June. http://www.huffingtonpost.com/michael-zimmer/facebooks-censorship-prob_b_852001.html.