# Principles of Information Neutrality and Counter Measures Against Biased Information

Valentin Burger[1], Matthias Hirth[1], Tobias Hoßfeld[2] and Phuoc Tran-Gia[1]

**Abstract:** The Internet serves as ubiquitous, prompt source of information in our society. Due to the giant amount of information, the access to the information relies on information processing mechanisms as implemented in search engines, product recommendation systems, or online social networks. As a result of the engineered preprocessing mechanism, the retrieved information is biased and does not represent a neutral view on the available information, while end-users often are not aware of this bias. In this article, we define the term "Information Neutrality", review current principles of information processing in the Internet and discuss influence factors hindering Information Neutrality as well as appropriate countermeasures. The main contribution of this article is to raise the awareness of Information Neutrality as an emerging key challenge in the Internet and to potentially consider Information Neutrality as a factor in the value-sensitive design of Internet-based services.

**Keywords:** Information Neutrality, information filters, recommendation, filter bubble, information bias, personalization.

## 1 The Internet as Information Source

In today's Internet users face the problem to cope with a too large amount of information. Current developments, like the increasing amount of location based data in 5G, sensor networks in smart cities and on wearable devices, and the ability to store and process data using big data techniques, eventually increases the volume of information available. The huge mass of information makes it nearly impossible to find items, which are personally most relevant. Back in October 2012, already more than one billion people were active on Facebook. Imagine how to find potential friends out of one billion people without any specialized means.

To make the selection easier for end-users, mechanisms are needed that generate manageable subsets from this large mass. To provide every individual user with information that is of interest, the subsets must contain suggestions that are tailored on the users' preferences. The volume of the subset has to be reduced and thus information may be aggregated. However, there is an obvious trade-off between the degree of aggregation and the information contained. Hence, there is a need for systems assisting the user to find relevant information, e.g., by recommendations.

[1] University of Würzburg, Institute of Computer Science, Chair of Communication Networks, Germany {forename.surname}@uni-wuerzburg.de
[2] University of Duisburg-Essen, Chair of Modeling of Adaptive Systems, Essen, Germany tobias.hossfeld@uni-due.de

Systems assisting users to find relevant information often base on the users past behaviour, which causes the user to end in a positive feedback loop based on its history. It gets hard for the user to find new eye-opening information, the chance for serendipity disappears and the user gets stuck in a so called “filter bubble” [Pa11]. This means web services do not treat users neutrally, but recommend only information tailored on the user. To address this issue, we define the term “Information Neutrality”.

> *Information Neutrality is the principle to treat all information provided (by a service) equally. The information provided, after being processed by an information neutral service, is the same for every user requesting it, independent of the user’s attributes, including, e.g., origin, history or personal preferences and independent of the financial or influential interest of the service provider, as well as independent of the timeliness of information.*

Sociologists are worried about the limited scope of users in the Internet. Books like “Filter Bubble” [Pa11] or “Siren Songs and Amish Children” address this problem of Information Neutrality. Leading Internet companies already react to these concerns. Companies like Google or Facebook improve their algorithms to avoid filter bubbles or provide means to access neutral information. Nevertheless, user profiles and tailored information will always be present, since customer oriented ads are very profitable.

Computer scientists as well as software engineers and companies have ethical responsibility developing and designing information neutral algorithms. Supported by recommendations, the large mass only propagates emotional issues and topics, suppressing serious topics. Appropriate algorithms can help to identify serious topics or counter discrimination by also showing contributions of minorities.

Examples for such mechanisms are already developed in related work. In [Ka12] an algorithm is developed that enhances neutrality of recommendations towards a viewpoint specified by a user. This is achieved by prohibiting the use of some information for the purpose of making recommendation by privacy policies. In [Ma12] semantic-based recommendation systems are used to increase the serendipity in different use-cases by enriching the data with user activities, interests and other meta data for a better alignment of user profiles. [Ng14] measure the filter bubble effect in terms of content diversity received by individual end-users and determine the impact of collaborative filtering-based recommender system, showing that recommendations have a positive effect on the user experience.

The bias on information processed by computer systems has already been analysed in the past [FN96]. In [IN00] the bias of search engines is investigated, showing that certain sites are favoured, which narrows the scope of the web and counters its values such as unrestricted access and widespread information. Investigating the bias of such mechanisms is especially important today, due to the vast amount of data and widely used and more sophisticated algorithms in today’s Internet.

Developers need to provide interfaces to enable the user to decide on the amount of tailored information confronted. This allows end-users to determine the degree of Information Neutrality and to adjust the trade-off between tailored information vs. neutral information.

In this work, we identify factors that influence Information Neutrality and present techniques that help end-users escaping the filter bubble and guidelines for providers to improve their service to be information neutral. Finally, we discuss key questions and challenges that have to be faced to achieve a good trade-off between privacy and usability of services.

## 2 Definition of Neutrality in the Internet

To get a clear definition of Information Neutrality we need to distinguish it from different kinds of neutrality concerning the Internet. Please note that we do not intend to evaluate the existing neutrality concepts, because there are still ongoing debates whether (and which kind of) neutrality is beneficial for which stakeholders. Neutrality in the Internet is often used with different meanings. In order to clarify our understanding of neutrality in the Internet, we distinguish the following four different types:

1. Network Neutrality also known as Net or Internet Neutrality: Equal treatment of all Internet traffic.
2. Transport Neutrality: Message exchange independent of transport protocol.
3. Search Neutrality: Even-handed treatment in search results.
4. Information Neutrality: Equal treatment of information (includes Search Neutrality).

Strict Network Neutrality "prohibits Internet service providers from speeding up, slowing down or blocking Internet traffic based on its source, ownership or destination." Network Neutrality ensures that users and applications are treated equally in a way that no user is discriminated depending on its origin or network usage. Hence, the access to the Internet is not limited depending on the location of the connection and on the application or platform used. [BEa] shows that Network Neutrality is violated by all providers in Europe by prioritizing throttling and blocking traffic. Peer-to-peer traffic, for instance, is throttled or blocked in many transport networks and Voice-over-IP traffic is blocked in certain mobile networks. Hence, non-profitable traffic, or traffic that produces transit costs or that is produced by applications that replace the services offered by providers, is blocked by providers. Throttling and blocking is commonly realized by deep packet inspection. A network neutral service forwards bits with equal priority, no matter to which flow, user, or operator they belong. An example for a mechanism that enables Network Neutrality is the transmission control protocol (TCP). The available bandwidth on a link is shared equally among concurring flows. However, emerging technologies like OpenFlow and software defined networking aim to dynamically control and prioritize flows by software. This new paradigm will drastically limit Network Neutrality, but will also open new business models and efficient resource allocation and utilization. Network Neutrality is also a part of

Internet governance, which tries to develop shared norms and rules that regulate the usage of the Internet among governments that can have very different interests.

Transport Neutrality is a concept that allows message exchange of web services independent of the transport. This is realized by a distinction between payload and headers, so that application protocol information can be captured without creating dependencies on transport protocols. Sender and receiver have to share an interpretation of headers placed in the message. For example, web services using the Simple Object Access Protocol (SOAP) are transport neutral, since the SOAP envelope provides a framework for separating a payload from accompanying headers. In contrary to transport and Network Neutrality that consider the data transmission, Information Neutrality considers the bias put on information by presenting information filtered based on user profile and personal preferences, history of user, location of user, timelines, etc.

An existing concept is Search Neutrality that considers even-handed treatment in search results. Search Neutrality is defined as “the principle that search engines should have no editorial policies other than that their results be comprehensive, impartial and based solely on relevance.” In contrast to Search Neutrality, Information Neutrality requires neutral information while consuming a service without explicitly searching for information, e.g. feeds in on news aggregators, recommendation of new friends, or hints on related products.

The system bias introduced by big providers of information systems poses a threat on equal treatment of users, businesses, and opinions in the Internet. Service providers determine what users read, consume, use, and purchase online. Hence, service providers have a huge influence, which can be abused to shape public opinions and to gain competitive advantages. Hence, this also has an impact on business and support monopoly. To maintain a fair market and free opinion making Information Neutrality matters. Regarding this, the European Commission has already raised an issue concerning the anti-competitive behavior by Google in changing the shopping search results for price comparison from a vertical service to directly monetized ads.

To be applicable to recent developments and applications in the Internet, the common layer model has to be extended considering the user and its social interactions as shown in in Figure 1. Depicted is the network stack on layers 1 to 7 as defined in the OSI model. Layers 8 and 9 are added corresponding to end-users and the social network. Layer 8 considers communication among end-users and provides interfaces to end-user as the Quality of Experience perceived or the personal preferences as well as personal information provided. The social interaction between end-users and their organization in social networks provides a vast amount of information, which is subsumed in layer 9. Structural information reveals the organization of end-users in communities and the formation of clusters where information flows within as well as across communities. Interfaces to the social network layer 9 are necessary to get access to the information provided, which is highly valuable for emerging business models. The different types of neutrality identified touch different layers of this extended model. Depending on the considered scope of Network Neutrality it can concern everything from the physical layer to the user layer, corresponding to equal treatment of bits and to equal treatment of flows of different users. Information Neutrality considers equal treatment of information provided in social networks, filter bubbles due

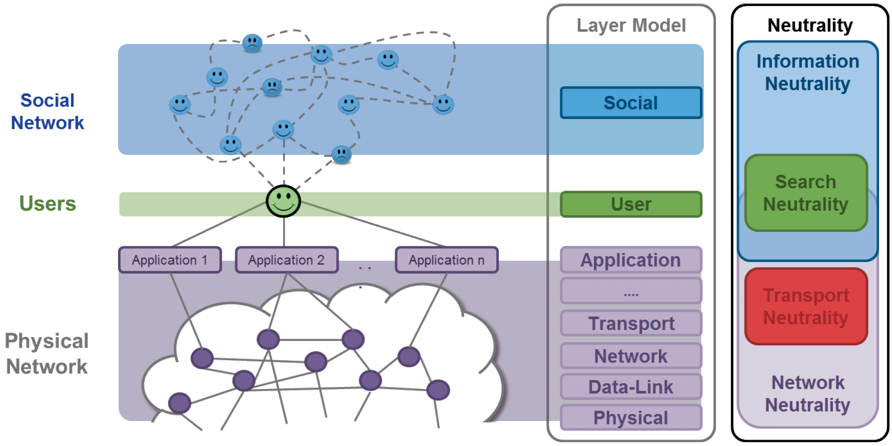


Fig. 1: Layer Model and Neutrality in the Internet.

to tailored information based on user profiles as well as the bias put on information by application interfaces.

## 3 Information Processing Mechanisms

There are different information processing mechanisms, which influence the Information Neutrality derived in the Internet.

The most basic mechanism, which leads to non-neutral information, is *information filtering*. In general, it is used to derive a small manageable subset from a very large set, which could not be processed in a decent amount of time. There are different policies to filter information, for example on most websites content can be filtered by popularity or recency. By confronting the user with filtered information, other topics are dashed and a bias is put on the information presented. Hence, the information the user derives is non-neutral. Simple mechanisms are collaborative and content based filters. An example for collaborative filtering is Amazon's product recommendation: "Customers Who Bought Items in Your Recent History Also Bought". It relies on the assumption that there is a high probability that a user who bought a specific product is also interested in products other users bought, who bought the same product. Content based filtering differs from collaborative filtering in using attributes of the items instead of using the collaborative intersection.

Secondly, *tailored information* is filtered information tailored to the user's profile. For example, to increase the click-through rate on advertisements and products, user profiles are set up, to identify the users' interest and to provide users information in which they are likely interested. The problem is that the information provided is based on the user profile, which is a reflection of the user's past activity. The information seen is similar to the information consumed in the past. This prevents discovering new unexpected topics and reduces the chance to find interesting.

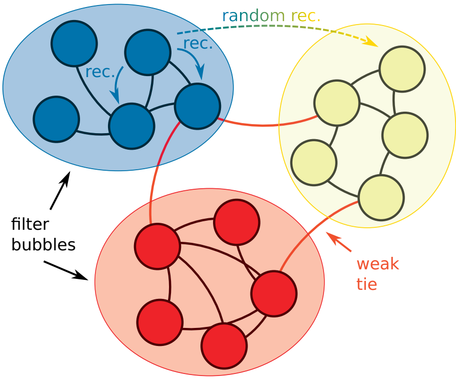


Fig. 2: Filter bubbles in a network with recommended items.

Thirdly, there are means to provide the user with an *aggregation/condensation of information*, e.g., summaries of different blogs, articles, products, which help him to get an overview of a large collection of objects belonging to a topic that he can process in an affordable amount of time. Given a large set of similar information objects, these techniques summarize the essence or locate the source of information. However, interesting additional information, e.g. details, thoughts or comments on a discussed topic, might be suppressed.

Finally, there are *recommendations*. In recommendation systems, attributes of users, their history, as well as recency and the network structure are used to predict which friend, product, video the user is going to connect to/consume next. Hence, a subset of objects is selected based on all considered attributes of the user and the surrounding network. This puts a bias on the recommended objects and therefore fades out different objects. Further, it supports segmentation of communities by strengthening ties between objects of shared interest. It also reduces the potential of building bridges to different communities, since until now, most recommendation metrics foster strongly connected components.

Consider for example the mechanism recommending friends on Facebook. The algorithm behind is based on a supervised random walk [BL11] that aims to predict which links in the networks are going to evolve in the near future. A rich set of attributes is used to guide a random walker on the network graph by adding strength to the edges. The strength of the edge determines the probability of the random walker to follow an edge. A learning task with the goal to find a function assigning strengths to the edges is used to make the random walker visit nodes more likely that the user is going to interact with. The underlying optimization problem cannot be solved exactly, because of the large amount of users in the network. Therefore, approximations are used and only a subset of nodes can be considered. In the case of Facebook only friends-of-friends are considered, which count already 40k on average. Figure 2 shows a network with different items, which can be persons, products, news articles. Links connect associated items, for example, by having the same interest or being purchased together. In the network are different clusters with items that are similar and are relevant with high probability. If a mechanism only recommends similar items that are directly connected to the initial item, it is hard to escape these filter bubbles. If random items are recommended, they might be in another cluster and the filter bubble can be escaped. Another possibility to escape the filter bubble is to recommend items connected to weak ties. Weak ties are links that connect different clusters. According to [Gr73] most interesting information is derived via weak ties, hence from people of different communities, which are only slightly in contact.

## 4 Non-Algorithmic Influence Factors on Information Neutrality

Besides the different information processing mechanisms, there are also non-algorithmic influence factors on Information Neutrality.

### 4.1 Application Interfaces

A technical, yet not algorithmic, influence factor on Information Neutrality are application interfaces. Application interfaces can have an influence on Information Neutrality, because they determine how the information is presented to the user. A simple example is the formatting of headlines in newspapers. Articles with large and bold headline immediately draw more attention even if they might not be as relevant and interesting to the readers as other articles with headlines that are not so prominent. This immediately puts a bias on how the available information is perceived. In an online scenario, the arrangement of search results and advertisements determines what a user is likely to click. Google places customer oriented ads as first results on its sites, because users click on them with high probability. To protect customers, Google had to change this and make advertisements clearly recognizable. The number of results per page can also have an influence on what a user consumes. Recent work shows that most users only click on first three pages of Google Search. Top sites like YouTube, Facebook or Amazon frequently change their interface. For instance, the interface of YouTube changed from the subscriptions being the front-page to recommended videos being on the front-page. The videos seen on first sight are no longer the videos from subscribed channels, but also recommended videos based on your profile.

### 4.2 Social Relationships and Proximity

Social networks use recommendation systems not only to suggest friends, but also to decide which posts users are confronted on first sight. Different topics, which are not presented on the front-page, are suppressed. Facebook for example filters the news on the front-page. Algorithms calculate and decide which of your friends or stories are interesting for you and show only these activities. This service helps to manage the large mass of posts by friends and their activities. However, if the user does not claim interest to a person, it might disappear from the feed and the user might totally forget about her or him. Further, users are only exposed with impressions of their social environment, which hardens their point of view and offers less space for revising their opinions. This can also harden fronts between opposing parties. Further on, emotional issues and topics are more likely to be propagated by the large mass, suppressing serious topics.

Proximity has a high impact on our social relationships and the communities we join determine our environment. Users are interested in people and events that are locally close. Hence, there is also a bias on these topics. The authors of [LYW15] call this phenomenon the "majority illusion" and show that behavior that is globally rare may be systematically overrepresented in the local neighborhoods of many people, i.e., among their friends.

### 4.3 External Influence Factors

Information reaches the users from both, mass media and from personal social network, meaning the friends that the users meet in person. Myers et al. [MZL12] investigated information diffusion of the OSN Twitter, they find that "only about 71% of the information volume in Twitter can be attributed to network diffusion, and the remaining 29% is due to external events and factors outside the network." That means that about one third of the information users get is from external influences.

## 5 Guidelines Towards Neutral Information

Both service providers and users can counter tailored information. Service providers can add means to disable information filters or make them adjustable to the personal preferences. End-users have several possibilities to overcome tailored information, even if providers do not directly support this. In the following, first, guidelines for end-users and second, guidelines for service providers are described.

Tailored information can be avoided and recommendations can just be ignored. A user can visit an alternative web site handling information in a more neutral manner. Being aware of weak ties, a user can target information sources accordingly. Such sources can be Facebook walls of friends that are not in the everyday clique, that you met while travelling, that are of different culture or religion. However, for most users it is hard to assess, which web service is most neutral.

Internet and services can be accessed anonymously to prevent recommendation systems to set up user profiles. One prominent tool providing anonymous access is "Tor", which is a network of virtual tunnels developed with the U.S. Navy. Its primary purpose is protecting government communications. One of its use cases is preventing websites from learning location or browsing habits.

The virtual tunnels are routed via three proxies, so called relays, that change for every connection. Each relay has a different key for decryption and removes the corresponding decryption layer of the message. The layered encryption referred to as onion routing. This hides the origin and prevents the traceability of the request, but also adds a delay depending on the location of the relay, c.f. Figure 3, which can be a problem for non-delay tolerant

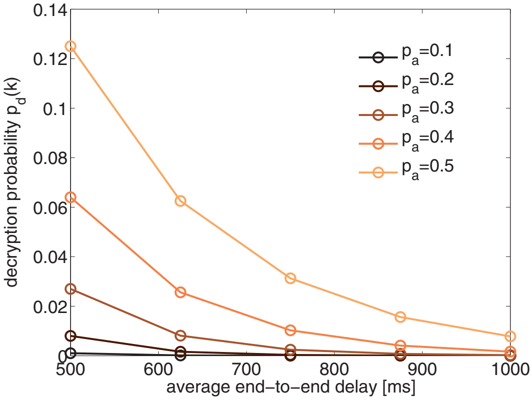


Fig. 3: Trade-off between decryption probability and average end-to-end delay for different attack probabilities. If the attack probability $p_a$ is low, the decryption probability is low independent of the end-to-end delay. If $p_a$ is high, a high delay has to be tolerated for a secure connection. Hence, anonymity can either be bought with waiting time or by deploying a high number of relays.

applications. An attacker can still decrypt the message, if it controls all relays on the route. So adding more relays on the route increases the probability that at least one relay is not controlled by the attacker.

Another mean towards information neutrality considers tracking cookies. In order to track users, web services store cookies, with user IDs on your local machine. These tracking cookies can automatically be blocked or deleted to avoid tracking. Tools like "NoScript"[5] or "AdBlock"[6] can block scripts on websites that read the cookie to forward the information. Modern browsers offer a "Private Browsing" function, which allows running a clean browsing session that does not store cookies or deletes them after closing the session. However, there are very persistent cookies like "evercookie"[7] that store the cookie data in several types of storage mechanisms and reacreate the cookie, if any of the stored cookie types has been removed. A website that lets you check how safe your browser is against tracking is provided by "Panopticlick"[8].

In general, to get out of the filter bubble one can simply look beyond the front page of websites. An example is the Facebook front page. The default setting for the Facebook news feed is "Top Stories". The "Top Stories" presented are selected by a filtering algorithm according to the user's interests, recent activity, and contacts. To get rid of this filter the setting can be switched to "Top Stories", which shows the unfiltered and neutral version of the news feed. In [LYW15] an interactive method to visualize the personalized filtering is presented, which provides awareness of the filter bubble in online social networks.

[5] https://noscript.net/
[6] https://adblockplus.org/
[7] https://samy.pl/evercookie/
[8] https://panopticlick.eff.org/tracker

It would not be necessary to use tools like "TrackMeNot"[9] if service providers add means to disable information filters or make them adjustable to fit the personal preferences. In the following we give examples for guidelines that can be adopted by service provider for fostering neutral information.

A profile is visible as soon as some queries concerning a certain topic are used more frequently. Tools like "TrackMeNot" inject random search queries, so that every topic has equal weight, to destroy the user profile. A service provider can add random recommendations to provide neutral information. In the end there will be always user profiles constructed to increase profits of the commercial industry. For instance online social network users' web histories are tracked even if they are not logged in. This is realized by tracking cookies that are integrated in the popular share-buttons. These share-buttons are distributed on websites all over the Internet.

Service providers can use more sophisticated recommendation mechanism. In [Zh12] a tool for music recommendation "Auralist" is developed, which aims to mimic the actions of a trusted friend or expert. Such accuracy, diversity, novelty and serendipity of the recommendations shall be improved simultaneously and balanced. Therefore, they investigate three algorithms, namely Latent Dirichlet Allocation, Listener Diversity and Declustering. The algorithms are briefly described in the following.

Latent Dirichlet Allocation clusters users with similar preferences. Composition vectors define the listener base of each artist. A similarity metric for the composition vectors is then used to calculate recommendations for a user. Listener Diversity promotes artists with diverse listener communities. Such users are encouraged to explore beyond a given niche. Declustering directly addresses Information Neutrality by identifying "music bubbles" in form of similar artist clusters and counteracts them by recommending least clustered or "boring" items for a user. The result of the combination of all three algorithms especially enhances serendipity in the recommended music.

Soylent [Beb] is an extension for Microsoft Word that uses a crowdsourcing platform. One of the functionalities of Soylent is text shortening, which can be used to filter the most valuable information out of a document. The key concept of Soylent is to divide the text-shortening task in different subtasks that can be submitted to a crowdsourcing platform. The subtasks are identifying paragraphs that can be shortened, shortening the paragraphs and proofreading the shortened paragraphs. A number of different workers perform each step to crosscheck results to assure a certain quality of the condensed text.

Summly is a news app for smartphones that condenses news to 400 characters in readable sentences, such that it keeps the main message. The algorithm identifies, extracts, and combines the most important sentences of the article. The shortened news text with less than 400 characters fits on a smartphone display and contains enough information to get a quick overview. A longer summary or the original article is linked for detailed information. Summly uses a genetic algorithms and machine learning for text summarization. The algorithm learns from the well-formed structure of news articles.

[9] https://cs.nyu.edu/trackmenot/

## 6 Key Questions and Challenges

Information neutrality is important for end-users as well as service providers. Its characteristics are dependent on every single stakeholder. First, convenience and the possibility to get a quick and good selection from a vast set of items in a short time are important for end-users. But, end-users need to be aware that the information presented is filtered and that there is a bias on the results. The question is, to which extent information is filtered and how to show the bias and that the information is filtered?

Second, it is also important for end-users to discover new items and topics to experience serendipity. This can be accomplished by better mechanisms or random results. Hence, the question is how to improve mechanisms and which amount of random results to add?

Further, users must be able to decide on how tailored the information presented is. A user might just want to access neutral results, to get the unbiased view. Therefore means, like sliders that range from “neutral” to “personal” could be provided by service providers, so that the user can adjust according to its personal preferences. The question is how to measure the degree of information neutrality and how to provide means to let the user decide on the degree presented?

Metrics are necessary in order to measure the degree of information neutrality. A problem to define appropriate metrics is to define which information can be considered as totally neutral. A value determining the information neutrality of a service could for example be measured by the amount of data used from user profiles to process the information.

Users have to be aware that the personal information they provide is of high value for service provider. Service providers may not abuse the capital they get, which is the information base collected from all users. The information base can be used for opinion shaping as well as provisioning of non-competitive services. Public opinion has an impact especially on democratic societies. The information base and the access to many users can be abused to form the opinion according to the interest of the company. Provisioning of non-competitive services helps huge companies to further strengthen their position. This threatens the free market and will lead to monopolies of the giants. But companies can also take the chance and use their means to identify serious topics or counter discrimination by also showing contributions of minorities. The challenge is to provision fair, neutral and inspirational services, Companies as well as developers of information systems need to be aware of that. Information neutrality has to be considered by companies, organizations and government to establish policies.

## 7 Conclusion

Like all commercial companies, information providers seek to maximize their revenue. For web services, the revenue is often tightly connected to the click through rates on the offered products or content. Good recommendations and customized content can increase the click through rate of less than 1% to more than 5%. To enable customized content, user data has to be accumulated. However, users of web services do not know and cannot control what is collected. It is not clear how the collected data is processed and exploited by companies and how much they profit from it. Hence, there is a value of the personal data we provide, which is not rewarded by service providers.

The massive data accumulation may not only have severe consequences on information neutrality, but also on the privacy of the user. The information collected by recommendation systems is personal and can be highly sensitive. To protect the users' privacy it has to be clear, which personal data may be collected, who may access personal data and how it may be used.

Algorithmic driven user profiles limit the information neutrality as they preserve models of every user. People change and so do the models have to change. Falsifiability of the models has to be considered to train them efficiently. Web services must experiment with their users and present them new and random content to even refine the users profile, keep it up to date and help the user encounter serendipity.

There is a trade-off between the privacy and serendipity level of a service and its usability. Strong user profiles and tailored recommendations lack privacy and serendipity, but provide a good set of results quickly. Interfaces have to be provided that make the bias on the information visible to the user and let the user decide on the degree of neutrality in the information provided. It is part of future work to investigate how to enable user empowerment in web services.

## References

[BEa] BEREC: , A view of traffic management and other practices resulting in restrictions to the open Internet in Europe, 2012.

[Beb] Bernstein, Michael: , It's People! Meet Soylent, the Crowdsourced Copy Editor.

[BL11] Backstrom, Lars; Leskovec, Jure: Supervised random walks: predicting and recommending links in social networks. In: Proceedings of the fourth ACM international conference on Web search and data mining. ACM, pp. 635–644, 2011.

[FN96] Friedman, Batya; Nissenbaum, Helen: Bias in computer systems. ACM Transactions on Information Systems (TOIS), 14(3):330–347, 1996.

[Gr73] Granovetter, Mark S: The strength of weak ties. American journal of sociology, pp. 1360–1380, 1973.

[IN00] Introna, Lucas D; Nissenbaum, Helen: Shaping the Web: Why the politics of search engines matters. The information society, 16(3):169–185, 2000.

[Ka12] Kamishima, Toshihiro; Akaho, Shotaro; Asoh, Hideki; Sakuma, Jun: Enhancement of the Neutrality in Recommendation. In: DecisionsRecSys. pp. 8–14, 2012.

[LYW15] Lerman, Kristina; Yan, Xiaoran; Wu, Xin-Zeng: The majority illusion in social networks. arXiv preprint arXiv:1506.03022, 2015.

[Ma12] Maccatrozzo, Valentina: Burst the filter bubble: using semantic web to enable serendipity. In: The Semantic Web–ISWC 2012, pp. 391–398. Springer, 2012.

[MZL12] Myers, Seth A; Zhu, Chenguang; Leskovec, Jure: Information diffusion and external influence in networks. In: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp. 33–41, 2012.

[Ng14] Nguyen, Tien T.; Hui, Pik-Mai; Harper, F. Maxwell; Terveen, Loren; Konstan, Joseph A.: Exploring the Filter Bubble: The Effect of Using Recommender Systems on Content Diversity. In: Proceedings of the 23rd International Conference on World Wide Web. WWW '14, ACM, New York, NY, USA, pp. 677–686, 2014.

[Pa11] Pariser, Eli: The filter bubble: What the Internet is hiding from you. Penguin UK, 2011.

[Zh12] Zhang, Yuan Cao; Séaghdha, Diarmuid Ó; Quercia, Daniele; Jambor, Tamas: Auralist: introducing serendipity into music recommendation. In: Proceedings of the fifth ACM international conference on Web search and data mining. ACM, pp. 13–22, 2012.