Title: Personal Web Search in the Age of Semantic Capitalism
Subtitle: diagnosing the mechanisms of personalisation

Authors: Martin Feuz, Matthew Fuller, Felix Stalder

Affiliation: M. Feuz / M. Fuller, Centre for Cultural Studies, Goldsmiths College, University of London, F. Stalder, Zurich University of the Arts

Contacts:
M. Feuz (cu702mf@gold.ac.uk), M. Fuller (m.fuller@gold.ac.uk), F. Stalder (felix.stalder@zhdk.ch)

Address:
Centre for Cultural Studies, Goldsmiths, University of London, New Cross, London SE14 6NW, UK
Specialization Media Arts, Department of Media & Arts, Zurich University of the Arts, Sihlquai 131, 8031 Zurich, Switzerland

About the authors:
Martin Feuz is a PhD researcher at the Centre for Cultural Studies, Goldsmiths, University of London. His research focuses on exploratory search interactions and how such interactions can be meaningfully supported.

Matthew Fuller is David Gee Reader in Digital Media at the Centre for Cultural Studies, Goldsmiths, University of London.

Felix Stalder is lecturer in digital culture and network theories at the Zurich University of Arts and senior researcher at the Institute for New Culture Technologies, t0 in Vienna. His work is accessible at felix.openflows.com.
Abstract

Web search engines have become indispensable tools for finding information online effectively. As the range of information, context and users of Internet searches has grown, the relationship between the search query, search interest and user has become more tenuous. Not all users are seeking the same information, even if they use the same query term. Thus, the quality of search results has, at least potentially, been decreasing. Search engines have begun to respond to this problem by trying to personalise search in order to deliver more relevant results to the users. A query is now evaluated in the context of a user’s search history and other data compiled into a personal profile and associated with statistical groups. This, at least, is the promise stated by the search engines themselves. This paper tries to assess the current reality of the personalisation of search results. We analyse the mechanisms of personalisation in the case of Google web search by empirically testing three commonly held assumptions about what personalisation does. To do this, we developed new digital methods which are explained here. The findings suggest that Google personal search does not fully provide the much-touted benefits for its search users. More likely, it seems to serve the interest of advertisers in providing more relevant audiences to them.

1. Introduction

Google’s mantra is ‘to give you exactly the information you want right when you want it’. They operationalize this by providing ‘personalised’ search results and recommendations. This is achieved on the one hand through logging of interactions whenever a person uses one of the many Google services and on the other hand by techniques such as collaborative filtering to generate group and user profiles based on which Google produces ‘personalised’ search results and recommendations (see,
Stalder & Mayer, 2009).

Such a situation raises a number of profound questions, both methodologically in terms of researching such a flexible entity, and politically, most notably in how to draw the line separating individual filtering as a convenient service from outright distortion, be it for overt reasons (censorship, propaganda, social engineering\(^2\)), or because certain aspects of reality and behaviour are more readily computed and thus feature more prominently within such systems.

We examine these questions by proposing three hypotheses and corresponding methods to test their validity. The hypotheses reflect the suggested character and value of personalisation based on assertions made on it by Google as well as popular writers in the realm of personalisation (Negroponte, 1996; Anderson, 2006) that may have influenced what we expect from personalisation. We test these hypotheses by developing a novel Digital Method. The paradigm of Digital Methods 'is not simply to import well-known methods - be they from humanities, social science or computing. Rather, the focus is on how methods may change, however slightly or wholesale, owing to the technical specificities of new media.'\(^3\)

This research thus provides a diagnosis of the mechanisms of personalisation in the age of semantic capitalism. To our understanding, this is the first such diagnosis within the domain of universal web search. Given the pervasive spread of the search paradigm to all aspects life, developing an understanding of forms of mediation by means of personalisation of web search results seems all the more urgent. The findings of our research are in many ways surprising, significant and disappointing with reference to what the grand promises of personalisation currently amount to. Disappointing also, precisely because it comes at the cost of giving up masses of personal information\(^4\).

Our findings contribute to an improved understanding of the quantitative effects personalisation has and proposes several lines of interpretation of the quantitative data. Based on this, we conclude that more research is urgently needed to develop a wider understanding of the social and cultural implications of personalisation of web search in people's everyday life.

### 2. The rise of the personalised search engine

Very soon after the World-Wide-Web was developed in early 1990s, the problem of locating relevant documents within the burgeoning information-space emerged. The WWW lacks an in-built indexing and categorizing mechanism. It is only threaded by one-way hyper-links (though this has changed a bit with the advent of social platforms). As the space expanded exponentially, simply following links became impractical almost immediately. Thus, indexing and categorizing were created as an extra service within this space, rather than being part of the protocols that create the space in the first place. Thus, the providers of such indexes have been very prominent actors from early on (Havalais, 2009).

Initially, in 1994, it was Yahoo! which offered a very familiar looking directory, compiled by experts, like a library catalogue. While this traditional format made the relatively unknown space of the Internet seem less alien to many, it quickly ran into deep problems, both in terms of scale (impossibility to keep up with the growth of the WWW) and ontology (the categorical system could not contain the complexity and dynamism of the information space it claimed to organize). In 1995, AltaVista, the first

\(^2\) Google, for example, tries to actively prevent people from committing suicide through it tweaking its ranking. [http://searchengineland.com/google-trying-to-help-suicidal-searchers-39389](http://searchengineland.com/google-trying-to-help-suicidal-searchers-39389)


\(^4\) ([Stalder & Mayer, Second Index / M. Zimmer First Monday](http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2136/1944))
full-text search engine built on automated information gathering and indexing appeared. It quickly overtook the human-compiled directory, because it was faster and more comprehensive. It also established the now-standard interface paradigm of a relatively empty page with a simple search box, in which one could enter a query and receive a ranked list of search results. The late 1990s was a time of great diversity in search providers. Dozens of search engines were competing for market share. A relative late entry into this burgeoning field was Google, started in 1996 and incorporated in 1999. Yet, it quickly came to dominate the field, particularly in the West, partly because it was based on superior indexing and search algorithms, partly because it remained focused on search, without much of a business plan. Most other search companies, trying to find ways to create revenue, transformed themselves into 'portals' – central points of access for heterogeneous information sources and services – providing content in cooperation with a range of media partners and neglecting search. This strategy failed almost across the board (Yahoo! being the most prominent example and partial exception). In 2000, Google introduced its own business model: targeted advertisement, which since became its single source of revenue. Thus, over time, it transferred itself into an advertising company, producing not search results, but audiences as its primary commodity. Economically speaking, search results are an expenditure like TV programming, given away for free in order to attract an audience, which then can be sold to advertisers by the provider of the TV channel. Google's success helped to established advertisement as the dominant business model in the field and restructured advertising according to the logic of semantic capitalism, “in which any word of any language has its price, fluctuating according to the laws of the market”. (Bruno, 2006). Thus, virtually all search engines today have a dual purpose: they provide search results to users and users to advertisers. Initially, the single variable to connect the two sides was the search query, which produced the same results for everyone. All users got the same search results to a given query, and all advertisers got the same users to a search term for which they had purchased advertisement space.

As the range of information, context and users of Internet searches grew, the relationship between the search query, search interest and user became more tenuous. Not all users were seeking the same information, even if they used the same query term, and even individual users did not connect the same search interest with a particular query at all times and locations. Thus, the quality of search results, at least potentially, decreased. Search engines began to respond to this problem by trying to personalise search, promising to deliver more relevant results to the user, whose query is now being considered in the context of his/her search history and other data compiled into a personal profile. The very same approach also addresses the problem of the advertisers, who are interested only in the most relevant users, that is, those most likely to be influenced by their message.

The introduction of personalised search marks an important moment of intensification in semantic capitalism (White, 2010). Not only is it the case that every word in every language now has its price that fluctuates according to the laws of the market, but additionally, both search results and the corresponding advertisements shown are now optimised according to their potential market value based on pre-emptively calculated individual ‘user relevancy’.

On Google’s AdWords Help Center site, (an information guide for advertisers) Google states the following among the top reasons why their platform ‘is the best to reach your potential customers’:

---

5 Google's market share in 2008/9 (for internet searches) ranged from 96% in Belgium to 3% in Korea. See [http://googlesystem.blogspot.com/2009/03/googles-market-share-in-your-country.html](http://googlesystem.blogspot.com/2009/03/googles-market-share-in-your-country.html) [accessed 13.09.2010]. Generally speaking, in Western language markets, Google's share is very high, in Russia or Asian language markets, it is much lower. This probably reflects linguistic biases in the search algorithms. See, Vaidhyanathan (2009)


7 The concept of the audience-commodity was initially developed for mass media by Dallas Smythe (1981), but it can be applied to any advertisement driven business.
Tools: Enhanced keyword tool provides lists of additional phrases to consider and most popular synonyms based on billions of searches. Results in better targeting and higher click-throughs.

Ranking: [Ad] Rank is determined by a combination of several relevance factors including CPC [Cost-per-Click] and click-through rate. If an ad is irrelevant to users, they won't click on it and it will move down the page. Your relevant ads will gain higher positions on the page, at no extra cost to you.

In order to produce this context, vast amounts of personal information need to be collected, organised and made actionable. Within the fast receding limitations of storage space and computing power, profiles can never be too comprehensive, too detailed, or too up-to-date. Google, the most advanced player in this field, is compiling personal profiles in three dimensions: the knowledge person (what an individual is interested in, based on search and click-stream histories), the social person (whom an individual is connected to, via email, social networks and other communication tools) and the embodied person (where an individual is located in physical space, and the states of the body) (see, Stalder & Mayer, 2009). Together, these three profiles promise to provide a detailed, comprehensive and up-to-date context for each search query, with the potential to deliver precise results, that reflect not just the information 'out-there', but also the unique interest a user has at any given moment. Personalised search does not simply aim to provide a view onto existing reality, which is problematic enough (see, Introna & Nissenbaum, 2000). Rather, personalised search promised an 'augmented reality' in which machine intelligence interprets the user's individual relationship to reality and then selects what's good for each. Today, we are at the border of this 'augmented reality', which promises to be more personal by becoming more opaque and more abstract. Attempting to make such conditions more transparent and more concrete raises a number of difficult methodological questions.

3. Methodological Considerations

Search engines and related systems are of increased importance because of the role they play as parts of contemporary abstract infrastructure and this has implications for the way in which they must be researched. They are infrastructural to the present world because, like water mains and roads, numerous other systems now take them for granted as a means of operation. But what do we mean by saying they are abstract? Abstraction in this sense refers to the diagrams effectuated between entities embodying them, and which may precede them. At a certain scale, the abstract machine may simply be interpretable as the algorithms, data structures and other elements of a search engine, whether or not it is populated with data. But it also refers to another scale, that of the modes of knowledge, classification and ordering that inform and shape the engine, and in turn, by a further scale - that things in the world may be seen as computable data and information (Chun, 2011; Golumbia, 2009). The further interpolation of such systems within mechanisms of governance such as private companies and state governments and the reserve of such agencies in rendering their ordering accountable to users or others, defended in turn by the norms of intellectual property and the necessities of market competition, renders such infrastructure particularly inscrutable (Becker & Stalder, 2009).

The second scale mentioned above is of particular methodological interest here. As attention in social sciences and material culture increasingly develop means of recognising the multiple kinds of agency of things, their potential methodological agency, their epistemic and ontological valence, also becomes apparent as a possibility to be worked. That is to say that not only do things have kinds of complexity that need attending to, but that also, as Celia Lury and Nina Wakeford suggest, they have some substance as means of knowing (Lury & Wakeford, forthcoming).
Between such contexts, that of the high stakes of the forms of abstract governance enacted in such engines – where assemblages of code and hardware are crucial - and the possibilities of working with material forms, such as those of software, in devising means of interrogating them, there is an interesting confluence of methodological means and problematic. Using software-based methods to understand the reality-forming nature of search engines and related systems is thus a point at which several initiatives converge. A short overview of these follows.

A number of notable tendencies have emerged in studying the Internet using software-based methods. Whilst summary cannot do them justice, they provide an important contextual dimension to the work done here and it is useful to orient the present work in this regard. Firstly, work in Network Analysis tends to provide mathematically and topologically oriented means of enquiry into the composition of networks (Newman et al, 2006). Whilst such work has many applications, including a mapping of the increased centralisation of the Internet, it tends to remain largely detached from wider cultural analyses.

A certain set of rather more guerrilla-oriented practices, in Search Engine Optimisation, (Search Engine Watch, n.d.) can also be said to engage in methodologically rich, highly empirical though under-documented, interactions with search engines. The recognition of search engines as non-neutral systems that are inherently being gamed at multiple levels is core to SEO operations and their fine degree of attention to the nuances of such systems provides useful understanding of their stakes.

Providing another form of context for the present work is the field of software art, a current of work developed for more than a decade at this point, in which software is reflexively articulated in aesthetic exploration of emerging digital cultures. Search engines and their operations have been key to certain key works and initiatives in this field (Fuller 2003, Ludovico 2009) and the often irreverent and inventive methods deployed in software art are inspirational to our approach here, which, whilst not constituting such a project in itself harmonises with a certain sense of the tendentially absurd in software that is abundant in software art (Bruno, 2006).

Two further methodological contexts are worth noting here. Digital Methods and what might be called ‘Investigative Computing’. Digital Methods is pioneered by Richard Rogers at the University of Amsterdam, a researcher with a long engagement with search engines (Rogers 2000) and with efforts to pursue ‘digitally native’ research methods (Rogers 2004) that are appropriate to research on the Internet, in a way which counts its technical specificity as being inherently significant. The Digital Methods Initiative provides a ‘collaboratory’ in which, “Scripts to scrape web, blog, news, image and social bookmarking search engines, as well as simple analytical machines that output data sets as well as graphical visualizations” (Digital Methods Initiative) are placed in carefully chosen conjunction with the factors of particular questions. Digital Methods use software to study worlds partly or wholly composed of such systems.

Investigative Computing aims to uncover particular aspects, hidden conjunctions of forces in contemporary life as they are manifest in networks and other systems. Part of this current can be seen to have affinities with the tradition of hacking to discover malfeasant actions of the operations, whether publicly or privately owned that govern computing and telecommunications. Examples of such work are widespread and are typified in the reporting of 2600 magazine, (2600) and the activities of collectives such as Chaos Computer Club and others. Another tendency in this field is more academically inclined and typified by the work of Ben Edelman at Harvard. Edelman is both a programmer and a lawyer and has developed a highly significant programme of work probing the activities and claims of online media companies and service providers of many kinds, often in the area of advertising. His work typically attempts to gain means to answer a specific question, such as
whether or not a particular search engine can be shown to inflate advertisers’ costs, (Edelman, 2010) to serve users with deceptive advertising (Edelman, 2009), or track users’ online activities despite statements to the contrary. (Edelman, 2010). Such methods work with new kinds of evidence and documentation to ‘reverse engineer’ the policies and actions embedded in software allowing precise questions to yield telling answers.

The method proposed here aims to complement such work with an analysis of the linguistic manifestations of such powers, something which immediately spills into further complexities. Thus, the central aim of the method is limited to rendering visible some of the effects of such operations. Given the highly dynamic environment of the Internet and web search engines more specifically, warrants highlighting some of those dynamics as they surfaced in this research project.

In designing the method and setting-up the research project we had to take into account the particular conditions which this object, which reacts to user interaction, can in fact be studied. For example, Google employs a control system geared towards identifying automated search queries. From what is unofficially known (http://www.squigglesr.com/spip/spip.php?article33#timing), more ‘natural’ patterns of search queries occur in a ‘burst-like’ pattern. As Google generally operates on the premise of ‘security through obscurity’, it is unknown what constitutes such a burst-like pattern in their view. The stated rationale is to block Search Engine Optimisers, that aim to improve search rankings for interested (and paying) parties. But of course, such behaviour also impedes all other reasons for doing automated search queries, such as our research.

While performing the research, that is the training and testing-sessions that will be described shortly, the study object reacted to being studied in a way that may be considered mild but nevertheless annoying. Mainly, this reaction took the form of IP address blockages during testing-session when examining how personalisation has evolved in terms of search-results position distance (see Section Research Findings: Hypothesis 2). Ironically, Google’s control system considered our search result click behaviour, that is clicking beyond search result page 1 and 2, as looking ‘similar to automated requests from a computer virus or spyware application. To protect our users, we can’t process your request right now.’ (See footnote 14 for more details.)

After performing our research, the conditions on how the study object can be actively researched have changed dramatically. Since 4. Dec 2009, Google has introduced personalised search results for all search queries, including those of users that are not logged into any of their services. This means that this research could no longer be performed as there is no longer an ‘objective’ baseline of ‘non-personal’ search results to work from.

4. Description and discussion of research method

The aim of the research method was to generate data in order to analyse and render more transparent and interrogable some of the specific ways ‘personalisation’ is already shaping current search results provided to users. To our knowledge no such study has been undertaken previously. We focused solely on Google for two reasons. Firstly, Google’s market share means that most people will actually be affected by the personalisation it provides. Secondly, the extent of services Google offers, such as

---

8 Work in network analysis is often allied with the broader field of Web Science, a context in which the establishment of new standards and inventions in the internet itself, for instance the Semantic Web, also involves the recursive move of analyzing the network in various ways. It is perhaps the facilitatory nature of web science with its emphasis on the disinterested development of technologies of enlightenment that distinguishes it from the present approach.

Three philosophers were selected, one from the 18th, one from the 19th, and one from the 20th century, namely, Immanuel Kant, Friedrich Nietzsche and Michel Foucault. The search terms for the generation of the Web History profiles for each philosopher were based on the indexes of seven of each philosopher’s books using the standard English translations. This amounted to approximately 6,000 search queries per philosopher during the training-sessions.

Each philosopher had his own Google account. A training-session consists of performing the search queries compiled from one of the indexes of a book from each philosopher whilst logged-in to a Google account with the Web History feature activated. For the training-sessions only, each search query was followed by a random visit to (the equivalent of a click-on) one of the search results provided by Google. This is accounted for in Google’s Web History and thus further intensifies the ‘personalisation’ process. There were seven training-sessions per philosopher, each of which is represented by one book. The goal of the training-sessions was to record how subtle ‘personalised’ search results would develop over the course of the seven training-sessions.

After performing a training-session for each of the three philosophers, a testing-session followed, thus seven in total. Each search query, whether for a training or testing-session, was concurrently performed for a so-called anonymous user. An anonymous user is technically constituted by the absence of any login credentials or other previously tracked user data reported to Google. This method allowed the comparison of search results received by the profiles to a generic Google baseline and determine, whether the search results for the profiles were ‘personalised’ and if so, how.

The goal of the testing-sessions was to record and compare if and how ‘personalised’ search results would develop differently for the three profiles. The testing-sessions were based on 40 search terms, which remained constant over the course of the seven testing-sessions. The search terms for testing were drawn from three groups:

- The first group of test search terms was based on a pool of such terms from the training search terms set, terms which the three philosophers have in common. The terms were: aesthetics, causality, dialectic, ethics, freedom, immortality, knowledge, morality, obedience, punishment, reflection, Sophists, virtue, welfare.
- The second group was based on popular tag words from the social bookmarking service delicious.com, which can be said to vaguely represent contemporary Internet culture. The terms were: software, diagrams, travel, neuroscience, open source, programming, art, blogs, learning, information, knowledge, technology, video.
- The third group was based on Amazon’s ‘Statistically Improbable Phrases’ from three contemporary books concerning surveillance, network theory and global democracy. The three authors have some overlap of terms in their index
  - Foucault and Nietzsche: 10% keywords overlap
  - Foucault and Kant: 20% keywords overlap
  - Nietzsche and Kant: 22% keywords overlap

Alternatively, we could have visited always only on one of the ‘personalised’ search results provided by Google as a purely positive feedback signal and consequent valuation of those ‘personalised’ search results provided by Google. Designing our method along those lines would have been in stark conflict with the actual navigation options available to users made of flesh. Also, this would actually be more likely a kind of reverse engineering, which this project has no interest in.

Footnote SIP: Amazon.com's Statistically Improbable Phrases, or "SIPs", are the most distinctive phrases in the text of books in the Search Inside!™ program. To identify SIPs, our computers scan the text of all books in the Search Inside! program. If they find a phrase that occurs a great number of times in a particular book relative to all Search Inside! books, that phrase is a SIP in that book. SIPs are not necessarily improbable within a particular book, but they are improbable relative to all books in (Amazon’s) Search Inside! (Program).

The books are: a) ISpy: Surveillance and Power in the Interactive Era, by Mark Andrejevic (2007); b) Network Power:
terms were: immaterial labor, global multitude, immaterial property, asymmetrical conflicts, global second language, networked information economy, linguistic coordination, dominant network, indirect force, virtual enclosure, interactive era, citizen publicity, monitoring gaze.

Training and testing sessions were performed in July 2009 over the time span of three weeks. In terms of its technical traits, the method was designed and performed along the following lines: For each philosopher profile, a Google Gmail account was opened. The country settings required when opening a Gmail account were set to UK. As part of this process, the Google Web History feature is activated by default. The Gmail accounts were never used for any other purposes. In order to secure rigour of the method, the search queries were performed from a server with a fixed IP address from central London, UK. All search queries were performed explicitly on the google.co.uk domain. The search queries were defined as searching “the Web” (rather than choosing the ‘from United Kingdom only’ option available on the browser of the Google search interface) and ‘Safe search’ mode was turned off. For the training-session only, each search query term was entered in quotation marks, so as to establish a ‘specific’ search history. Only the search results which Google provides on the first page of the browser were considered. Advertisements were specifically excluded in the method. Testing-session search queries were deleted in the corresponding profiles’ Web History after each testing-session.

Software was developed to compare search results between profiles and the anonymous user with regards to ranking position for search results that have the same URL, and reports absolute differences of search results between anonymous user and profile based on URL. The following information is compiled per search query:

- Number of search results that are identical in terms of URL and position rank for anonymous user and profile
- Number of search results that are identical in terms of URL but different in position rank for anonymous user and profile
- Number of search results which anonymous user and profile do not have in common based on URL’s of search results returned
- Parameter to indicate whether search results for profile were personalised and what type of personalisation had occurred (none, re-ranking, different search results, or combination of the two)
- Parameter to indicate in % the degree of intensity of personalisation of search results for a philosopher profile.

Based on the method, a total of 18,211 search queries were performed, containing 195,812 individual search results. Put differently, the approximately 6'000 search queries per philosopher translate into a bit less than 4 search queries per day over four years, a threshold many will have passed since the original inception of personal search by Google in April 2005 and Google’s corresponding data collection.

In designing the method, we aimed at creating profiles likely to be different enough to produce diverging results, yet similar enough to remain comparable. Thus the selected profiles would have to belong to a larger but semantically clearly distinguishable subject area, while at the same time being different enough from each other within that field. It is for this reason that, we selected three authors from the field of philosophy, each with his own typical semantic idiosyncrasies, based on subject areas, historical period and personal interests. This approach also had the advantage of being able to interrogate the level of group cluster profiles that are applied by Google's ‘personalisation’ algorithm. Group cluster profiles are different from individual profiles (i.e. one of our philosophers) in the sense

that they also contain characteristics which are assumed to be relevant for the entire group, but for which personal data for a specific member of this group is lacking (e.g. putting Immanuel Kant to the upper-middle class consumer category and personalizing results accordingly in subject areas outside philosophy.)

The web search profiles established for this research can obviously only partially resemble those of a natural person, as the interests of natural persons go well beyond professional interest in a subject field. Nevertheless, it can be said that selecting the works of philosophers as compared to say a writer on Information-Retrieval do some justice in resembling the interests of a natural person, as they also cover topics as broad as sex, war and personal health etc..

5. Research findings. The ambiguities of Personalisation

Hypothesis 1: ‘Personalisation is subtle—at first you may not notice any difference’. 14

What this hypothesis means:
Since these are the words of Google's PR machinery, we can only speculate what they may mean. Nevertheless, two aspects stand out: 'subtle' and 'at first'. We take 'subtle' to be a quantitative measure and in this case referring to the extent to which a search query is being served with altered search results. That is, how many of the ten search results provided for a search query are exchanged for personalised ones. 'At first' seems to refer to the frequency with which a person’s search queries are served with personalised results, in other words the rate at which it deviates from those of un-personalised search.
Thus, we understand the hypothesis as personalised search results to appear in low magnitude as well as infrequently, especially at the beginning ('at first') of a persons search history.

Method applied to test the hypothesis:
In order to test this hypothesis, we analysed the results from the seven training-sessions. These represent the search and web history for the three philosopher profiles. As discussed in the Method Section 4, each training-session search query for a philosopher was immediately followed by a search query for an anonymous user. The search results of the philosophers were then compared to those of the anonymous user based on URL and search result position (1 to 10). This procedure allows to identify both the frequency with which the philosophers search queries were served with personalised search results as well as the extent to which the ten search results for the philosopher profiles search queries were exchanged for personalised ones.

Research Finding: Reject the hypothesis
As our findings indicate, Google personal search begins to have effects rather quickly. Concerning frequency, for Foucault's first 453 search queries (Training-session 1), slightly more than every 10th search query, 11% to be precise, was served with personalised search results. However, at 1,585 search queries (Training-session 2), every third search query was served with personalised search results. In fact, for both Kant and Foucault’s profiles, the first ‘personalised’ search results appeared within the first ten training-search queries performed.
At 2'809 performed training-search queries (Training-session 4) for Nietzsche’s profile, on average
more than every second search query was ‘personalised’ (56%).\(^{15}\)

With regards to intensity or 'subtleness', an even more surprising finding arises. The 52 (11\%) personalised search queries within Foucault's first training-session show an average intensity of 64\%.

That is, on average 6.4 search results out of the 10 provided by Google for a search query were altered.\(^{16}\)

Kant reaches an even higher level of intensity by his second training-session, that is after 1,335 search queries. Theoretically, subtleness could also be assessed from a qualitative standpoint. That is as a direct comparison of the search results being exchanged for another or re-ranked. Necessarily, such an undertaking rests upon the individual person being exposed to personalised search results and generally cannot be usefully judged. Unfortunately, it is precisely this option, which has been foreclosed by the current way in which Google delivers results, with compulsorily ‘personalised’ results.

We thus conclude that Google personal search is, from a quantitative perspective, not very 'subtle', even if it might feel qualitatively subtle to users. However, this qualitative subtlety is at least partially due to the fact that the user has no way of detecting the degree of personalisation. This raises, a number of questions regarding accountability of the personalisation process as such. If even relatively abrupt changes due to personalisation remain essentially undetectable to end-users, how can they trust the results? The subjective quality of the search results – whether the users feels them to be relevant – will not suffice, since this subjective test can only be applied to domains that the users knows very well and can thus recognize if something is missing. In the vast majority of cases, users will not be able to judge the quality of the personalisation with any degree of critical autonomy.

**Hypothesis 2: The more user search history is gathered, the more long-tail content is retrieved.**

What this hypothesis means:

Personalisation schemes are thought to generate increasingly 'personalised' or relevant recommendations the more they know about a person's interest. In our case this is reflected by the extent of the search and web history collected by Google. The assumption of this hypothesis then is, that as Google collects information on more and more of a person’s interests, the likelier it should become that personalised search results will be found beyond the first hundred search results for an anonymous user and thus represent truly long-tail content. Greater variety in search results, after all, is the core promise of personalisation.

Method applied to test the hypothesis:

In order to test this hypothesis, the ranking positions of personalised search results received for the philosopher profiles were identified from the perspective of the anonymous user. The analysis is based on URL equivalence. Only personalised search results within the first ten search results provided by Google for the philosopher profiles were considered. We performed this test based on testing-session 7.

At this point, more than 5,000 training search queries were performed for Kant and more than 6,000 for Nietzsche and Foucault respectively.

Concurrently to each search query for the philosopher profiles carried out in the testing-session, the

\(^{15}\) Training-session 1 for Nietzsche's profile shows a very high frequency and intensity of personalised search results. This is due to a time lag of 22 hours when comparing the search results to those for the anonymous user. This lead almost certainly to the high numbers of personalised search results. Nevertheless, we included this data, as it makes better understandable the underlying dynamics of the research object.

\(^{16}\) Intensity of personalisation refers to the percentage degree by which the search results for a search query have been personalised. This percentage is calculated as the count of search results that were re-ranked in position and/or altogether different (based on URL) when compared to a simultaneously performed anonymous search query.
same search query was performed for the anonymous user. Based on the URL of the personalised search results for the philosopher profiles, we searched the search result pages of the anonymous user and noted its search result position.\textsuperscript{17}

Figure 2: Infograph Hypothesis 2

Research Finding: Reject the hypothesis

Testing-session 7 produced 35 ‘personalised’ search results in total. 5 of these were based on content from Google that was only recommended to the philosophers but never within the search results of the anonymous user (bear in mind that Google will only ever make approximately the first 1,000 search results for a query available). 37\% of the these personalised search results were found on the second page of search results (positions 11 to 20) and can, thus, be said to represent an exchange of very dominant positions. A further 43\% were found between search result pages 3 and 10 and can also be said to occupy relatively dominant positions.\textsuperscript{18} Only 7\% were found to be between pages 11 and 100 (on rank 108 and 123 respectively) and another 13\% of search results were not available in the approximately 1,000 search results provided by Google for the anonymous user.

Our research finds that Google personal search does not seem to be able to make long-tail content available in a substantial manner. This can be interpreted in numerous ways. First, the weight of personalisation, vis-a-vis other ranking methods, is currently relatively limited, predominantly confined to re-ranking already highly ranked results. This could suggest that Google puts relatively little emphasis on personalisation, applying it to relatively trivial dimensions. This might be a sign of a complexity a the personalisation process and the relatively novelty of its introduction. Following this

\textsuperscript{17} Performing this analysis wasn't without difficulties. Searching for the specific URL’s within the anonymous user's search results required to click through multiple search result pages for a search query. This behaviour isn't quite like the that of the average Google search user. Numerous times our IP address was blocked by Google, displaying a comment that the 'query looks similar to automated requests from a computer virus or spyware application'. To continue the analysis, we had multiple IP addresses available and could swiftly change them. To ensure methodical robustness, all IP addresses were from a central London based proxy server.

\textsuperscript{18} This holds all the more so when their ranking is considered in percentage ratio reference to all the search results available for the search query as indicated by Google (top hand right corner of the search interface). All of them were found to be in the top tenth per mill (0.01\%) of all results.
interpretation, we can expect this to change over time. Another interpretation would suggest that the Google ranking methods do not work very well on the long-tail and thus they are unsuited to find relevant content there. This could be a legacy of Google's focus of finding the one best document (the “I feel lucky” option on the front page is testament to this focus) rather than structuring broad knowledge domains. We can experience this when we look beyond the first 50 results. The relevancy drops quickly and noise increases rapidly. This interpretation would suggest that even with improvement of the personalisation algorithms the situation is not likely to change, thus, in general, personalisation is of limited value to users. The third interpretation would be that personalisation is not primarily about improving search. In this interpretation, what is primarily personalised is not the search result but the delivery of advertisements and the precision with which users can be matched to advertisers seeking to influence their behaviour. Thus, personalisation is about fine-tuning the relationship between users and advertisers, further consolidating the market dominance of Google for online advertisement. In his perspective, personalisation is an intensification of semantic capitalism, using control over words and symbols as a means of expanding capitalistic logic. Of course, these three interpretations need not to be mutually exclusive, but can apply at the same time, to varying degrees.

Hypothesis 3: Personalisation reflects only an individual user's past search and web interests.

What this hypothesis means:
Personalised search results are only served for those search queries that reflect a user’s semantic history as recorded by Google. In reverse this means, that Google does not serve personalised search results for queries for which it has no direct means of assessing potential relevancy based on the individual user's factual search and web history.

Method applied to test the hypothesis:
The seven testing-sessions constitute the basis for evaluating this hypothesis. We only considered the test-terms based on popular tag words from the social bookmarking service delicious.com (group 2) for our test, as they exhibit the necessary semantic distance from the philosopher profiles. These terms (Software, diagrams, travel, neuroscience, open source, programming, blogs, information, video) are not reflected within the philosopher's training-session search queries.

For every testing-session we counted the number of search queries for which the philosopher profiles received personalised search results (count #1). Furthermore, we counted the number of test-search queries performed per philosopher and testing-session (count#2). We then divided count#1 by count#2 to calculate the percentage ratio per testing-session and philosopher for which the philosopher profiles receive personalised search results.

To confirm this hypothesis, the three philosopher profiles should not see personalised search results at all.
Research Finding: Reject the hypothesis
As the above data clearly illustrates, in every testing-session all philosophers received personalised search results for some search queries, even if there was no relationship between the search history and the test query. In fact, the data not only demonstrates that Google applies personalisation to search queries outside the user’s domain of recorded search and web history, but this tendency increases over time, as the upwards trends in percentage rates indicate. Once more, we have to refute the hypothesis.

Similar to our discussion of the previous hypotheses, we can only speculate about the underlying, undisclosed processes that generate these results. To us, the most likely interpretation is that Google does not only rely on a user's personal semantic history, but that it extrapolates from what it knows about a person to his or her association with statistical group profiles that Google has built up over time. A strong interest in philosophical terms – which can be gleaned from the semantic history – could, for example, be associated with certain age and income groups, which, in turn, become associated with certain preferences in, say, holiday destinations. In such a way, Google infers Immanuel Kant's taste in hotels, or Friedrich Nietzsche's bias for or against open source software. The result of such group patterning in the background, unseen and undetectable to the user affected by it, could be an inversion of the promise of personalisation. Rather than seeing what is of most interest to the user as an individual, we are presented with a preselected image of the world based on what kind of group the search engine associates us with. Rather than increasing diversity, this might well lead to a subtle homogenisation, and the adherence to a preselected world becomes a self-fulfilling prophecy. If Kant chooses one of the hotels Google has preselected for his income bracket, he will deliver data to Google showing that this pre-selection was correct and thus anchoring him more deeply in this group to which he perhaps wouldn't otherwise have belonged.

6. Conclusion and further Questions
This research has indicated that personalisation is a far from unambiguous process simply delivering better results to the user. At the moment personalisation is both taking place to a surprising extent (hypothesis one), but with relatively trivial results (hypothesis two) most likely reflecting the fact that we are in the early stages of the process and that, at least, some of the benefits of personalisation will
not accrue on the side of the end-user, but on the side of the advertisers and thus Google itself, which sells these personalised audiences. Furthermore, we have produced first evidence (hypothesis three) that Google is actively matching people to groups, which are produced statistically, thus giving people not only the results they want (based in what Google knows about them for a fact), but also generates results that Google thinks might be good to users (or advertisers) thus more or less subtly pushing users to see the world according to criteria pre-defined by Google.

Each of these preliminary results presented in this article, raises a host of further questions about the ever-increasingly opacity of the process by which search results are generated and the growing influence of Google in shaping this process to further its commercial agenda. Thus, questions regarding the transparency of the personalisation process and the boundary between service to the users and their manipulation in the interest of advertisers are becoming increasingly crucial.

These questions are closely intertwined with methodological questions. How can we study a distributed machinery that is both wilfully opaque and highly dynamic? One, which reacts to being studied and takes active steps to prevent such studies being conducted on the automated, large-scale level required? Furthermore, personalisation makes inter-subjective testing of hypotheses impossible. The dynamism of search engines makes this even more complex over time. Thus there is an urgent need to develop digital methods further, an endeavour that must be multi-disciplinary, given the fact that automated processes need, at least on some levels, be studied by means of other automated processes.

Unless we can update our research methods and tools, we cannot adequately address the social and political issues connected with personalisation and the power of search engines more widely. But we urgently need to do this, otherwise the knowledge and power differentials between those on the inside of search engines and those who are mere users of a powerful but opaque machine are bound to grow.

Acknowledgements
Martin Feuz would like to thank Florian Bös ch, Daniel Boos, Megan Hall, Raphael Perret, Daniel Unger and the great folks at DMI Amsterdam, especially Esther Weltevrede, Erik Borra and Richard Rogers. I am most thankful to Monya Pletsch.
licen ce: cc by-nc-nd
Bibliography


Celia Lury and Nina Wakeford, eds. in press, Devices, the happening of the social. London: Routledge.


