The Spirits That We Called

by Sabine Hossenfelder

MARCH 13th 2008: Until Wednesday, the Presidential candidate [insert name here] scored high in the polls. Then a Google search for his name showed up as first hit a report on an alleged child abuse committed by the candidate, published by Mary S. (name changed) on her personal website. The story was backed up by the following highly ranked hits that indicated two similar events during his youth, though reliable sources were missing. Within less than one hour, the reports were echoed on thousands of weblogs, appeared on digg and reddit, the original websites received 200,000 hits within the first 6 hours, until the server crashed down. Immediate press releases by the candidate's PR groups did not appear on the Google listing, and could only be accessed by secondary links. It took until the next day that printed newspapers could attempt to clarify the situation.

NOVEMBER 9th 2011: Two independent eye witnesses report on their weblogs about Chinese military violently overtaking the government in Khartoum, Sudan. The reports score first hit on the key words 'world news', 'news' and 'foreign politics' at Google, later also on 'Sudan', and 'China'. Reports by the Chinese government denying the events did not appear on the Google ranking. The events were picked up by various TV stations, using the provided YouTube videos of extremely bad quality and doubtful sources. Dozens of reporters asked the White House for a statement. The President said he would not tolerate China getting a grip on Sudan's oil resources. The Shanghai Composite Index fell 541.12 points.

DECEMBER 30th 2015: Six months after Google and Yahoo was bought by Frederic F., multi-billionaire and president of several global companies, it was officially announced that Google will further improve the quality of search results, and counteract the drawbacks of information overload. Beginning New Year's day 2016, the algorithm will filter out "low quality sites, sites of obscure origin, and doubtful content", as the press release states. Yet it remains unclear who sets these criteria. Frederic F. answered inquires with "Customers trust us. We will not disappoint

them, and remain truthful to our philosophy to do no evil." He explained the need for such a change with the accumulation of outdated and irrelevant information on the web and added "Google will do its best to provide the user with correct information. Our employees are working hard to provide an excellent service to foster global knowledge."

I. Information Overload

Information is one of the most important resources in today's world. In a rapidly changing environment that gets complexer every day, the availability and accuracy of information is essential already to preserve the status quo, and indispensable to further progress.

Information overload isn't just an error message my brain produces when I check the arXiv, and an expression that I've made up for fun, but a rather unsurprising and well known side effect of a tightly connected world. The human brain's capacity to process input is limited. Today you are confronted with more information than you a) need and b) can deal with. The challenge today is not to collect all information you can possibly get, but to filter it and extract the relevant bits.

You can notice how your brain has learned to deal with information overload: by only skimming this side, losing attention already at this paragraph - because it's not obvious to you what it might be good for listening to me [1].

> "[I]nformation overload is not only caused by the sheer volume of information, but also because of the complexity or confusing structure of information that might overtax the users cognitive skill to focus on relevant information ... Therefore Helmersen et al. (p. 2) characterize information overload as "difficulties in locating, retrieving, processing, storing and/or reretrieving information due to the volume of available information." Information overload may lead to stress, health problems, frustration, disillusionment, depression, as well as impaired judgment and bad decision making ... From an ethical perspective, these consequences of information overload are problematic, because they undermine several basic principles, especially the requirement of participants autonomy/self-determination and the nonmaleficence principle."

---

[1]And now you're confused because I don't tell you ;-)

Behr, Nosper, Klimmt & Hartmann (2005) *Some Practical Considerations of Ethical Issues in Virtual Reality Research*, Presence Teleoperators & Virtual Environments 14:6, 668 (2005)

The internet collects and hosts an increasing amount of data. Besides potentially resulting in "frustration, disillusionment, and depression" as claimed in the above quote, a database without tools to find the relevant information however is above everything else useless: Badly ordered information is no information. Envision a library without any cataloguing. Of what use is it if you're told everything you need to know can be found somewhere on these four floors, filled with bookshelves up to the ceiling?

Luckily, thanks to ingenious software masters, we have today powerful search engines that help us structure the available information.

Und nun komm, du alter Besen!	*Come on now, old broom, get dressed,*
Nimm die schlechten Lumpenhllen	*these old rags will do just fine!*
Bist schon lange Knecht gewesen:	*You're a slave in any case,*
nun erfülle meinen Willen!	*and today you will be mine!*

II. Filtering Information

It is 2008. Today's school kids have grown up with the internet. It promises answers to all questions you can possibly have. And if you can't find an answer, ask the expert. Even better, you will find support no matter which opinion you happen to hold, or which side of an argument you want to defend. And you will somewhere come across a forum of likeminded friends that confirm your convictions.

What criteria is it that people use to filter information? A high Google ranking is without doubt useful to pass a first filter. Note, I neither said a high Google ranking is an indicator for quality, nor do I assume people are not aware of that. It is just a fact, that what ranks highly on search engines is more likely to be read [2]. And what is more likely to be read is more likely to stick. Esp. children who haven't been taught how to deal with information they find on the internet are prone to make mistakes in judgement, but confirmation bias is a fairly wide spread habit among all ages.

Besides this, people give a higher value of importance to information just because they hear or read it repeatedly (2). What's in your face is in your mind. There

---

[2]If you're one of the 100 visitors per day that comes to my blog searching for 'Map of America' you know what I mean.

must be something going on when many people point into the sky. That's what advertisements take advantage of, that's what meta filters like digg and reddit do, and that's what search engines do: directing attention, filtering your information. Now you can tell me everyone of us ¡em¿should¡/em¿ be rational, we ¡em¿should¡/em¿ always check sources, doubt unverified reports even if repeated several times. We shouldn't believe what we read without questioning it. We ¡em¿should¡/em¿ seek accuracy and not easy entertainment. We should, we should, we should [3]. But face it, many people don't. Because they just don't have the time, or are not interested enough, and the most commonly used criteria in this case is to follow the masses. Read what others read (the posts with the most comments?) go where many people link to, talk what others talk about, pay attention to what many people consider relevant. Majority offers security, Wikipedia is trustworthy, Google has proved useful. No go back to my opening line: "Information is one of the most important resources in today's world." Accuracy and availability of information is essential for the progress of our societies. You can direct people's opinions with the information you given them, and in the way you provide it. You don't need hard censorship for that, its more efficient to leave people the illusion of knowledge. It doesn't matter if there's a right for free speech, if you can make sure little people listen to what they shouldn't hear. Majority offers security, Wikipedia is trustworthy, Google has proved useful?

The preface of this post are three examples of how easily tempering with search engine algorithms can today affect opinions. Effectively, this interferes with our political systems since information is the basis for our decisions. Note, our decisions are *not* based in the information that is 'theoretically' available - somewhere, somehow - but on the information that is 'practically' available in our head, because we've read it, because we recall it, because we consider it consciously or unconsciously relevant.

Ein verruchter Besen!	*Be you damned, old broom,*
der nicht hören will!	*why won't you obey?*
Stock, der du gewesen,	*Be a stick once more,*
steh doch wieder still!	*please, I beg you, stay!*

III. Politics on the Web

The internet today has aspects of different political systems: capitalist anarchy and direct democracy, that reflect in the most frequently used services on the web.

---

[3]New Year's resolutions, anybody? Yawn.

III a. Googlism

Information has always been filtered by the media, and there has always been an influence by this on our political opinion making process. People have always fought for attention. New is

1. The necessity. The increasing need for such a filter, and the the relevance this ordering mechanism obtains through this. Consider Google, Yahoo, and MSN were down for 24 hours and the consequences.

2. The centralization. Google isn't the only search engine, but without doubt the presently most popular one. Millions of people world wide rely on it. How many would even notice if all hits after page 3 were missing?

Combine that with the problems into which old-fashioned print media runs, because they have trouble selling yesterday's news. Those who structure the information of many people have political influence. This is not a virtual, but a very real reality. The internet affects our daily lives, and it is still mostly a legal vacuum [4]. If I was a terrorist, I'd overtake the Google headquarter, and prominently place a couple of fake reports causing the US economy tumbling down, setting the stage for another war. You think that wouldn't work? Think twice. We life in a world where a couple of cartoons can kill dozens of people.
The Google ranking of a website can be pushed by various means. If a company, or lobby, can afford to hire an expert in search engine optimization, they can literally buy a good ranking. Even better if they invest in paid links, or further advertisement. The internet is presently mostly a capitalist anarchy with communist (shareware/no private properties) areas, that are struggling for structure. It offers the possibility to focus lots of influence in the hand of little people. The archetypical nerd community the web started with is today a small minority among those who just use the net they are being offered.
One can hope that there are self-regulatory mechanisms that save our societies from being influenced by a small groups of people because users would just chose different companies, different information sources. Or maybe some nerdy guys would set up their own 'better' search engine. That might work. But there is no guarantee it will. It is far from clear what a majority of people would consider 'better' for

---

[4]And each time I have to read through insults in blog's comment section, its the first Basic Law that sounds in my ears. "Article 1: (1) Human dignity shall be inviolable. To respect and protect it shall be the duty of all state authority."

whatever reason. If I see what a majority of people considers interesting on digg, I have my concerns about relying on a self-regulatory mechanism to work out. Relying on the good will and rationality of a majority of people is a decent approach, but it is naive. I hope that by now you see how much power lies in the hand of those who order, filter and structure our information, and that this has an impact on our political system, the opinion making process, and the decisions that we reach. In my nightmares I see the President's consultants putting together their briefings by Googling some keywords. Yes, maybe users would indeed just chose different companies, and everything would work out fine, but do you want to rely on it? People make mistakes, and the majority doesn't always make the right choice. That's why we have a representative democracy [5]. That's why our countries have constitutions that can't be changed from one day to the other. That's why we have laws to protect our freedom, that's why we have "executive and judicial officers" that are "bound by oath or affirmation, to support this Constitution"[6]. We shouldn't privatize part of the executive and we shouldn't hand over filtering information to private companies. It is almost tragically comic to me that all concerns about Google that I find in the media circle around economical power (3):

> One senior executive at Time Warner, who did not want to be identified, because Time Warners AOL division is a Google partner, says, "Sometimes I dont know what to think of Google. We have the best relationship of anyone with Google. On the other hand, you always have to worry when someone gets so much more powerful than all the competition out there. This is why I come down to this: I hope the government starts understanding this power sooner rather than later."

Yes, I too hope the government starts understanding this power - the political power.

IIIb. Wikiarchy

I find it quite interesting to follow the developments of non-profit collaborative projects like Wikipedia. Wikipedia has certainly proved useful, and in my perception the quality of articles has tremendously improved over the last years. It

---

[5]The overwhelming majority of visitors on my blog comes from North America and Central Europe. Apologies to those who can not identify with my use of 'we' when it comes to the political system.

[6]The basics of the politial system are quite similar in Germany, with exception of the President's role and details of the election processes.

has been a while since I came across a statement that I could immediately tell was blatantly wrong. It is a good quick reference, and I prefer it often over other sites if only for the simple reason that the sites are well maintained, easily readable, and cross-referenced.

However, here as much as with Google, I find the influence exerted by these sites worrisome, because I believe many people are not sufficienly aware of this influence, especially the younger generation. In a certain sense, Wikipedia seems to appear very trustworthy and likable, up to the point that I have to feel bad for criticising it and expect some comments to vehemently defend it. Isn't it after all a community project? Anti-authoritarian? Democratic? GOOD in capital letters? Danah Boyd from Many 2 Many expressed her concerns as follows:

> "My concern - and that of many of my colleagues - is that students are often not media-savvy enough to recognize when to trust Wikipedia and when this is a dreadful idea. They quote from it as though it cannot be inaccurate. I certainly distrust many classic sources, but i dont think that an anti-elitist (a.k.a. lacking traditional authority and expertise) alternative is automatically better. Such a move stinks of glorifying otherness simply out of disdain for hegemonic practices, a tactic that never gets us anywhere."

> Danah Boyd, Academia and Wikipedia, Jan 4th 2005

It's not only a tactic that doesn't get us anywhere, but a tactic that can simply go wrong for the same reason I mentioned above: the majority isn't always right. Wikipedia works as long as those who are not experts realize they are not experts, know that they don't know, respect the rules and don't execute their potential power. In his article 'Digital Maoism' (4), Jaron Lanier formulated his concerns like this

> "The problem is in the way the Wikipedia has come to be regarded and used; how it's been elevated to such importance so quickly. And that is part of the larger pattern of the appeal of a new online collectivism that is nothing less than a resurgence of the idea that the collective is all-wise, that it is desirable to have influence concentrated in a bottleneck that can channel the collective with the most verity and force. This is different from representative democracy, or meritocracy."

I recommend you read the full article, it's one of the most insightful writings I've come across for a long time [7]. I find it so to the point, I'll borrow another paragraph:

---

[7]Except that the title doesn't make sense to me, I can't see any Mao in the game.

"A core belief of the wiki world is that whatever problems exist in the wiki will be incrementally corrected as the process unfolds. This is analogous to the claims of Hyper-Libertarians who put infinite faith in a free market, or the Hyper-Lefties who are somehow able to sit through consensus decision-making processes. In all these cases, it seems to me that empirical evidence has yielded mixed results. Sometimes loosely structured collective activities yield continuous improvements and sometimes they don't. Often we don't live long enough to find out."

The belief that problems will be corrected I find very nice because it puts faith in mankind, but I wouldn't want to rely on it. It is quite an interesting trend that people rely so much on the common sensus. It is a trend though that can go wrong exactly because it does not necessarily have a self-correcting mechanism. Relevant for people's decisions is not only the information they have. But how much information they believe they have, and how accurate they believe it to be. Combine that with the faith in Google and Wikipedia. Is there information not on the internet?

On the other hand, what I find very interesting in these developments is that Wikipedia isn't just a direct democracy, it has guidelines for editing, and de facto indeed has a power structure (5). This is interesting to me because it is pretty much like witnessing the formation of a political system. Though it is quite clear where it should go, if Wikipedia wants to remain a high quality source of information, and that's what you'll read in the next section.

Herr und Meister, hör' mich rufen!
Ach, da kommt der Meister!
Herr, die Not ist gross!
Die ich rief, die Geister,
werd' ich nun nicht los.

Lord and master, hear my call!
Ah, here comes the master!
I have need of Thee!
from the spirits that I called
Sir, deliver me!

IV. Representative Democracy

In the early nineties, I was a member of the social democratic party in Germany, and I was a strong believer in direct democracy. The internet was fresh and new, and it seemed to me like the perfect tool to make reality what I thought had been given up for practical reasons: decisions being made by all the people. It wasn't difficult to extrapolate that internet access would catch on like a fire and that in

some decades almost all households would be connected, fast, easily, with access to all the information they need to make decisions.

At this time I was really excited about it, registered the first domain I ever had (demokratie-im-netz/democracy on the web) and tried to get a critical mass of people behind me. I didn't get very far though. Essentially, nobody was listening to me. What I should have expected given that the average age of the people I was faced with was somewhere in the upper 50ies, and most of them had not the faintest idea what I was talking about. World wide what?

Either way, though I am not generally easy to discourage there were two reasons I gave up on this. For one, at some point I had to make a decision between politics and physics. The latter won, that's why I am today where I am [8]. The other reason was that I came to realize that the adjective 'representative' is an essential ingredient to the democratic system.

The tasks in our society have become increasingly specialized. Most jobs require a years long education. If you want a good performance you look for a specialist, for an expert, somebody who has experience where you don't. I wouldn't want to make important decisions if I don't have the time or the education to do it well. So we elect representatives to do this job, people that have a good qualification for this. (I am afraid though big part of the frustration with politics/politicans today is a result of the low expertise status in the government. What can I say. It's a democracy. You get what you vote for.)

Either way, the election of representatives is beneficial for two reasons. The one is to increase expertise in decisions, above what could possibly be done by all people - most of which have a day job and other things to do. The other reason is that people's opinions are easily influenced by events with large emotional impact, and are prone to irrational fluctuations on too short timescales. Lanier put it like this: *"One service performed by representative democracy is low-pass filtering."* The drawbacks of direct democracy, and the reasons why we today have representative democracies however haven't yet been fully acknowledged. Let me quote from the Google Corporate Philosophy (1):

> 4. Democracy on the web works.
>
> Google works because it relies on the millions of individuals posting links on websites to help determine which other sites offer content of value. Google assesses the importance of every web page using a variety of techniques, including its patented PageRank algorithm which

---

[8]Unpacking the boxes of my 5th move in 4 years, writing this blog post on a Sunday afternoon, while my husband is on the other side of the atlantic ocean. Wait, who ordered that?

analyzes which sites have been 'voted' the best sources of information by other pages across the web.

Determining 'content of value' and 'assessing the importance' of webpages is limited by the way the system operates, which presently does no use the advantages of the small adjective 'representative'.

The point is that the problems we are facing on the internet have already been solved. Read your constitution. What is missing are elected representatives whose task it is to pursue the majority's goals, formulated as a set of rules/criteria/regulations. I will give you a concrete example: What I would consider a useful search engine is one that has a rating of websites by varios criteria like accuracy, entertainment, visual appeal, whatever. I don't want such a rating to be done by everybody clicking on a scale of stars. I don't want judgement to be made by anonymous people, nor an algorithm, nor an algorithm modified by anonymous people.

I do not care how many links go to a site if it is only an echo of another article, or - even worse - contains nothing than advertisements and links to other sites. I want there to be a group of people who is ¡¿responsible¡/¿ to provide a certain quality, a group of people whose names are known, and who explain their qualifications, opinions and decisions (I don't mind pseudoanonymity if it suffices to prove expertise). The internet is the ideal tool to provide representatives with feedback, and a useful platform for people to explain their qualifications, and to convince a majority they are trustworthy.

The recently launched Wikia Search, a search engine wikipepia-style with open source algorithm and user feedback, is an interesting experiment. The concept doesn't really convince me though for the reasons that should have become clear by now. There is the possibility the user feedback will optimize popularity (like digg), and not quality in the site's ranking. And though popularity is one interesting criteria to order sites, it shouldn't remain the only one. The "wiki-style social ranking" they advertise doesn't seem to me to a sufficient guarantee that expertise will be increased among those who provide the feedback.

V. Summary

The Sorcerer's apprentice that I have quoted here is a poem by J.W. Goethe (6). The apprentice is excited about the power he has witnessed, and while the master is away plays around with the broom. Unfortunately, things don't go as expected. He involuntarily causes trouble he doesn't know how to deal with, eventually culminating in the famous line *"Die ich rief, die Geister, werd' ich nun nicht los,"* which

translates roughly into: *"I can't get rid of the spirits I called."* In the poem, the master comes to help and sends the broom back into the closet.

The internet is a great invention and a powerful tool. It has a large and increasing influence on our daily lifes, as well as on our opinion- and decision making processes that eventually affect the quality of our living. It is as much an opportunity as a danger. We should be very cautios to ensure that self-organized structures on the internet - that are presently (still!) operating mostly in legal vacuum - do not interfere with our political systems. And this potential very really exists.

"In die Ecke	*"Back now, broom,*
Besen, Besen!	*into the closet!*
Seids gewesen,	*Be thou as thou*
denn als Geister	*wert before!*
ruft euch nur zu seinem Zwecke	*Until I, the real master*
erst hervor der alte Meister!"	*call thee forth to serve once more!"*

References

[1] www.google.com/corporate/tenthings.html

[2] Weaver, K., Garcia, S.M., Schwarz, N., & Miller, D.T., *Inferring the popularity of an opinion from its familiarity: A repetitive voice can sound like a chorus.* Journal of Personality and Social Psychology, 92(5), 821-833 (2007).

[3] *The Search Party*, by Ken Auletta, The New Yorker, Jan 14th 2008.

[4] *DIGITAL MAOISM: The Hazards of the New Online Collectivism*, An Edge Original Essay by Jaron Lanier, 5.30.06, www.edge.org

[5] www.itwire.com.au/content/view/4666/127/

[6] *Der Zauberlehrling*, by Johann Wolfgang von Goethe – *The Sorcerer's Apprentice*, Translation by Brigitte Dubiel, german.about.com/library/blgzauberl.htm