Simone Ganß

# We are googled to death

June 2006

# 1 Internet research

Internet research has long become a central cultural technique and an unavoidable piece of advice is: „google it“. Besides reading and writing, communication in mother tongue and foreign languages, mathematics and natural sciences as well as socio-political education, access to information has become relevant for the-so called knowledge society. With the universal interface "search engine" the internet offers an unprecedented access to information. Never before there has been a broader, publicly accessible "knowledge pool", and never before there has been a flood of information complicating the finding of relevant information.

Everything worth knowing is known by Google? It may be true that Google delivers the most important peaces of information because - thanks to PageRank - the most popular pages are listed on top of the ranking. Inverting this assumption is much more discouraging: What Google does not know, is not worth knowing. Or in other words: Information that does not appear in Google's ranking, does virtually not exist.

## 1.1 Construction of reality

While "research" stands for "inquiry" or "investigation" in the journalistic sense, the internet research is developing more and more to a „construction kit“ for constructing reality.

If an author wants to support his theories or opinions by internet research, he will find competent partners in Google and Co. With a vast amount of information - written by the rather different authors - the chances rise that an author finds supporting documents for his thesis.

Just as the author Mathias Bröckers who wrote a book about conspiracy, conspiracy theories and secrets all around the terrorist attacks of 9/11. He has written about his research:

„To get the information in this book, I had neither to dispose of special relations, nor arrange myself with slouch-hats and turban bearers to clandestine meetings - all sources lie open. To find them, the search engine Google did me incalculable services.
[...]
A user who searches for publications in a certain context, has only to combine two or

three terms or names and gets results in seconds.“[1]

Regarding Bröcker's work some questions are suggesting themselves: To which extent does human "logic" create reality? Does Google create reality? Is reality necessarily the same as truth, or is truth what we assume to be real?

Bröckers joins various information which seem suitable for his conspiracy theories, and thus construct a consistent simulacrum of his reality. It has been easy to find "supporting documents" for this on the internet, a medium on which every user is both producer and recipient, and in which every point of view has a chance.

„Bröcker's book (diary) is an excellent example for the fact that divergent and even contradictory sources must be first joined by a specific (human) logic to make a consistent image that adds up to the intention of the subject / author.“[2]

Finally, many "right/ real" images can be composed from the huge number of information. The question whether there is  a "real" image is hard to answer. Search engines can be used as a starting point for a search which can be used as a basis for every conceivable theory. Internet research is only made possible by billions of websites which are indicated by Google and Co. Without search engines the gigantic potential of the Internet could not be used.

### 1.2 The googled reality

People act according to their knowledge of reality. If the information space Internet is compared to other information spaces, there is often made a distinction between the "virtual space" and the "reality“. Both spaces have lbeen making up reality for a long time. Research with search engines is everyday practise to „check something up“, to purchase objects by auction, to read newspaper, to see stock market exchange rates, to rummage through apartment advertisement or „to google“ unpopular colleagues for most people. Information found this way is integrated into their own image of reality. Reality and truth based on that information are created if the source of information is accepted.

1 Mathias Bröckers: "Verschwörungen, Verschwörungstheorien und die Geheimnisse des 11.9.", 35. Auflage, Zweitausendeins (Frankfurt am Main),  2002, S ??

2 Michael Schetsche: "Die ergoogelte Wirklichkeit", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.),  1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 114

Search engines offer access to this information which is constructed to form a reality. Indeed, every internet user can publish his contents to other users, but they cannot be found if they are not indexed by a search engine. For most users the way to internet starts with the use of a search engine.

The first page of results is obviously the most important one. Then the second one. But who is going to look at the fourth page, or even the tenth? What can't be found by Google and Co. and does not appear on top of the list of results is virtually not existing for the practice of knowledge. It is not registered by the user and is therefore not real. This information does not add something to reality. The knowledge about reality originates from the use of searching technology. Thus search engines are developing from helpful tools to machines of reality and power.

### 1.3 Machines of reality and power

The machine of reality and power decides on indexing or not indexing, matching or not matching the query, and arranges the results logically. This logic is, as in all machines, based on an algorithm. However, the user doing the search has a logic of his own with which he is processing the machines-generated result. His logic is not based on algorithms, but associations. He tries to sort the results by asking himself whether or not the results are fitting the intention of his question. Even rather divergent sources can be joined by "human logic" to a consistent image. An increasing number of presentations at school and universities and even non fictional books as Bröcker's are written this way. There seems to be a strong need to join scraps of information to a greater whole.

In this context the user's intention is of vital importance: He will always join the information from different sources to an image appealing to him. A user intending to write an article about the Scientology sect and strongly deprecates it, will prefer sources which likewise publish a negative attitude towards the sect.

There is not only one „correct" image. Different perspectives and intentions lead to different ideas deriving from the same information. Thus a number of different consistent images based on the user's beliefs emerges.

Does this make knowledge nothing more than a collage of "matching" information? A collection of information that fits to what the user is believing. Do we rather believe

instead of know?

„There are some ideas so wrong that only a very intelligent person could believe in them.“ George Orwell

### 1.4 Knowledge versus information

At this point we can conclude that the term „knowledge“ is defining a stock of data, while information is both required and acquired in a more urgent way. Thus it would be more precise to talk of an information-based society rather than of a knowledge-based society when referring to the internet.

Brockhaus encyclopedia defines "knowledge" the following way:

„Assisted by knowledge an individual is not only enabled to understand the world, but also to affect its environnement selectively. Despite being the result of a sequence of learning processes, knowledge is also a prerequisite for processes of learning, thinking and solving problems. Knowledge acquisition is dependent on both cognitive and motivating factors and social processes.“[3]

Knowledge as a stock of data helps people to classify the world and the complex connections within it. Knowledge as a resource sets up a relation to reality, defines reality in general. Knowledge is the medium of our experience of the world. Thus knowledge is not delivering facts, but interpreting observations.

Brockhaus encyclopedia defines „information“ as:
„general term for communication or message with news value“[4]

Information disposes of uncertainty or ignorance, initiates learning and (immediate, future or/and lasting) behavioral change. Information can only be useful when it is well accessible. In contrast to knowledge information is rather fragmentary in its structure. Information cannot interpret, it can be only the foundation of knowledge.

Today's search engines deliver fragmentary results, as they can only search for

3 Stichwort: "Wissen", In: "Der Brockhaus multimedial 2005", Version 7, Bibliographisches Institut & F. A. Brockhaus AG, Mannheim, 2005
4 Stichwort: "Information", In: "Der Brockhaus multimedial 2005", Version 7, Bibliographisches Institut & F. A. Brockhaus AG, Mannheim, 2005

strings of letters. They cannot arrange documents in a complex context, they produce a fragmented simulacrum of reality. Thus search engines can only deliver information - not knowledge!

What the so-called "knowledge-based society" needs, is knowledge, not information. What the search engines of the future will have to deliver, is far beyond the search for letters. It is important to see the information that is found in its context. E.g., details on the author - who is the author, which further publications did he do, where was the document published for the first time etc.
The author of a web document is has the responsibility to provide this information and to communicate with meta-information.

Information is the structure and communication of the meta-information is the process. The interaction of both factors makes up knowledge. This step cannot be made by technologies. It requires human logic, because information is not intended to be used by machines, but by people. Sociological methods will become relevant in this context. Only this way it will be possible to program search engines in future which can show complex connections, link information associatively, and thus bring it into a shape understandable for human beings.

Knowledge is a of a higher quality than information. Knowledge is enriched by personal experiences, associations, morality as well as by socio-cultural background.

"The computer and its information cannot answer any of the fundamental questions we need to address to make our lives more meaningful and humane. The computer cannot provide an organizing moral framework. It cannot tell us what questions are worth asking."[5]

Although search engines only deliver information and the internet has not become a knowledge-pool by the use of the search engine technology, it is still much more than a big computer network.
Its structure is changing the cultural ways of handling knowledge, because the unique accessibility of information influences the development of knowledge. The way of how information is dealt with has changed since the the network medium internet has become more widespread:

5 Neil Postman: "Informing Ourselves To Death", German Informatics Society, 11.10.1990, Stuttgart, In: "http://www.eff.org/Net_culture/Criticisms/informing_ourselves_to_death.paper" Stand: 22.03.06

- information is created differently, that is digitally
- information is accessible more easily
- The search for information is mainly done by using search engines
- The representation of information is adapted in a way that they are found more quickly
- Information is processed more quickly, even though often fragmentarily
- information is spread quicker

### 1.5 Society of believers

People come to belive things where they cannot comprehend. They construct realities because there is no all-embracing truth. For this neither the internet, nor the search engines are responsible. It is a cultural phenomenon of the modern times. Beginning with the Reformation by Martin Luther 1517 an epoch, in which people lived in a clearly structured truth- and reality-structure was replaced: The Middle Ages. The papacy stood above all, a guilty conscience could be calmed by the purchase of a letter of indulgence, one did not have to read, because the priest in the church told all the facts worth knowing. The earth was the center of the universe and everything happened according to God's intention.

Today the earth is one of many planets in the universe. There is no divine plan, but earthly chaos. Anything is possible, anything can be true: people can believe everything they choose to believe.
And the search engine belongs since the middle of the 1990s to the most important instruments for constructing a reality of choice.

"The point is that, in a world without spiritual or intellectual order, nothing is unbelievable; nothing is predictable, and therefore, nothing comes as a particular surprise." [6]

#### 1.5.1 It's magic

Search engines and other web applications are mysterious for most users. Algorithms are „black boxes“ whose way of working remains concealed. Everyday vast numbers of people use a technology which they do not understand making it a central source for information and knowledge. And, in fact, the strategy of

6 Neil Postman: "Informing Ourselves To Death", German Informatics Society, 11.10.1990, Stuttgart, In: "http://www.eff.org/Net_culture/Criticisms/informing_ourselves_to_death.paper" Stand: 22.03.06

search queries is resembling methods of magic: Charms and spells for the magic world of the internet. „Words of power“ to conjure the search engine. The parallels are quite obvious: Only the right way of linking the right words will lead to the desired result. This is how search engines are working, aren't they?

This is how Google and Co. become magic machines and the user an everyday magician.
If these search strategies are effective - whether in the virtual world of magic or the "real" world - is questionable.
Concerning the knowledge there is no division in a virtual and a real world. Information from the internet is also having an effect on the practice of knowledge beyond the net. If a user does not know how to arrange and evaluate information on the internet, he will lack this competence beyond the net as well. The internet does not create a second, virtual world, but a joint reality. It is therefore important to examine people's practice of knowledge on the internet, because the still very young medium has already gained great influence on the information society and knowledge-based society and will increase this influence further.

### 1.6 New media structure

The very young medium internet brought forth an absolutly new structure of media which has to be seen as the cause for a new type of creating knowledge and reality. Network media created the so-called polydirectional isomorphic media.[7]

The word "isomorph" comes from the Greek and means „from the same guise or structure“, "poly" (in Greek) means „ a lot“ or „more“, "direktion" stands for "direction"[8]. Communication with polydirectional isomorphic media takes place in several directions. This arises from the network structure of the internet: A server is connected to many other servers which are likewise linked with others. The elements of communication are of the same structure. In spite of Flash, images, films and music the internet is mainly text-based. The description of the appearance of the website is determined by writing - the markup language HTML.

From this new type of media new social structures are arising: On the internet every

7 Michael Schetsche: "Soziale und kommunikative Ordnungen", In: "Netzwerkperspektiven", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, Roderer, Regensburg, 2003, S. 213-223
8 wissen.de GmbH: Stichwort „isomorph“, In: „http://wissen.de/wde/generator/wissen/services/print,page=3378974.html“, München, Stand: 28.03.2006

user is both sender and recipient. This double-role has consequences for dealing with information and knowledge.

Whether information is classified as „true“ or „wrong“ often depends on social relations. If the newscaster of ARD makes a message, he is more likely to be trusted than the baker around the corner telling the most recent gossip.

On the internet everybody can give himself the appearance of competence. A self-employed person working alone in is home-office can speak on his website of "our company“ and „about us“ suggesting a somewhat bigger company. Checking the validity of a piece of information on the internet can be a rather difficult task.

# 2 Universal interface search engine

## 2.1 Statistical and cultural meaning of Google

On 9/7/1998 the search engine Google started operating. Google has had (and still has) a rather simple design: An input field for search terms and a childlike, colored logo. This was completely different from its competitor Altavista, who sacrificed the clarity of its homepage for an abundance of new features and services. Google's design definitely contributed to the success of the company.

Today clue to the large amount of theoretically available information on the internet, most user's way of searching leads across a search engine.
For more than 80% of the German users the search engine "Google" is the search engine of choice.[9] Many of them use the search engine exclusively, insteadof consulting alternative sources.

Taking a closer look at the search engine does not only show its importance as an instrument of research but also its cultural importance:
The 23-rd edition of Duden knows the search engine Google, as well as the verb "googeln" (engl. „to google“)[10]. This is, by the way, one of few cases in German in which a verb is derived from a brand.

9 Peter Decker: "Web-Barometer", In: „http://www.webhits.de/deutsch/index.shtml?webstats.html“, Stand: 21.02.2006
10 Duden: Das große Wörterbuch der deutschen Sprache in 10 Bänden: Stichwort: „googeln“, Aktualisierte Online-Ausgabe, In: „http://duden.de/“, Mannheim, Leipzig, Wien, Zürich: Dudenverlag 1999-2004, Stand: 29.03.2006

The main competitors Yahoo! and MSN Search hardly get two-digit proportional market share. And nobody „yahooes!“ for information and nobody says: „ msn for it“. These search engines do not have such a great meaning as Google, neither statistically nor culturally.

It is therefore no exaggeration to speak of a monopoly position of Google. Google monopolizes the access to information on the internet, because without the interface of the search engine a specific search is getting impossible due to of the sheer size of the internet.
People used to surf the net, today they google. Once the user drifted, clicked themselves from hyperlink to hyperlink and on the way they did or did not find relevant information. People who google want specific information, the more exactly and quickly found, the better.
Regarding the search engine as a new medium of information-access the monopoly of Google equals a monopoly of knowledge on the internet.

The universal interface offers access to a wide range of information. However, many users lack the knowledge for dealing with a search engine.

Some features:
A search is not limited sensibly
Too few search-terms are used
Too general search-terms are used
Too few search engines are used, often only just one

Unfortunately, many users (also experienced ones) are not aware of this problem. The creed: „I always use Google. That works! “ the ability to filter relevant from irrelevant information deteriorates. In fact, using the search engine in a wrong way counteracts its original purpose, that is filtering relevant from irrelevant data, as it is no longer seen as an assisting tool.

Most users are too comfortable to expand their research, to go beyond the first results or to consult other search engines.

Along with the growth of the internet the need for filters is growing as well. The users

want to get a first overview by using the search engine which is hardly possible when looking at statistic data:
The number of freely accessible websites is estimated at 10-15 billion pages[11]. This estimate does not include information about databases, dynamic websites and linked documents.

Only about 60% of the estimated 320 million German websites are found by Google[12]. One the one hand this is more than other search engines find, yet is but a fractional amount of the actual web. Despite all troubles and limits more than 90% of the internet users use search engines to find their way trough that fractional amount of information[13].

While new pages are permanently added, others are deleted. Compared to the speed of change, search engines react very slowly. Until a page that actually has been deleted is deleted from the Google index, some time passes. This is how the known phenomenon that a link refers to the famous page „ HTTP 404 - file not found“ occurs. The simulacrum of the web transmitted by the search engines is always a simulacrum of the past.
In spite of these „delays“ there is of course preserved a "recency" which is achieved by no other medium. A recency, that contains terabites of data.

### 2.2 Change via new cultural practices

Like every new technology, search engines bring both advantages and disadvantages. As an interface between humans and knowledge it brings the advantage of a global accessibility. At the same time users are annoyed by an unmanageable mass of information which often turns out to be irrelevant on the search for knowledge. The users get quick access to information, but have to invest time in selecting relevant information.

"The invention of the printing press is an excellent example. Printing fostered the modern idea of individuality but it destroyed the medieval sense of community and social integration. Printing created prose but made poetry into an exotic and elitist

11 "Statistik: Aktuelle Zahlen zum WWW", 01/2005, In: http://www.ub.uni-bielefeld.de/biblio/search/help/statistik.htm Stand: 03.03.2006
12 "Suchmaschinen-Verein sucht 'Wege zum Wissen'", 23.11.2004, In: http://www.heise.de/newsticker/meldung/53568 Stand: 03.03.2006
13 Hari Obermeier: "Gütesiegel für Suchmaschinen", In: http://www.politik-digital.de/text/econsumer/verbraucherschutz/guetesiegel.shtml, Stand: 03.03.2006

form of expression. Printing made modern science possible but transformed religious sensibility into an exercise in superstition."[14]

Today new media and technologies like the internet and the search engine change cultural practice as Gutenberg's invention of the letterpress did in his time: The letterpress made it possible to publish information on a large scale and to spread ideas in a speed not known before.
The same applies to the internet: information can be accessed in split seconds. What can't be found immediately, is irrelevant for modern knowledge practises, what can't be found via Google is virtually not existing. As the knowledge of reality is based on the use of the search technology.

### 2.3 Gutenberg-galaxy versus Turing-galaxy

The change in the information-structures by the origin of the network media is cut into the knowledge practise. The Term "Gutenberg galaxy" (coined by Marshall McLuhan) faces at this point to the term "Turing galaxy". The Gutenberg galaxy stands for the epoch which is characterized by the leading medium "book". The Turing galaxy stands for the epoch in which the book is removed as a leading medium. An extended term of medium and causalities between different media are importent in this context.

#### 2.3.1 Gutenberg galaxy

The signs of the Gutenberg galaxy were:
- Linearity (e.g., the the letterpress in contrast to hypertext)
- Textualism (the document has properties, - e.g., several sentences which have a connection in form and content – which make the text clearly recognizable as a text.)
- individual knowledge production
- Authorship
- Clarity
- binary construction of reality (clear separation of producer and recipient)

Like every new medium the book as a leading medium of the Gutenberg galaxy has influenced the knowledge practise and construction of reality of society. This influence was not to be foreseen by the fathers of the medium, sometimes it was

14 Neil Postman: "Informing Ourselves To Death", German Informatics Society, 11.10.1990, Stuttgart, In: "http://www.eff.org/Net_culture/Criticisms/informing_ourselves_to_death.paper" Stand: 22.03.06

not even wanted. Neil Postman writes about Johannes Gutenberg:

" [...]Johann Gutenberg was by all accounts a devoted Christian who would have been horrified to hear Martin Luther, the accursed heretic, declare that printing is "God's highest act of grace, whereby the business of the Gospel is driven forward." Gutenberg thought his invention would advance the cause of the Holy Roman See, whereas in fact, it turned out to bring a revolution which destroyed the monopoly of the Church."[15]

The Gutenberg galaxy, that is the invention of the letterpress by Johannes Gutenberg, was a development which has brought many advantages to mankind. However, like every new technology it disposed of old values and cultural practices and has thereby had a massive influence on the knowledge practise of the society.

In his book „ the Gutenberg galaxy“ McLuhan predicts that the world will change into a "global village“ which has its origin in the used electronic mass media. He uses the term „global village“ to refer to the modern society which has overcome spatial and temporal barriers of communication by using electronic mass media, and is thus transforming into a "village". The term is often used today as a synonym for the internet, which is, in wide parts, fulfilling the prediction of the „global village“.

### 2.3.2 Turing galaxy

Signs of the Turing galaxy:

- Diffusion (dispersion of the sources and reflections of information)
- Intertextuality (texts are linked (via by hyperlinks in the web) with each other, the texts are no longer standing for themselves)
- dialogical and collaborative production of knowledge (knowledge is produced by communication and collaboration)
- Ambiguity
- Hyperreality (simulacrums of something not existing in reality are produced)

In 1995 the German social scientist and media expert Volker Grassmuck

15 Neil Postman: "Informing Ourselves To Death", German Informatics Society, 11.10.1990, Stuttgart, In: "http://www.eff.org/Net_culture/Criticisms/informing_ourselves_to_death.paper" Stand: 22.03.06

suggested "to call the emerged horizon of the binary-digital media "Turing galaxy" as it was Alan Turing who formulated its two central concepts [...]".[16]

Alan Turing (a British mathematician, logician, and cryptographer, one of the most influential theorists of the early computer development and informatics) is the author of two concepts which are fundamental for today's definition of media:

1. The Turing machine disposes of the problem of always having to build new machines sharing the same size and completely by using a limited set of commands for the universal Turing machine, thus transforming it into the new machine.
Theoretically, the human brain also is a limited machine and thus could be emulated by a universal Turin machine. The human beings using a Turin machine establish a new type of human beings which use the computer as they thinking-device.

2. The Turing-Test:
This test is supposed to provide an answer to the question of whether or not machines have the capacity to think.
It involves a human being sitting in front of a computer. The participant types in questions, which are answered by a machine or another human being. After a set time the person asking the questions has to tell which questions were answered by the human being and which ones were answered by the computer. Up to today there has been no program which has passed the Turing-Test.

### 2.4 Search engine darwinism

The search engine technology is a powerful one. It structures the perception of the world by defining of which website matches a search-term best, even if there is a lot of advertisement at the top of the list. For about 70% of the users only the first hits are relevant.[17]
For those who succeed in adapting their website to the algorithms of search engines, will find their site on top of the ranking-list. Survival of the fittest – a kind of search

16 Volker Grassmuck: "Die Turing-Galaxis. Das Universal-Medium auf dem Weg zur Weltsimulation". In: Lettre International, deutsche Ausgabe, Heft 28 (1. Vj. 1995), S. 48-55
17 Nadine Schmidt-Mänz, Christian Bomhardt: "Wie suchen Onliner im Internet?", 02/2005, In: http://www.absatzwirtschaft.de/pdf/sf/Maenz.pdf Stand: 03.03.2006

engine Darwinism. There is a big demand for books like „Internet Intern Suchmaschinen-Ranking optimieren“ and „ Besseres Webseiten-Ranking bei Google & Co“ with Amazon-sales-ranks of 11.399 and 7.366 (amazon.de: 3/30/2006). After all the own website has to be on top of Google's result, after all. The nature, the search-engine-darwinists are adapting their website to, is the market leader Google. The PageRank algorithm is the holy Grail everybody is searching for.

Despite everybody knowing that there is a high amount of advertisements on top of the list of results, the hits on top of the list are still perceived as the most relevant ones by the user, as stated above. This phenomenon is due to the people's absolute belief in science. After all, search engines have been developed scientifically and are based on technology which is mainly understood by information scientists.

"George Orwell remarked that the average person today is about as naive as was the average person in the Middle Ages. In the Middle Ages people believed in the authority of their religion, no matter what. Today, we believe in the authority of our science, no matter what."[18]

### 2.5 PageRank and the others

Search engines promise to treat pages just. But that does not mean that all pages are treated equally. Not all websites are equal before the search engine . The principle is rather a democratic one: each hyperlink is counted as vote. At least the market leader Google's search algorithm "PageRank" is said to work that way.

All websites are processed by the same algorithm. The algorithm treats the websites unequal insofar as popular pages become even more popular by a good positioning in Google:

„The heart of our software is PageRank™, a system for ranking web pages developed by our founders Larry Page and Sergey Brin at Stanford University. And while we have dozens of engineers working to improve every aspect of Google on a daily basis, PageRank continues to provide the basis for all of our web search tools. [...]

PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But, Google looks at

18 Neil Postman: "Informing Ourselves To Death", German Informatics Society, 11.10.1990, Stuttgart, In: "http://www.eff.org/Net_culture/Criticisms/informing_ourselves_to_death.paper" Stand: 22.03.06

more than the sheer volume of votes, or links a page receives; it also analyzes the page that casts the vote. Votes cast by pages that are themselves "important" weigh more heavily and help to make other pages "important."“[19]

Google judges the popularity of a page by counting the hyperlinks leading to it. Search engine-specialist Stefan Karzauninkat says that by doing so, Google eventually had a human ranking, despite its being done exclusively by machines.[20] Popular is, what achieves approval and is generally known. Just like a special kind of trousers suddenly worn by many people, or a band whose music is bought by many people. Popular is, what is compatible with a certain trend. The term "trend" stands for a complex, multidimensional, social phenomenon, enclosing many classes of population, and changes values, behaviour patterns, shopping behaviour etc. with lasting effect[21]. Regarding this, Karzauninkat is absolutely right if he speaks of a "human ranking“ through PageRank.

With PageRank a pioneering success came off in 1998 which made Google within the shortest time the most popular search engine. The PageRank algorithm is possibly protected as well as the recipe of Coca Cola.
With PageRank Google became the monopolist of the search engines. If there is only one (relevant) search engine, PageRank decides on to be or not to be. If there was a huge number of search engines there would also be a huge number of rankings. The results would thereby become possibly better.

John Kleinberg, an American mathematician, is one step ahead. He suggested that not all hyperlinks should be counted, but only those who come from pages that deal with the subject of the search query[22].
But such an algorithm needs a long arithmetic time for the search query. Google delivers thousands of results in less than one second. This is possible by the fact that the search process is independent of the search term in this respect, as that only one level of the interlinking is determined by PageRank.
With Kleinberg's algorithm the calculation is depending on the search term and must be determined again for every new query. The mathematical challenge for PageRank

19 „Our Search: Google Technology“, In: „http://www.google.com/technology/“, Stand: 02.05.06
20 Christoph Drösser: "Ausgetrickst und zugemüllt", 18.03.2004, In: DIE ZEIT Nr.13/04: "http://www.zeit.de/2004/13/C-Google?page=all" Stand: 24.03.2006
21 Stichwort: „Trendforschung“, In: „Brockhaus Multimedial 2005“, Bibliographisches Institut & F. A. Brockhaus AG, 2005
22 Heike Faller: "David gegen Google", 06.10.2005, In: DIE ZEIT Nr.41/05: "http://www.zeit.de/2005/41/Suchmaschinen_2?page=all" Stand: 25.03.2006

competitors is to find a formula which shortens the arithmetic time.

Apostolos Gerasoulis is the founder of the fourth-biggest American search engine "Ask Jeeves". The machine receives about ten million search queries a day. Gerasoulis is professor for applied mathematics at Rutgers University, the national university of New Jersey. He is a an expert for algorithms and works on a mathematical solution to accelerate the Kleinberg-idea. Should he succeed, the Kleinberg-algorithm could become a serious competitor for Google's PageRank.

### 2.6 Censorship

Not only the different algorithms of the search engines are responsible for different search results and rankings. The search engines must submit to the laws of the countries in which they are offering their service. Thus the results of google.com and google.de differ because Google has to stick to the German constitution. Accordingly, contents which deal, e.g., with the glorification of national socialism or certain forms of pornography are censored in Germany[23]. In accordence to German law Google has to delete pages from its index which could tempt to commit a criminal offence, or provide inciting or endangering content, e.g.:

The Committee for Open Debate on The Holocaust
www.codoh.com
Holocaust denying contents

Skinheads of the racial holy war
www.creator.org
Racist contents

Danmarks Første Patriotiske Hjemmeside
http://www.patriot.dk/
Racist contents

These and similar pages are not listed in the Google-results referring to legal reasons. The filtering of the contents in the search engine's index happens in collaboration with "Chilling Effects Clearinghouse", developed by the " Berkman centre for Internet and Society “ and supported by programs of the universities of

23 Klaus Platzwaldt: "Suchmaschinenlandschaft", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 80

Harvard, Berkeley, Stanford, University of San Francisco, University of Maine, George Washington School of Law, Santa Clara University School of Law clinics and EFF.

„Chilling Effects aims to support lawful online activity against the chill of unwarranted legal threats.“[24]

Besides legal restrictions there are also restrictions made by search engines themselves, like in the case of the German version of MSN Search, which does not deliver results of queries related to sexually associated queries.[25]

These conditions whether by law or by the search engine providers, are not published. The quality of the search-results suffers from the fact that the evaluation criteria of the operating authorities are not made public and, thus are not discussed by public. A medium filtering the general public's access to information should be more open to public discussion, but as the publication of censorship criteria would also publish information about the algorithms, it would reveal the secret of success and by that endanger the economic future of the search engine.

### 2.7 Commercial foundation

Today, all successful search engines work on commercial basis, even though the models of founding are different. Google finances itself to 95% from advertising. A main source of income are the so-called "Adwords" - those adverts which fit to the entered search term and are placed on the right - beside the results.

In contrast, Yahoo! Is increasingly focussing on commercial services like DSL internet access. This way Yahoo! has a financial basis which is more independent of advertising.

The association for the support of search engine technology and free knowledge access "SuMa-eV" (www.suma-ev.de) thinks of commercial search engines as a problem. The main aim of the association founded in July 2004 is promoting knowledge access on the internet regardless of commercial interests.

---

24 Chilling Effects Clearhousing: „About Us“, In: „http://www.chillingeffects.org/about?print=yes“, Stand: 01.04.2006

25 Alfred Krüger: "A tergo – Sex durch die Hintertür bei MSN" (18.11.2004), In: „http://www.heise.de/tp/r4/artikel/18/18836/1.html“, Stand: 22.02.2006

And free knowledge access is indeed threatened when economic factors gain too much influence. Last year Google's turnover has almost doubled to 6.1 billion dollars, its profit has risen by the factor 3.5 to 1.46 billion dollars. And when at the end of February 2006 the financial executive George Reyes explained that Google would not grow that fast forever, Google's stock exchange rate declined by about 13 percent.

„Since it's highest level of 394 euros in January the stock has lost about 20 percent because the uncertainty about Google's future has increased. The company is gaining 99 percent of its turnover with online advertising. And although there are new services they still bring no money."[26]

This shows how the access to knowledge depends on the economic situation. With the headlines of price declines memories of the dotcom crash when companies disappeared as quickly as they had appeared, come alive. Economy experts say that the Google stock was overrated anyway. And if the announcement that Google will not always grow as quickly causes a fall of prices of 13 percent, what will happen when Google will not multiply its profits as usual? Which consequences will that have on knowledge practise?

## 3 Global library of knowledge

The internet as the library of human knowledge – a powerful vision which many people worldwide are trying to realize. The advantages this library of knowledge are obvious: In theory this library will be available anywhere in the world, everybody can make his knowledge accessible easily and information from the library is easily accessed.

### 3.1 Producer and recipient

The Internet has quite a special characteristic in the organization of knowledge: The abolition of separation of producer and recipient.
There are no authorities like in big media corporations, there is no "pre-censorship" sorting out content before publication. There are no commentators who quote the sent material, there are no presenters who lead the viewer through the content of a program. Nobody says what is to be sent and what is not. There is no one in charge of the program, there are no broadcasting times.

26 Götz Hamann: "Expansionsgelüste", 09.03.2006, In: DIE ZEIT Nr.11/06: "http://www.zeit.de/2006/11/Kasten_Google", Stand: 24.03.2006

Every user makes his own decision on which information he is transmitting and at which he will have a look at. Everyone is both producer and recipient at the same time and every point of view has a chance. Great deviations from the common sense can make information even more appealing.

However, the drawbacks of this achievement are obvious, too: A mass of information which is not easily grasped; contents which can't be found by search engines; pornographic, racist and criminal content, a lack of permanence of content.

### 3.2 Alternative concepts

There are many alternative concepts for structuring the internet. In the following some theoretical attempts will be introduced in the following. Many of these concepts are dealing with the question whether or not the Internet is in need of a controlling authority.

Prof. Dr. Marcel Machill from Leipzig university is involved in the development of a "code of conduct"[27]. Which involves that search engines should put themselves under certain obligations, a kind of law for search engines. Unfortunately, the rules of the code of conduct would need a neutral control authority. And the question who could be that authority is still not be answered.

Another model is the "public law search engine“ which got into public discussion by Dr. Wolfgang Sander-Beuermann of the computer center Hannover[28]. As with the more traditional media the state would take over a leading role. However, the state must not be a monopolist. With this kind of a search engine the delivered contents would be covered by committees and advisers.

Nina Degele puts the question: „Will classic encyclopaedias the future pilots through the sea of the knowledge?“[29]. Many users already go to wikipedia instead of using search engines to do research. They get an overview about the searched subject

27 Interview mit Prof. Dr. Marcel Machill aus Kai Lehmann: "Blackbox Suchmaschine, Politik für Neue Medien", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 57

28 Interview mit Dr. Wolfgang Sander-Beuermann aus Kai Lehmann: "Blackbox Suchmaschine, Politik für Neue Medien", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 59

29 Nina Degele: "Neue Kompetenzen im Internet", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 72

there without having to filter advertisement and inappropriate information. The articles of the online encyclopaedias are linked well and lead to the desired information more quickly, than search engines. It is very well possible that research will start at Wikipedia, Brockhaus online and Encyclopedia Brittanica instead of starting at Google, Yahoo! and MSN.

Additionally there are hundreds of alternative search engines which want to change the access of knowledge, "Nutch" and "Yacy" beeing two of the most popular projects.
"Nutch" is a project promoting free search engines. It is based on an open source software which can be used for commercial as well as for non-commercial search engines.
"Yacy" is a project on P2P basis. Provided a sufficient number of users installs the software on their computers, the result would be a global search engine with a decentralized index.

Many of these projects share the characteristic that they involve their users in the process of developing their engines, the search engine index is for example compiled by users. Nevertheless only ambitious users with at least basic programming skills and an understanding for the structures of search engines can participate and thus influence and control the search engine.

The concept of these search engines is therefore the opposite to the development of the use of the internet, because search engines are primarily just used - only few people understand them or even have the knowledge to program or develop search engines. The search engine remains a black box for most users – even by taking into account new search engine projects. A black box which is used, but it is not understood.

What the average user needs, are intelligent search engines which can recognize what the user really has meant with his question. A machine which can understand, how words are connected, which can find connections on subjects, which has of linguistic abilities, far beyond Google's "did you mean...". Briefly: A semantic web.

### 3.3 The Semantic Web

The idea of the semantic web is not a new one, but it's development is still in its

infancy. Nevertheless, the semantic web could be an answer to insufficient search results and chaos of documents on the internet.

The basic idea is quite simple: Meaning is assigned to all information on the internet in form of meta-data. This meta-data are readable for machines. That means, that search engines can search on the internet for semantic connections instead of scanning the index for character strings.

With the semantic web a search engine would recognize that the search query „Tokio Hotel“ has different semantic connections, e.g., "hotels in Tokyo“ or the German band „ Tokio Hotel “. On the question „ Who was German Chancellor in 2004“ such a search engine would deliver the answer "Gerhard Schröder", while today's search engines deliver newsletters, newspaper articles and Wikipedia entries in which the searched terms may occur without any connection.

The semantic web generates connections between disordered information by processing information about the information. The internet could become a network of real information which would finally be more than the sum of its data. A search engine which can - at least roughly - "understand" the structure of the human mind.

### 3.3.1 Technologies for the semantic web

To make the semantic web work, different technologies which, combined with each other, allow a semantic search are required:

XML - Extensible Markup Language
A standard for the creation of documents in form of a tree structure. The elements of a document are sorted hierarchically with XML. This way, the document receives a syntactic frame and a defined content.

URI - Uniform Resource Identifier
A uniform identifier for resources. URIs are used to the name resources on the Internet. URI determines the exact location of a document in the web.

RDF - Resource Description Framework
RDF is a language which serves the publication of information. The framework has been designed to support the activities in the sphere of the semantic web by

the W3C. RDF is intended to be a basic format for the representation of taxonomies and ontologies - formal vocabularies generally. RDF joins meaning, location and content.

The semantic web offers advantages which today's web cannot offer. Indeed, it is more of a concept for closed user groups. A new Google service à la "Google Semant“ is hardly imaginable, because global meanings for documents are hard to define.
The classification of documents in a hierarchy of meanings requires - to be worldwide understandable – an editorial staff, which would conflict the idea „Everybody is a producer and recipient at the same time“.

### 3.4 Search engines and media policy

As is mentioned above every user on the internet is both producer and recipient. For the majority of documents published on the internet applies, that the author has published his document out of his own initiative, no editorial staff has defined or checked its content, no executive has sanctioned the document. The loss of this controlling body makes it hard for many users to arrange content, to evaluate their relevance or validity. Other users don't even have any awareness of the problem.

"The loss of editorial filters and controlling bodies is not perceived by the majority yet - the media-political classification of search engines hardly takes place."[30]

As the media structure of search engines is hardly reflected by users, it is not surprising that search engines have gained a huge influence on the rest of the media scenery.
Controlling levels like editorial staffs of online newspapers are heavily influenced by Google and Co.. A good example for this is the arrest of the hacker responsible for the computer worm "Sasser", in a village called Waffensee in 2004. The following introduction to an article dealing with the incident was published by Spiegel online:

„In the village Waffensee near Rotenburg an der Wümme in Lower Saxony the world still seems to be in order: The 'Eichenhof' inn lures with a comfortable chimney, farmer Poppe sells meat and jam from own production on the corner , and up to now

30 Hendrick Speck, Frédéric Philipp Thiele: "Goggle, Gossip und PR-ostitution", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 181

only the Shanty Choir brought a hint of the big, wide world to the North German 'Heideflecken'"[31]

If one searchs with google.de for the place "Waffensee", one receives in 0.13 seconds the following result ranking:
Position 1: Eichenhof
Position 3: Farmer Poppe
Position 12: Shanty choir

Google possibly becomes the new landmark of journalism. What Google finds on a subject, has to be taken into account by the journalist for his research. Insufficient research for articles can be veiled by searching Google. And more importantly: Google creates a mock-reality which is reflected by the media scenery.

In this context the question arises, how search engines will change the knowledge practise of the media scenery in the long run. The efforts of classic media corporations to face the internet are rather weak up to now. Every important newspaper has an online editorial staff publishing a web version of the newspaper. Television stations have websites on which users can sometimes watch bad streams. And there is internet radio as well. But that the internet-offensive of the classic media comes to an end at this point. Not only new media itself, but also the structure of its use and its users are a novel phenomenon.

Managers of traditional media companies are concerned about this development. The way internet users deal with media is the opposite of the mere viewer. This threatens the business model and with it the existence of television broadcasting stations, radio stations, magazines and newspapers, based on the idea that professionals are responsible for the content and the interested, but passive consumer is paying for it.[32]

Will information and knowledge no longer be generated by central institutions from now on? We used to have newspaper on the breakfast table, are we going to have a laptop with Google News in the future? We used to ask Brockhaus, we are going to

31 Jochen Wegner: "Die Googelisierung der Medien", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 235
32 Götz Hamann: "Die Eingeborenen des Internets", 16.03.2006 In: "DIE ZEIT Nr.12/06: http://www.zeit.de/2006/12/memedia?page=all" Stand: 23.03.2006

ask Wikipedia instead? We used to go to the library, are we going to use Google Print from now on?
Since January 2005 Google is scanning about 50,000 book pages a day. In 10 years time Google Print is planned to enclose a total of 4.5 billion book pages, cost 200 million dollars and is founded by advertising.
Google representative says that thy would like to make mankind's knowledge available to the general public.[33]

Even though Google does not yet charge for the use of the book pages, the president of the French National Library Jean-Noël Jeanneney sees the peril that one day only titles are digitized that Google's advertising customers pay for. The border between information and advertising which even today is hardly noticed by users would possibly become totally foggy.[34]

Again the tremendous influence of the policy of the search engines on the classic media scenery appears. It can hardly be said, which consequences on information and knowledge practice this will have. Nevertheless it is possible to make presumptions derived from an analysis of the structure of network media.

### 3.5 Unbound information

The Internet as a medium has a structural specific feature: Knowledge is stored in archives for an uncertain amount of time. Documents are updated, substituted or deleted in rather short intervals. New documents are added on a daily basis. The question: „Will the website still be there tomorrow on which I support my knowledge?“ can only be answered "maybe".

The reason for this uncertainty is that the internet freed information from its physical limitations. In contrast to books which are physically inert due to being printed and bound, a web document can be changed quickly and easily. Both performance and integrity of the document cannot be vouched for. Information can be loosened from its context. Information scraps are joined to new, possibly even absurd "information".

Of course that is not a new phenomenon, scientists always have gathered

33 Michael Mönninger: "Die Google-Bibliothek", 04.08.2005, In: DIE ZEIT Nr. 32/05: "http://www.zeit.de/2005/32/Google-Bibliothek?page=all", Stand:24.03.2006
34 Michael Mönninger: "Die Google-Bibliothek", 04.08.2005, In: DIE ZEIT Nr. 32/05: "http://www.zeit.de/2005/32/Google-Bibliothek?page=all", Stand:24.03.2006

information from several books, have summarized and compiled them to new books. This practise of knowledge is not novel to the network media.

But what has changed is the accessibility to this information and the availability of knowledge. A computer with internet access – and theoretically all information that were fed to the internet is available. A central knowledge and information pool to everybody can access directly.

Already in 1945 Vannevar Bush's vision of the Memex created a structure of the "Memex" very similar to the structure of the internet and its knowledge archiving. [35] The Memex is a machine hidden in a kind of desk which provides an improved exchange and access to information and contributes an extension of knowledge of the mankind.
Documents, like images etc. are available on microfilm in the Memex, enabling the user to take so-called "knowledge strings" from the context and to join them to new pieces of information.

Creating knowledge, collecting and to sharing it is an important characteristic of human beings. With the culture the access to and the dealing with knowledge changes: Every new medium has an influence on the cultural practise of knowledge and changes the availability of knowledge.

### 3.6 Origin of knowledge

At this point the question arises how knowledge is produced by the use of media in the first place and how information is dealt with.

Observing the internet the development of the new medium internet, one will find that the development of knowledge has not substantially changed: Knowledge is generated by causal chains. Pieces of information which, at least seemingly, connect well are linked in the brain. These connections of the nerve cells (synapses) can change: New connections can be formed between nerve cells and other, unused ones, recede. Most of the contacts or linkings of nerve cells are formed in the childhood. A newborn child has virtually no linkings between his nerve cells. That is why newborn children react rather slowly to stimulations. In the course of the first months of life nerve cells get increasingly connected.[36]

35 Vannevar Bush: "As We May Think", In: "Atlantic Monthly", 176(1):101-108, Juli 1945
36 Stichwort: "Nervensystem", In: "Der Brockhaus multimedial 2005", Version 7, Bibliographisches Institut &

If external influences can cause changes in the brain, regular use of media definitively belongs to these influences. The medium Internet with its novel structure is playing a special role. The above mentioned abolition of the physical restraint of information on the internet and the resulting instability, lead to a dissimilarity of information which following a kind of modular design principle: One is always fitting to another. And the access to these elements has accelerated.
If an author intends to write an article about the dangers of the sect "Scientology", he will be finding information with Google within one second. If he has the intention to write an article about the positive aspects of the community of "Scientology", he will also be able to access that information in less than a second with Google.

One always fits to another. Information is loosened from its context and is joined to new information. The product is published again (as seen with the Spiegel-Online article dealing with the arrest of the Sasser hacker). Thus the value of information is becoming increasingly provisional and information itself is becoming more and more manipulable.

From this structure information, one can conclude how knowledge is generated. If single pieces of information are pulled out of their context, are combined with other information and are re-composed, this is not specific to the medium internet. But the speed of this process on the internet is new.

If information has no permanence, knowledge is fragmented. Especially the use of the search engine technology leads to an accelerated permutation of knowledge. Every user is forming an individual knowledge. Everybody changes and recombines information - the instability of the "knowledge pool" has its roots in the fact that digital objects of knowledge are moving permanently.[37]

### 3.7 Indexing and assessing

Google doesn't index information systematic-hierarchically like classical search engines which sort the websites in their index by categories, but dissociative-egalitarian, and thus apart from categorical relations as elements of the same kind.

F. A. Brockhaus AG, Mannheim, 2005

37 Jeanette Hofmann: "Digitale Unterwanderungen: Der Wandel im Innern des Wissens" (2001), In: http://www.bpb.de/publikationen/MHNZMN,1,0,Digitale_Unterwanderungen:_Der_Wandel_im_Innern_des_Wissens.html Stand: 21.02.2006

Without a doubt a new order of knowledge originates from the knowledge indicated by search engines. Accessibility is rated higher than the relations of content.

The assessment of knowledge on the internet is no longer done via the code „true/false", but by accessibility. The information which is accessible most easily, will gain the highest attention.
Factors of economic relevance are not product, price and positioning, but attention, accesibility and relevance.[38]

Web contents are completely arbitrary, widely uncontrolled and unstructured. The authenticity of the content or the sender cannot be checked. The data is not organized in a linear way, there is no set broadcasting time, format or content organization. And with a virtually infinite amount of information search engines are representing the only guidance.

„People's inability to understand political or economic changes in an increasingly complicated world lets them grasp explanations where confusing structures of cause and effect and feedback-circuits are substituted by simple causes and causal chains.“[39]

Although the internet could theoretically provide all conceivable information which can be seen in exhaustive, complex connections, the user cannot manage these complex possibilities. However, the search engine – especially Google's search algorithm - is hardly an adequate means of acquiring knowledge.

## 4 Conclusion

38 Michael Domsalla: "Wissensmanagement und neue Wissensordnung", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 175
39 Michael Schetsche: "Die ergoogelte Wirklichkeit", In: "Die Google-Gesellschaft", Kai Lehmann, Michael Schetsche (Hrsg.), 1. Auflage, transcript-Verlag, Bielefeld, 2005, S. 116