

# SEARCH ENGINES

Selected data protection issues



Candidate number: 1

Patrick Unger

Supervisor: Lee A. Bygrave

Deadline for submission: 09/01/2009:

Number of words: 17,982 (max. 18.000)

27.08.2009

UNIVERSITY OF OSLO

Faculty of Law

# Content

<b><u>1</u></b>	<b><u>INTRODUCTION .....</u></b>	<b><u>1</u></b>
1.1	Presentation of the issues .....	1
1.2	Background .....	2
1.3	Demarcation of thesis .....	5
1.4	Legal method .....	5
1.5	Definitions .....	7
1.6	Structure of the thesis .....	9
<b><u>2</u></b>	<b><u>IP-ADDRESS AS PERSONAL DATA .....</u></b>	<b><u>10</u></b>
2.1	Problem statement.....	10
2.1.1	Composition and character .....	10
2.1.2	Dynamic versus static IP-address .....	12
2.1.3	A Log file.....	14
2.2	Legal analysis.....	14
2.2.1	Definition of “personal data” in light of Directive 95/46/EC.....	14
2.2.2	Working Party’s perspective .....	17
2.2.3	Google’s / Peter Fleischer’s perspective .....	18
2.2.4	Evaluation concerning a dynamic IP-address.....	20
2.2.4.1	A legal approach / Analysing European legislation.....	20
2.2.4.2	Interpretation of recital 26 of the Directive .....	26
2.2.4.3	Relevant agent of identification .....	32
2.2.4.4	Auxiliary elements contained in the identification process .....	34
2.2.5	Evaluation concerning a static IP-address .....	36
2.2.5.1	A legal approach .....	36
2.2.5.2	Auxiliary elements contained in the identification process .....	39

2.2.5.3	Intermediate result .....	42
<b>3</b>	<b><u>SEARCH ENGINE AS A “CONTROLLER” .....</u></b>	<b>44</b>
<b>3.1</b>	<b>Problem statement.....</b>	<b>44</b>
3.1.1	Working Party’s perspective .....	44
3.1.2	Google’s perspective .....	45
3.1.3	Statement on the issue .....	46
3.1.3.1	A practical approach .....	48
3.1.3.2	A legal approach .....	50
<b>4</b>	<b><u>CONCLUSION.....</u></b>	<b>54</b>
	<b><u>SELECTIVE BIBLIOGRAPHY .....</u></b>	<b>56</b>
	<b><u>ANNEX.....</u></b>	<b>A</b>
	Whois Record.....	A

# 1 Introduction

## 1.1 Presentation of the issues

The thesis elaborates on a topic which has attracted a lot of discussion in recent years and still is the subject of an ongoing debate. A big controversy has flourished between the company Google Inc. and several privacy groups' ostensibly led by the "Working Party on the Protection of Individuals with regard to the Processing of Personal Data".<sup>1</sup>

Deep data protection concerns have been raised in this debate by the use of a search engine. There exists a legal conflict on how to qualify the status of an IP-address in light of it being considered as personal data under European legislation.

The debate is continued in evaluating the status of a search engine provider in terms of classifying it as a controller under European legislation. The thesis will elaborate and focus on those data protection issues.

While examining on the legal classification of an IP-address several sub-issues are being touched upon. An in depth interpretation of recital 26 of the "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data"<sup>2</sup> will occur. The aim of the thesis is to present clear guidelines, especially if a lawful approach of the stakeholder involved is necessary when acquiring additional data on a data subject and if auxiliary material has to be taken into account when determining the identifiable criterion set out by European legislation. While examining those legal aspects it will be necessary to come up with clear guidelines who of the stakeholders involved should be considered as the entity responsible for the assessment of an identification process.

---

<sup>1</sup> Normally referred to "the Art.29 Working Party". I will use the abbreviation WP

<sup>2</sup> From now on referred to the Directive

## 1.2 Background

While using the services offered by Google, vast amount of data is gathered by the company in intention for reusing them for marketing purposes suited on the specific user's interest.

A reason for dealing with Google is their dominant position which they possess in the search engine sector. A study conducted of search share rankings in the U.S. in March 2009 shows an overall of 9,522,853 searches made, Google having a search query of 6,113,906, representing nowadays 64.2 percent of all search queries conducted during the given time period.<sup>3</sup> In Germany Google had a market share of 90% at the end of October of 2007, whereby the second rated search machine operated by Yahoo! had a market share of 3.2% and Microsoft's Live Search just reached 1.2%.<sup>4</sup> Those figures were nearly the same in March 2008 when Google reached a market share of 93 %.<sup>5</sup>

The WP<sup>6</sup> and several other privacy groups<sup>7</sup> have several times requested upon Google to change their data privacy protection, especially to refrain from storing data belonging to the data subject.

Due to the above described dominant market position Google has obtained in the search engine sector, the omnipresence Google has achieved combined with the ongoing debate between Google and the WP has persuaded me to intensively deal with Google.

Another reason is the circumstance that Google is the first company involved which publicly announced a new system of gathering data, a system called "behavioral advertising". In March 2009 Google proclaimed on their "Official Google Blog"<sup>8</sup> that they are of the opinion that ads are a valuable source of information. They proclaim that by making ads more relevant and suited for their user they can create a higher value for

---

<sup>3</sup> Nielsen//Netratings, [http://www.nielsen-online.com/press.jsp?section=ne\\_press\\_release&nav=1](http://www.nielsen-online.com/press.jsp?section=ne_press_release&nav=1)

<sup>4</sup> Kuehling, 2007, p. 881

<sup>5</sup> <http://googlesystem.blogspot.com/2009/03/googles-market-share-in-your-country.html>

<sup>6</sup> WP document 148, p. 3

<sup>7</sup> "Data Protection and Privacy Commissioners' Conference", 2006,

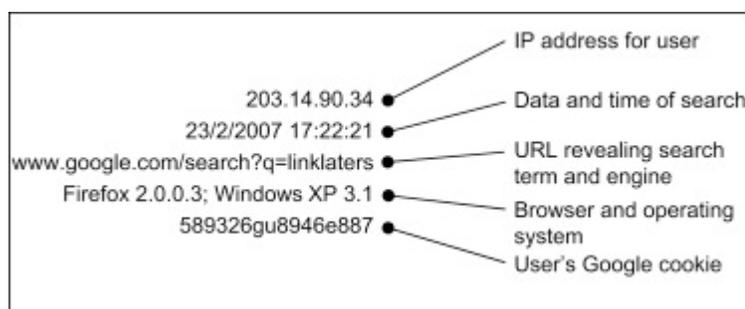
[http://ec.europa.eu/justice\\_home/fsj/privacy/news/docs/pr\\_google\\_annex\\_16\\_05\\_07\\_en.pdf](http://ec.europa.eu/justice_home/fsj/privacy/news/docs/pr_google_annex_16_05_07_en.pdf)

<sup>8</sup> <http://googleblog.blogspot.com/2009/03/making-ads-more-interesting.html>

everyone involved. By achieving this goal and by creating ads that are suited on the preferences, Google needs to create a profile of every single user. After the preferences are singled out the marketing business comes into play, targeting users with their preferences extracted out of the gathered data.

To create such a profile Google has to store the user's search term with the help of a server log. In the following a graphic will present the data contained in Google's log.<sup>9</sup>

**Fig.1: Search Log**



The graphic shows that the log file does not just contain of different search requests but also contains traffic data. The thesis will focus on the extraction and processing of the IP-address contained in the search log. The accumulation of all the above elements shown will allow identification of the user, thereby leading to the processing of personal data.

Google has stated that they as a mere website holder-opposed to an Internet Access Provider- are not processing personal data if just taking account of the extracted IP-address, since IP-addresses can not be considered as personal data in relation to their status as a website operator.<sup>10</sup> It is my opinion that it is impossible to come up with a clear indication that an IP-address as such -irrespective whether categorised as an Access-, Service- or Content Provider- can be considered as personal data. A more in depth analysis has to be approached considering technical, practical and legal interpretations of relevant norms.

---

<sup>9</sup> Church & Kon, 2008, p. 462

<sup>10</sup> Google's Response to the WP Opinion 1/2008, p. 8

IP addresses can be used as static and dynamic ones. I will stress that while there are technical differences in the operation system it is important not to lose sight of a legal approach too.

It is my opinion that a static IP-address can not be considered as personal data. A dynamic IP-address will neither reach the status of being considered as personal data. It is in these particular aspects that my work aims at contributing to the existing debate in how to qualify an IP-address.

The challenge of my work is that there is no broad discussion of this legal topic. Courts have been tentative in approaching a clear line which would lead to a final clarification. The discussion in academic literature is also far from more stringent either, showing more vague approaches in just examining an IP-address as such, not taking account of a legal separation in dynamic and static ones. The same approach was ascertained in European jurisprudence.

I want to highlight that the result of this examination can lead to a situation in which legal and technical aspects do not harmonize leading to a circumstance in which legal aspects are hard to implement with technical means.

Google further proclaimed that even if an IP-address should be considered as personal data the company still does not feel bound by European legislation. Google's EU-based data centres can not be considered as a controller like the European legislation requires it for being taken into responsibility.<sup>11</sup> To be classified as a controller active manual steps have to be involved in the operating system. Google's EU data centers are just used for the storing of data, in particular indexing web pages, so all processing activities are carried out by Google Inc. located in the US.

The thesis elaborates on what kind of action is taken in those EU-based data centers and if those actions qualify Google to be treated as a controller on the territory of a European Member State laid down in Article 4 (1.a) of the Directive.

---

<sup>11</sup> Google's response, op.cit.

### 1.3 Demarcation of thesis

Service Providers store, collect and process a huge amount of user data, thereby using technical aid by installing “cookies”. This means that the data gathered varies from the IP-address up to the extensive history of a searching behaviour of the user. The thesis will consider and legally analyse if the individual user’s IP-address can be considered as personal data, thereby not taking account of additional information stored on a “cookie” device. It will be analysed if it is possible to carry out an assessment just in light of taking account of the IP-address as such, or if European legislation requires that further auxiliary data has to be taken into consideration as well.

The scope of the analysis carried out in this thesis will be limited to the relevant material provisions in the European legislation available to the Member States in protecting the individual’s personal data. I will not address further European Directives in depth, nor touch on commercial regulations in this field.

### 1.4 Legal method

In discussing the legal interpretation of the term personal data in connection with the classification of an IP-address, thereby taking into account the different opinions of the stakeholders involved, my point of departure is the legal method the European Court of Justice<sup>12</sup> adopts when an interpretation of a legal norm is carried out.

The ECJ has jurisdiction to rule on interpretation of the Brussels Convention, where the ECJ has taken the purpose and teleological view to rule interpretation that characterizes its decision making generally.<sup>13</sup> The ECJ is of the opinion that concepts are to be interpreted autonomously of the meaning they are given in the domestic law of the Contracting States, although significant weight is also placed on the intentions of the drafters of the Brussels Convention.<sup>14</sup>

---

<sup>12</sup> From now referred as the ECJ

<sup>13</sup> Foss & Bygrave, 2000, p. 104

<sup>14</sup> Foss & Bygrave, op.cit.

The ECJ has not developed a maxim<sup>15</sup> which adheres to the wording of a legal norm, as long as no irrational reasons are visible which prohibits the justification of another criterion. Particularly there is no prima facie preference of a grammatical element. The ECJ more or less tends to expand the scope of a legal application even if it is not covered by the mere literal meaning. In their previous decisions the ECJ considered it necessary to adhere to the wording of the law, the context and its meaning. Nowadays the ECJ approaches another tendency which was developed in the 1970s to put weight on the mind, composition and wording of the norm plus recognizing the system and aim. This means that grammatical, systematical and teleological arguments compose an equal standard of evaluation in interpreting a legal term. The comprehension of European legislation, multilingualism, an autonomous concept followed by the dynamic character of Community Law seems it to be appropriate to put weight on the aim and the purpose of the relevant term and its contextual demand.

The thesis will work with this interpretation method when analysing legal terms.

While analysing the statements given by the WP a critical examination will occur, like with those given by Google and Peter Fleischer, due to their composition and legal coverage. The WP is composed of representatives from each Member State's data protection authority which acts as an independent body and has advisory competence only.<sup>16</sup> The meaning of their pronouncements can therefore be compared with those given by an administrative and not judicial body that determines what valid law is.

I want to point out that this thesis is not a comparison of different European jurisprudence. As a German lawyer I have focused the debate on German jurisprudence. For the reason of having a broader evaluation I decided to work with jurisprudence from other Member States as well. The debate will therefore incorporate Swedish and Norwegian jurisprudence as well.<sup>17</sup>

---

<sup>15</sup> The following presentation is based on Vogenauer, 2000, 400ff.

<sup>16</sup> Bygrave, 2002, p. 73

<sup>17</sup> I have chosen to incorporate Norwegian and Swedish legislation due to my Norwegian background that enables me to understand the written academic literature.

When discussing the legal classification of an IP-address I want to make clear that some statements are the opinion of the company Google represented by Peter Fleischer, Google's Global Privacy Counsel. Some statements will just represent Peter Fleischer's own opinion. I will differentiate between referring to Google as such and to Peter Fleischer as a private person expressing his own estimation. Jurisprudence from the ECJ is slightly touched upon. The ECJ has not been confronted to decide on the legal classification of an IP-address yet. It is important to have in mind that proponents of different opinions often do not differentiate between an Internet Service Provider and a Content Provider in a strict sense. If stakeholders involved talk about an Internet Service Provider they tend to incorporate the term Content Provider in that certain term. Peter Fleischer even incorporates an Access Provider in the term Internet Service Provider. It is my opinion that this creates an uncertainty in the debate as well. It is important to come up with clear definitions of the stakeholders involved which will be shown in the legal examination below.

## 1.5 Definitions

It is important to identify clear definitions that enable me to distinguish between the various stakeholders involved and between technical aspects which are essential in elaborating on this topic.

-Search engine: a computer program that retrieves documents or files or data from a database or from a computer network (especially from the Internet).<sup>18</sup>

-Internet Access Provider: enable the user –directly or indirectly- access to the computer network, especially access to the Internet.<sup>19</sup>

---

<sup>18</sup>At

<http://wordnetweb.princeton.edu/perl/webwn?s=search+engine&sub=Search+WordNet&o2=&o0=1&o7=&o5=&o1=1&o6=&o4=&o3=&h=>

<sup>19</sup>Hoeren & Sieber, 2008, Fn. 17

- Internet Service Provider: any natural or legal person providing an information society service.<sup>20</sup> Information society services span a wider range of economic activities which take place on-line; information society services are not solely restricted to services giving rise to on-line contracting but also, in so far as they represent an economic activity, extend to services which are not remunerated by those who receive them, such as those offering on-line information or commercial communication, or those providing tools for allowing for search, access and retrieval of data.<sup>21</sup>
- Content Provider: place and offer own information on the server of the host provider or their own server.<sup>22</sup>
- Host Provider: save external information on their computer systems (server) at the instigation of a user.<sup>23</sup>
- IP-address: An Internet Protocol (IP) address is a numerical identification and logical address that is assigned to devices participating in a computer network utilizing the Internet Protocol for communication between the nodes. Although IP addresses are stored as binary numbers, they are usually displayed in human readable notations, such as 208.77.188.166 (for IPv4) and 2001:db8:0:1234:0:567:1:1 (for IPv6).<sup>24</sup>

---

<sup>20</sup> Article 2(b) of Directive 2000/31

<sup>21</sup> Recital 18 of Directive 2000/31 for a clarification of the term "information society service".

<sup>22</sup> Hoeren & Sieber, op.cit.

<sup>23</sup> Hoeren & Sieber, op.cit.

<sup>24</sup> DoD Standard Internet Protocol, <http://tools.ietf.org/html/rfc760>

-Transmission Control: a transport layer protocol used by applications that require Protocol (TCP) delivery of data.<sup>25</sup>

-Controller: means the natural or legal person, public authority, agency, or any other body which alone or jointly with others determines the purposes and means of the processing of personal data.<sup>26</sup>

## 1.6 Structure of the thesis

Part 1 will analyse and discuss the legal classification of an IP-address as personal data in light of the Directive. A deep examination will be made of current jurisprudence and different opinions in academic literature concerning the status of an IP-address. In this context a deep analysis will be carried out of recital 26 of the Directive in interpreting the term “*all the means likely reasonable*” in the identification process and what the term requires and should encompass. In this context sub-issues will be touched upon on who should be the entity responsible in carrying out identification and if auxiliary material has to be taken into consideration in the identification process as well.

Part 2 of the thesis will elaborate if a search engine operator should be considered as a controller in light of the Directive. The basic function of Google’s operation will be dealt with: the allocation of the search engine as such.

In determining the term controller no account will be taken of additional instruments like the storing of personal data with the help of technical devices. The discussion will focus on the qualification of a search engine provider located with its headquarter outside the EU but operating with EU-based data centers.

Part 2 will further analyse what the term controller requires and what it should be composed of in order to be bound by European legislation.

---

<sup>25</sup> ”RFC Sourcebook”, <http://www.networksorcery.com/enp/protocol/tcp.htm>

<sup>26</sup> Article 2(d) Directive 95/58/EC

Part 2 is not going to elaborate on corporation law in analysing the status of an American Incorporation operating with European data-centers.

Part 1 and Part 2 will never lose sight in their legal examination of the ongoing debate between Google and the WP.

## **2 IP-address as personal data**

### **2.1 Problem statement**

The processing of personal data stresses concerns among privacy advocate groups and data subjects as well, regarding their possibility to exercise control and to be aware of personal information available about them on-line. How that information is being used is of special importance. To be able to determine if their concerns are justified and to analyse whether or not appropriate protective mechanisms are in place, a more general approach needs to be answered at first stake. What exactly is the IP-address about which the different stakeholders involved claim protection for? An IP-address is a complex asset and difficult to legally qualify due to its function.

An overview of these aspects will be given in the following before legally analysing an IP-address.

#### **2.1.1 Composition and character**

Computers have to rely on a specific communication standard like in any other communication mode. Those requirements of modality and the operation as such concerning the communication process between a computer and a computer network are determined through network-protocols, which result out of the protocol group called “Transmission Control Protocol/Internet Protocol (TCP/IP)”.<sup>27</sup> To transfer a data packet

---

<sup>27</sup> Ritterhoff & Neubert in Widmaier, 2006, Rn. 59

from a source host to its destination host, the Internet Protocol “wraps” the conditioned TCP-packages necessary for the data connection in IP-packages. The IP-addresses resumed from the TCP-package which contain the source- and destination host, are stored in an IP-Header.<sup>28</sup> The TCP/IP is therefore characterized through the following characteristics:<sup>29</sup>

- Information is split in transportable numbered data ready for transmission.
- Data packages are stamped with “*Headers*”, which contain the necessary information for transporting them, amongst other things the address of the source- and destination computer.
- Data packages are transported on different routes and different intermediate stations to the destination computer.
- The destination source checks the integrity of data packages and requests the source host on potential defective or during transport lost data packages.
- The destination source merges the received data packages in the right order and matches the original information together.

An Internet Protocol can therefore be compared with an envelope which consists of a dispatcher- and receiver address, comprising in its inside part its message, that is to say the TCP-package.<sup>30</sup> To ensure that the destination receives its data package, a computer which connects to the Internet is devoted an individual string of digits. This special string of digits is called the IP-address.

Every single computer has an individual and distinct IP-address. This gives each computer the above described possibility and function to act as a source- or destination computer.

An IP-address consists out of four different sections containing data ranging from 0 to 255.<sup>31</sup> An example for an IP-address is 62.156.153.38, which indicates the server belonging

---

<sup>28</sup> Hoeren & Sieber, op.cit., Fn. 52

<sup>29</sup> The combined list is based on Ritterhoff & Neubert, op.cit.

<sup>30</sup> Hoeren & Sieber, op.cit., Rn. 60

<sup>31</sup> Jähnel & Schramm, 2000, p. 67

to the official website of the German Police.<sup>32</sup> 8 bit are reserved for each data contained in the IP-address, meaning that an IP-address can reach up to 32 bit, labeling over 4 billion server with an individual IP-address.<sup>33</sup>

The thesis refers to the nowadays still overwhelmingly used protocol “IPv4”. Protocol “IPv6”, which will take over in future, will consist of IP-addresses covering 128 Bit. It has to be mentioned that the legal analysis and the outcome of it will not change through implementing protocol “IPv6”. Some technical changes will though occur. The goal of implementing “IPv6” is to create an additional global scope address based on a random interface identifier. Those temporal addresses will be used for a short period of time. After the usage the address will be deprecated, and further on used for already established connections.<sup>34</sup> The difference is that those deprecated addresses are not being used to initiate a new connection. This means that a constant flow of addresses can occur. This can eventually make it more difficult to track down a single user than it is today. The main advantage with implementing “IPv6” will be to attach more servers to the Internet, but the problem of an individuality assessment given by those assigned addresses will still remain. In being able to determine the legal qualification of the “IPv6” in future, it is essential to clarify the legal status of “IPv4” first, which can serve as a guideline for a further legal examination of “IPv6”.

### 2.1.2 Dynamic versus static IP-address

When talking about an IP-address it is essential to differentiate between dynamic and static IP addresses. Both of them fulfill their function as being considered as a server address.<sup>35</sup> A static IP address is a number that is assigned to a computer to be its permanent address on the Internet, assigned by an Internet Access Provider.<sup>36</sup>

---

<sup>32</sup> [www.polizei.de](http://www.polizei.de)

<sup>33</sup> Ritterhoff & Neubert, op.cit., Rn. 60

<sup>34</sup> “RFC 3041”, <http://tools.ietf.org/html/rfc3041>

<sup>35</sup> Holznagel, 2003, p. 221

<sup>36</sup> ‘static IP address/dynamic IP address’, <http://searchwinddevelopment.techtarget.com>

Due to the limited space of IP-addresses, a computer which is not permanently connected to the Internet will just get an individual and distinct IP-address necessary to communicate with computers. Those described IP-addresses are called dynamic IP-addresses. Such a dynamic IP-address is distributed to the computer from an Access-Provider out of a reserved contingent, composing an individual address just for the connected time on the Internet.<sup>37</sup> The next time the computer connects to the Internet a new IP-address will be assigned out of the Access Providers reserved pool. In my view it is important to differentiate between those two technical variations, even if the outcome of a legal interpretation might be the same.

There are statements in literature to be found which demonstrate that a different operational system exist when a computer connects to the Internet, but go further in proclaiming that this operational system is not relevant when analysing a legal classification. This view argues that it is just important-irrespective if a computer is devoted an IP-address permanently or not-that the address function is achieved.<sup>38</sup> I will argue in the following that it is essential to differentiate between these operational systems.

To emphasise my opinion a comparison can be made with telephone numbers. It would be inevitable to differ between a permanent telephone number belonging to an individual and an always new assigned number each time a caller dials a number.

In the majority of argued literature<sup>39</sup> a common conception prevails in considering a telephone number as personal data. The common notion prevails that a telephone number has to be considered as personal data because it is more than a momentary attribute carried with by the individual. It is my view that this common notion would advance a different view and approach if considering a telephone number which is just attributed for a single telephone conversation.

It is therefore essential not just to approach a technical, but also a legal examination when analysing an IP-address.

---

<sup>37</sup>Ritterhoff & Neubert, op. cit., Rn. 61

<sup>38</sup> Holznagel, op.cit., p. 219

<sup>39</sup> Norges Offentlige Utredninger, 1997, NOU 1997:19, p. 52

### 2.1.3 A Log file

Google operates with a log file belonging to a group called “non-reactive”-activity, opposed to a “reactive”-activity where the user takes active steps in providing personal data like signing in on an e-mail account. A log file consists inter alia of the IP-address. The log file does not contain information about the name of the user, thereby not enabling a direct identification if not combined with additional elements.

#### **Fig.2: Log File Entries in a Combined Log file Format<sup>40</sup>**

```
195.93.66.7 - - [27/May/2001:14:45:37 +0200] "GET /de/atlas-forum/index.html HTTP/1.0" 200
21871 "http://www.aeb.de/de/produkte/index.html" "Mozilla/4.0 (compatible; MSIE 5.0; AOL
5.0; Windows 98; DigExt)"
216.239.46.136 - - [27/May/2001:14:45:49 +0200] "GET /robots.txt HTTP/1.0" 200 171 "-"
"Googlebot/2.1 (+http://www.googlebot.com/bot.html)"
216.239.46.136 - - [27/May/2001:14:45:50 +0200] "GET
/en/events/bilder/AgendaSAP_Strategy_02_2001.pdf HTTP/1.0" 200 40955 "-" "Googlebot/2.1
(+http://www.googlebot.com/bot.html)"
```

It can be seen out of the log file<sup>41</sup> that the IP-address is registered on the left side followed by additional information like date and time. It is neither possible to conduct who exactly visited a certain web page nor which pages were read by a user. It is neither possible to ascertain how a user moved around on a specific page.

This opens the gate to elaborate if an IP-address can be considered as personal data in a legal context.

## 2.2 Legal analysis

### 2.2.1 Definition of “personal data” in light of Directive 95/46/EC

Article 1 (1) offers the scope of protection by presenting the aim of the Directive.

It states that *“in accordance with this Directive, Member States shall protect the fundamental rights and freedoms of natural persons and in particular their right to privacy with respect to the processing of personal data.”* Concerning the scope of the Directive

---

<sup>40</sup> Marcus Landau, 2001/2002, <http://www.ecommerce.wiwi.uni-frankfurt.de>

<sup>41</sup> Log Files are today still composed in the same composition, [www.e-teaching.org/didaktik/qualitaet/logfile](http://www.e-teaching.org/didaktik/qualitaet/logfile)

Article 2 (a) offers a definition on personal data stating that “*personal data shall mean any information relating to an identified or identifiable natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural, or social identity*”. Not the type of information is essential to determine personal information but the information which directly or indirectly can relate to an identifiable person.<sup>42</sup>

In interpreting this wording, information is a direct identification asset of a person if a name is combined with the age or address. Information is indirectly attached to a person with the help of distinctive marks of certain objects a person owns or possesses, like a car registration number or a telephone number. The following graphic will show how a distinction can be made.

**Fig.3:**

	Clear identification	Non-clear identification
Direct identification	A name combined with an address, picture and birth number	name
Indirect identification	Birth number	<i>IP-address</i> , car registration, job category

The graphic shows that there can be a situation where just a name is being possessed by an entity. Viewed in the abstract the name will not be a clear identifier, but it will reveal an identity if combined with additional elements like a birth number or address.

Important for the analysis is how, if just being in possession of an IP-address, a linking to additional or auxiliary information has to occur in order to qualify it as personal data.

---

<sup>42</sup> The following examination is based on Coll, 2000, p. 53-54

The Directive including its recitals do not offer more detailed information on how the term *identifiable* has to be construed. There are some issues which can be consulted and which can be seen as assistance in the identification process.<sup>43</sup>

- a) Are the information's easily accessible in order to identify a person?
- b) Should account be taken of legal / illegal means to carry out identification?
- c) Who is the legal relevant entity of identification?
- d) Is it possible to consult auxiliary material in the identification process?

Those issues will be of assistance when analysing the legal dispute between Google and the WP. There are intergradations between those four issues, especially between the first and second one. Point a) and b) will therefore be analysed together after having presented the ongoing debate. Point c) will be analysed subsequently due to its consequence out of the afore- mentioned examination. Issue d) will also be analysed in a separate sub item.

This means that those raised questions can be seen as a step ladder each giving an answer which has a consequence for the next one.

The examination will finally give an answer to the above presented graphic on how abstract or non-clear identification information can finally achieve the status of personal data.

In this context account must also be taken of recital 26 of the Directive.

Recital 26 highlights that “...*whereas, to determine whether a person is identifiable, account should be taken of all the means likely reasonable to be used either by the controller or by any other person to identify the said person...*”

An interpretation of the term *identifiable* can have a consequence for information which is solely an indirect identification, or where it is essential to gather auxiliary elements in order to achieve a direct identification. That information can fall outside the interpretation of the term. The question is therefore how far the term *identifiable* is and can be stretched.

---

<sup>43</sup> The issues are based on Bygrave, 2002, p. 42

This leads to the legal analysis of an IP-address and its criterion which determines it as personal data due to the fact that an IP-address as such is just an indirect, non clear identification asset.

The ECJ has not evaluated on the qualification of an IP-address in the sense of considering it as personal data yet. The closest decision which can be consulted here is the “Lindqvist-decision”.<sup>44</sup> The Court has taken the view that “*referring, on an internet page, to various persons and identifying them by name or by other means, for instance by giving their telephone number or information regarding their working conditions and hobbies, constitutes the processing of personal data within the meaning of Directive 95/46/EC*”. It appears that the Court takes the view that a telephone number has to be qualified as personal data.

This can have a consequence for the issue of an IP-address as well and therefore the question remains if an IP-address can pass the *identification* criterion test.

## 2.2.2 Working Party’s perspective

The WP states that “*an individual’s search history is personal data if the individual to which it relates, is identifiable. A search engine might not directly identify an IP-address, though identification may occur through a third person, like an Internet Access Provider holding IP-address data. In most cases therefore, including cases with a dynamic IP-address allocation, the data will be available to identify the user of the IP-address*”.<sup>45</sup> The WP further indicates that “*...unless the Internet Service Provider is in a position to distinguish with absolute certainty that the data correspond to users that cannot be identified, it will have to treat all IP information as personal data, to be on the safe side*”.<sup>46</sup> The WP highlights in document 148 that those considerations will apply equally to search engine operators.<sup>47</sup>

---

<sup>44</sup> Case-101/01 Lindqvist (2003) ECJ I – 12971, (27)

<sup>45</sup> WP document 148, op.cit., p. 8

<sup>46</sup> WP document 136, op.cit., p. 17

<sup>47</sup> WP document 148, op.cit.

They already addressed this approach in document 37<sup>48</sup> stating that “*Internet access providers and managers of local area networks can, using reasonable means, identify Internet users to whom they have attributed IP-addresses as they normally systematically log in a file the date, time, duration and dynamic IP-address given to the Internet user. The same can be said about Internet Service Providers that keep a logbook on the HTTP server. In these cases there is no doubt about the fact that one can talk about personal data in the sense of Article 2 a) of Directive 95/46/EC*”.

They continue in highlighting that there could be a situation where an IP-address will not be considered as personal data by stating, “*a particular case would be that of some sorts of IP-addresses which under certain circumstances indeed do not allow identification of the user, for various technical and organizational reasons*”, however in the following of the document making clear that “*...unless the Internet Service Provider is in a position to distinguish with absolute certainty that the data correspond to users that cannot be identified, it will have to treat all IP information as personal data, to be on the safe side*”.<sup>49</sup>

The question remains though what the term “*to be on the safe side*” should encompass. This will be analysed below where a legal approach is carried out.<sup>50</sup>

From the WP perspective there are practical, technical and legal assessments to be made in order to qualify an IP-address as personal data.

I will refer and analyse those assessments after having presented Google’s and Peter Fleischer’s response.

### 2.2.3 Google’s / Peter Fleischer’s perspective

In responding to the WP on the qualification of an IP-address as personal data, Google indeed points out that “*Google has always taken the view that IP-addresses should be regarded as confidential information that deserves a very high standard of protection*”.<sup>51</sup>

---

<sup>48</sup> WP document 37, p. 21

<sup>49</sup> WP document 136, op.cit.

<sup>50</sup> See 2.2.4.1 p. 21ff

<sup>51</sup> Google’s Response to the WP, op.cit., p. 5

This statement is however relativised when connecting an IP-address with Google's operational system, stating that *“indeed, Recital 26 of the Data Protection Directive states that determining whether a piece of data is personal information requires consideration of all the means likely reasonable to be used of identifying the individual. So sometimes IP-addresses should be considered personal data, for example in the hands of an Internet access provider that attributes those addresses to their own subscriber, whose personal details they hold, like name, address and billing address. On the other hand, IP-addresses should not automatically be considered personal data in the hands of any website that a user happens to visit, if that website has no ability to identify the user”*.<sup>52</sup>

In interpreting this statement, it can be concluded that Google itself is of the opinion that an IP-address has to be regarded in a special context and after having elaborated this context an assessment of the fact if personal data is involved has to occur. Google represents the view that the evaluation of an IP-address and its qualification as personal data is contextual based, highlighting that an Internet Access Provider has to be treated differently due to the circumstance that they have the ability to combine a certain IP-address with a specific user. This point of view is further supported by Peter Fleischer. He already addressed this topic in 2007.<sup>53</sup>

Fleischer highlights that *“considerations are of course contextual, based upon an assessment on a case-by-case basis of the likely chances that identification may occur in any reasonably foreseen set of circumstances”*, continuing in stating *“if a third party cannot receive assistance from an ISP in associating an IP-address with a particular user, the IP-address is not personal data as far as the third party is concerned. From the third party's perspective, the IP address is anonymous”*.

One year later Fleischer readdressed this topic<sup>54</sup>. He states that *“in order to get to that granular of a level, it would be necessary for Google to ask the ISP that issued the IP-address for the identity of the person that was using that IP-address (...), so again, on the balance of probabilities and taking into account any factors identified by the Working*

---

<sup>52</sup> Google's Response to WP, op.cit., p. 6

<sup>53</sup> Fleischer, 2007, <http://peterfleischer.blogspot.com/2007/02/are-ip-address-personal-data.html>

<sup>54</sup> Fleischer, 2008, <http://peterfleischer.blogspot.com/2008/02/can-website-identify-user-based-on-ip.html>

*Party as relevant, the most obvious conclusion is that the IP addresses obtained by Google and other websites are not sufficiently or revealing to qualify as personal data from the point of view of the EU data protection directive.”*

This opinion highlights that even if it might be possible to qualify an IP-address as personal data it can not qualify in being it in consideration of a Website Provider, due to the fact that the Website Provider has to get in touch with the Access Provider in order to get actual information on the user.

This summarised statement raises the issue highlighted in question c) above who should be the legal relevant entity of identification. The question will be addressed below.<sup>55</sup>

## 2.2.4 Evaluation concerning a dynamic IP-address

### 2.2.4.1 A legal approach / Analysing European legislation

It can be seen out of the above presented statements that numerous aspects are involved when analysing an IP-address. There exists a controversy between the stakeholders involved in how to verify the *identifiable* criterion set out in Article 2 (a) of the Directive. Some, like the WP and other stakeholders involved<sup>56</sup>, represent the opinion that a theoretical approach in being able to identify the user is sufficient, thereby evaluating Article 2 (a) in conjunction with recital 26 in a broad and literal way.

A theoretical approach means the possibility to acquire additional data even if prohibited by law. This means that even if a Website Provider himself can under no technical circumstances reveal a user's name, an IP-address should however always be considered as personal data.

To reveal the identity of that specific user an assistance of a third party is though essential and necessary.<sup>57</sup>

---

<sup>55</sup> P. 32 ff.

<sup>56</sup> Pahlen-Brandt, 2008

<sup>57</sup> Schaar, 2002, p. 37

The WP has positioned them by considering the means in how to *reasonably* identify someone. A cost of conducting identification should be considered as one factor, but also the advantage expected by the controller, the interest in stake and technical failures have to be considered.<sup>58</sup> The WP mentions that the test is a dynamic one, taking into account the special art in technology at the time of processing and the possibilities for development during the period for which the data will be processed.<sup>59</sup>

It can be indicated from this evaluation that by assessing the *likely-reasonable-test*, no consideration is taken of a lawful identification process. So what happens if this evaluation is taken for granted?

This objective test which takes into account key data, like a name or address stored in a database of a third party, would lead to a collective liability system of all data processors involved. The system does not consider what the different data processors involved exactly store. This leads to a situation that there just has to be one single data processor which could connect a data “X” to an individual with its stored data in a database, thereby treating the operator who just is in possession of data “X” as a processor of personal data.<sup>60</sup>

It is not necessary that the operator who is in possession of data “X” can by himself link this data with additional data.

The further consequence would be that an operator who just is in possession of data “X” is liable under the Directive, even if he is not aware of an entity possessing additional data referring to data “X”<sup>61</sup>. This leads to a situation in which the entity holding additional data can –deliberately or not- administer when the Directive and its scope of protection applies to the entity which again just holds data “X”.

Without an entity even being aware of an additional entity being in possession of further data which could allow a linking, the Directive opens its application as the case may be if additional data is held or not. This circumstance finally leads to a situation in a company

---

<sup>58</sup> WP Document 136, op.cit., p. 15

<sup>59</sup> WP Document 136, op.cit.

<sup>60</sup> Meyerdierks, 2009, p. 10

<sup>61</sup> Meyerdierks, op. cit.

holding data, even if a particular company holding the data cannot identify the individual to whom the data relates, finds itself regulated as a data controller.

This will even be the case if that company cannot or does not want to find out the identity of a person whose data they are in possession of.<sup>62</sup>

This in reverse has an impact on the data subject as well. Depending on whether an additional entity holds data - which in combination with data stored on him can reveal his identity- or not, decides if the data subject can challenge protection under the Directive.

In addition to this severe circumstance the operator who is in possession of data “X” is not able to determine the exact consequence on how to treat this data because he can not assess if the Directive is applicable or not.

There is some European jurisprudence to be found which has dealt with this theoretical approach.

The District Court of Munich<sup>63</sup> declared that IP addresses are non-personal data and speaks against the legal opinion of the Regional Court of Berlin<sup>64</sup>, the District Court of Darmstadt<sup>65</sup> and the WP.

In that case the plaintiff claimed injunctive relief against the defendant who operated an Internet portal. The defendant registered IP-addresses in certain log files, not just for the session in which the user active browsed that Internet portal but beyond that session.

The plaintiff argued that this operation method is illegitimate, revealing an Internet connection and thereby also revealing and allocating a certain user by storing an IP-address in a log file. This would constitute a violation of data protection law hence give reason for injunctive relief.

The defendant argued that a mere storing of an IP-address in a log file can not constitute a violation of data protection law.

Notwithstanding the claim being rejected due to a missing right to sue, the District Court went on in their justification and took a firm stand that the claim would have been rejected

---

<sup>62</sup> Tielemann for Covington and Burling, op.cit.

<sup>63</sup> Amtsgericht Muenchen: 133 C 5677/08, publication date 30.09.2008

<sup>64</sup> Landgericht Berlin: 23 S 3/07, publication date 06.09.2007

<sup>65</sup> Landgericht Darmstadt: 25 S 118/05, publication date 25.01.2006

out of a different reason as well. The Court went on in stating that a dynamic IP-address does not constitute personal data in light of data protection law.

The Court highlighted that an IP-address suffers from an intrinsic determinability.

Determinability is given if the data storing operator is able, without carrying out an unproportionate effort, to assess the person belonging to the stored particulars with their *normally available tools*. The Court further analysed those *normal available tools*.

They underlined that an IP-address is given to a single user by an Access Provider for a certain period of time. It is just the Access Provider who can determine, also after the user has finished his browsing session, the specific individual user. The defendant has to contact, in order to acquire additional data, the Access Provider. It is just with the help of the Access Provider that allows detection. Decisively is though that there exist no legal foundation, neither from the Access Provider to issue additional data nor from the Internet portal Operator to request additional data.

A final statement was made by the Court which contravenes the opinion of the WP.

The Court pointed out that a theoretical but illegal possibility in identifying a user through consulting an Access Provider, thereby forwarding this acquired data to the Internet portal Operator, is not consistent with the definition of personal data. Such an illegal approach can not be considered as a *normal available tool*, thereby not withstanding a proportionate assessment. The Court concluded that due to this given demonstration a dynamic IP-address can not constitute personal data in the hands of an Internet portal Operator who just stored the IP-address as such without further auxiliary material on a specific data subject.

The Court apparently distances itself from the opinion given by the WP.

In its judgement no legal explanation was given on the term *personal data* and what it should consist of in light of an IP-address.

The Court focused on the above presented theoretical possibility in acquiring additional or auxiliary data through an Access Provider.

Those thoughts expressed by the Court of Munich can be continued showing further problems the theoretical approach contains.

Even without the consultation of an Access Provider a Website Provider could acquire key data which is stored by an Access Provider. To acquire additional key data a Website

Provider would have to surmount an access protection, thereby making themselves accusable for spying out data<sup>66</sup>, which is punished under most European criminal law<sup>67</sup>.

The Amtsgericht<sup>68</sup> and Landgericht Berlin took another view on that issue.

Here the plaintiff claimed that the defendant had to refrain from storing personal data after each browsing session. The plaintiff claimed that by storing personal data the defendant had the opportunity in reconstructing which information he browsed and showed interest for. This ability could conclude political and religious opinions.

The defendant argued that a dynamic IP-address does in itself not constitute personal data. A storing of those IP-addresses is essential due to security reasons.

The Landgericht which decided in an appeal case followed the approach taken by the Amtsgericht Berlin at first instance. The Court basically ruled that a dynamic IP-address constitutes personal data. It does not make a difference if a dynamic IP-address is stored by an Access Provider or by an Internet portal Operator.

In explaining the reason for their decision they drew a line to recital 26 of the Directive, highlighting that all means have to be considered which either are adopted by the operator responsible or through a third party to acquire additional data. Due to the fact that a consolidation of personal data with the help of a third party is possible without actuating too much effort, it will be possible to identify a user in most cases. The Court continued that due to the ease in consolidating personal data, thereby denying a dynamic IP-address a status of personal data would mean that data protection law would not apply. This leads to a circumstance that personal data can be transferred without restriction to third parties.

The Court stated that a view which classifies determinability just in light of identifying a person with legal methods is not consistent with core data protection law. Here a line was drawn to the function of data protection law, thereby stating that the meaning and aim of it is exactly to protect data misuse, showing that such a modification of the term *determinability* is not justified.

---

<sup>66</sup> Meyerdierks, op.cit, p. 8

<sup>67</sup> See eg § 202 a German Criminal Code

<sup>68</sup> Amtsgericht Berlin-Mitte: 5 C 314/06, publication date 27.03.2007

This opinion takes an opposite approach in determining the classification of a dynamic IP-address. Here the view is advanced that by eliminating illegal means out of the determinability process would not be consistent with core data protection means and aims. The requirement of determinability does not only encompass legal means but also possible illegal handovers, as one of the aims of data protection law is to protect the individual from abusive use of data.

Finally the Landgericht Darmstadt elaborated on the storing of a dynamic IP-address. The difference in this case was that the defendant did not just offer online services but also access to the Internet as such. The focus rested on the criterion in storing dynamic IP-addresses for billing reasons. The Court stated that due to German law<sup>69</sup> there is no legal foundation for storing an IP-address in general. A storing is just permitted for detecting and identifying disturbances. This means that those arrangements are created for an *occurrence procedure*, which does not enable a storing of traffic data in general. Although the Court did not explicitly refer to the term personal data while elaborating on a dynamic IP-address, it can be seen that an IP-address has to be considered as an asset which belongs to the sphere of a user, thereby not allowing a randomly storing of it.

The legal uncertainty concerning the qualification of dynamic IP addresses carries on as can be seen from German jurisprudence for now. It is remarkable that all the statements concerning a dynamic IP-address solely focus on the legitimate grounds in processing personal data.

The same approach can be detected elsewhere in European jurisprudence.

In 2005 Stockholm's Lænsrätt<sup>70</sup> had to decide on the issue if an IP-address constitutes personal data. The Court took the same view as the Amtsgericht and Landgericht Berlin, stating that the essential and sufficient reason for regarding an IP-address as personal data is the ability of the Internet Access Provider in enabling identification with use of the IP-

---

<sup>69</sup> 'German telecommunication law', §100 I,III TKG

<sup>70</sup> 'Lænsrätten i Stockholms Län', reference number 593-2005, publication date 08.06.2005

address in connection with their stored data on the individual user. The Court addressed a circumstance which has not been touched upon by German jurisprudence.

While elaborating on the status of an IP-address they highlighted that it does not matter that the actual user of the computer necessarily has to be the physical person who signed a contract with the Access Provider. It is sufficient that the person signing up for a contract with the Access Provider can be identified. This contractual relationship allows a classification of an IP-address as being considered as personal data.

After having presented the different approaches, the thesis can not give a “black or white” answer for reasons which I will refer to later.

Personally I favour the approach taken inter alia by the District Court of Munich.

If considering illegal means in acquiring additional data, denotes to assume all stakeholders involved a tendency in breaking the law.<sup>71</sup> In addition it seems unreasonable to embrace an illegal approach in the identification process.<sup>72</sup>

#### 2.2.4.2 Interpretation of recital 26 of the Directive

The presented jurisdiction did not go into further details on the interpretation of the term *likely reasonable* contained in recital 26 of the Directive in achieving identification while elaborating on the classification of an IP-address.

As shown above, the attainment of additional data by consulting an Access Provider means fulfilling an element of crime. Doubts are appropriate here if the term wants to encompass an element of crime, which I personally challenge. A further evaluation and analysis of the term *likely reasonable* is therefore necessary.

It comprises inter alia terms like time, recourses and costs<sup>73</sup>, thereby assessing it in relation to a proportionality examination. I will refer to this in detail below.

---

<sup>71</sup> Meyerdierks, op.cit., p. 5

<sup>72</sup> See point 2.2.4.2

<sup>73</sup> Bygrave, op.cit, p .43

Like shown above<sup>74</sup>, an IP-address in form of a numeric combination assist in the addressing of a computer and does not constitute as such a direct identification of a natural person. This is the reason why the criterion of a *determinability* test plays an important role too. In considering the determinability assessment, the ability of the operator who stores certain data should be decisive.<sup>75</sup>

There are technical circumstances to be taken into consideration as well which the advocates of regarding a dynamic IP-address as personal data do not put any weight on. If an IP-address has determined the computer nothing is being said about the individual using that certain computer. There exists furthermore no clear indication that the IP-address can determine the exact used computer<sup>76</sup>, especially if a technical device called “IP-Spoofing” is inserted, which has the ability in counterfeiting the assigned IP-address<sup>77</sup>. Those systems are additional barriers in determining the individual user.

The only promising criteria to identify the user, viewed from the perspective of a Website Provider, is the log file containing inter alia the IP-address. In this context it is important to separate between “stock” and “user” data. The IP-address is just distributed by the Access Provider for a certain period of time which enables him to identify the user with the help of “stock” data. The “stock” data is transferred by the data subject to the Access Provider in the contractual signing period. The Content Provider never gets in touch with the “stock” data unless further technical devices like “cookies” are operated with, which are not taken into consideration here.

This again leads to the interpretation and scope of the term *likely reasonable* in recital 26 in order to identify a data subject.

Article 2 (a) of the Directive indicate different possibilities in achieving identification, inter alia by reference to an identification number or to one or more factors specific to his physical identity. In this context a line must be drawn to the *Explanatory Report of the Council of Europe Data Protection Convention* which points out that an identifiable person

---

<sup>74</sup> P. 8 above

<sup>75</sup> Koecher, 2007, p. 800

<sup>76</sup> Hoeren & Siebert, op.cit, Rn. 58

<sup>77</sup> Tanase “IP Spoofing-an Introduction”, <http://www.securityfocus.com/infocus/1674>

is a person “*who can easily be identified; it does not cover identification of persons by means of a very sophisticated methods*”.<sup>78</sup> It is important to notice that this institution does not constitute an instrument providing an authoritative interpretation of Conventions, although it may serve to facilitate the applications contained therein.<sup>79</sup>

The Convention does not elaborate on the term *sophisticated*. There are although several elaborations to be found which highlight that the identification process must inter alia consist out of an appropriate time, cost and manpower assessment.<sup>80</sup> This again means that information can not be considered as being able to identify a data subject if it requires an inappropriate effort of time, cost and work.<sup>81</sup> It is important to notice that this assessment is not a static one and can change throughout time due to the fact that technology always expands its possibilities. This indicates that the term *likely reasonable* always has to undergo an interpretation considering the actual art of technology.

The next step is to analyse how far the term *likely reasonable* should be stretched. In several occasions information will more be akin to anonymised information.<sup>82</sup> I will refer to this later. It can though be possible to achieve identification out of anonymised information as well. But again the question is what kind of effort has to be taken into consideration in order to attain identification. Here it is important to indicate a probability assessment that an identification can and will occur. It is my opinion that the term is unduly overstretched if illegal means are taken into consideration in the identification process. It is further my opinion that the term *likely reasonable*, as shown above<sup>83</sup>, consist inter alia of an element of time, cost and manpower but also and this is important, of a legal assessment. To achieve a personal reference it is my opinion that account should just be taken of the concrete possibilities the Content Provider has.

---

<sup>78</sup> §28 of the CoE Convention’s Explanatory Report

<sup>79</sup> <http://conventions.coe.int/Treaty/EN/Reports/Html/196.htm>

<sup>80</sup> Bygrave, 2002, p. 43

<sup>81</sup> Coll, op.cit., p. 55

<sup>82</sup> Coll, op.cit., p. 57

<sup>83</sup> P. 30 above

Those means *likely reasonable* in achieving a link to the individual user must further be analysed in light of a *proportionality* test. This test is now well established as a general principle of Community Law. Article 5 of the European Community Treaty<sup>84</sup> contains the proportionality test, which provides that action by the Community shall not go beyond what is necessary to achieve the objects of the Treaty.<sup>85</sup> It is my opinion that the term *necessary* in Article 5 EC contains of the same degree of an assessment evaluation like the term *likely reasonable* in the Directive does, just formulated with the former in a negative and the latter in a positive delimitation.

This means that the term *likely reasonable* contains a proportionality assessment as well. The proportionality test consists of three elements, an appropriate-, necessary- and adequate element<sup>86</sup>. Each element has to be considered in the identification process. Concerning this matter the actual ascertainment remains that the Website Provider, who in contrast to the Access Provider does not assign a dynamic IP-address to the user, can in most cases not assemble a personal reference if not operating with a “cookie” device.<sup>87</sup> There is no general claim for disclosure of a dynamic IP-address in terms to whom a certain IP-address was assigned to on a certain day and time.

There is various legislation in place where it is explicitly regulated who and when it is possible to get information on a certain person. Those are law enforcement agencies and Courts<sup>88</sup>. It is regulated in addition that this information is protected by secrecy of telecommunication<sup>89</sup>, which means that an illegitimate dissemination of certain information which can reveal a personal reference is punishable by law<sup>90</sup>, thereby as an *argumentum e contrario* having in mind that while not specifically regulated, dissemination is unlawful.

---

<sup>84</sup> From now on referred to the EC

<sup>85</sup> Craig & de Burca, 2007, p. 544

<sup>86</sup> ‘proportionality principle’, <http://www.saarheim.de/Anmerkungen/verhaeltnismaessigkeit.htm>

<sup>87</sup> Hoeren & Sieber, op.cit, Rn. 54

<sup>88</sup> See eg. §100 German Code of Criminal Procedure

<sup>89</sup> See eg. § 206 German Criminal Code

<sup>90</sup> Hoeren & Sieber, op.cit.

Given this background knowledge it is necessary to examine the *proportionality* test contained in the term *likely reasonable* to identify a person.

It is my opinion that while considering the above mentioned facts an unlawful theoretical approach in getting in touch with personal data can not be read into an interpretation of Art.2 (a) in conjunction with recital 26 of the Directive.

Further on, an unlawful approach would not stand the *proportionality* test as well. An illegal approach in revealing personal data could never pass the necessary and adequate element test.

Proponents favouring that a dynamic IP-address should be considered as personal data due to the theoretical possibility in attaining additional data from a third party, misinterpret the legal term and scope of data protection law. By stating that there exist a possibility in attaining data, thereby enabling them in affiliating a personal reference, displace the term *personal*. Even if the elements of time, effort and manpower could pass the likely reasonable test due to the fact that just a consultation of a third party in possession of additional data is sufficient, would though overstretch the term if an illegal approach is taken into consideration as well.

In addition there is an inconsistency to be seen when taking account of another legal discussion. The WP's approach in classifying an IP-address as personal data contrasts with an approach adopted in the Safe Harbour Agreement and in some Member States.

A different approach was taken to key-coded clinical trial data. The Commission and Member States represent the view that a transfer of key-coded data from an EU-investigator in a clinical trial to a sponsoring US pharmaceutical company is not being treated as a transfer of personal data if the company in the US does not hold the key.<sup>91</sup>

Here the view is manifested that the EU-investigator, who is in possession of the key-coded data which enables him to identify the individual who "hides" behind the stored data, has to be treated as being in possession of personal data.

---

<sup>91</sup> "Frequently Asked Questions 14-Pharmaceutical and Medical Products", at <http://www.export.gov/safeharbor/FAQ14PharmaFINAL.htm>

Argumentum e contrario this means that the U.S. company who does not possess data which enables an identifying of a certain individual is not being treated as holding personal data due to the fact that this data is being considered as anonymised data.<sup>92</sup> Auxiliary material in possession of a third party, here the EU-investigator, is not being evaluated and not taken into consideration. It is my opinion that this conclusion has to be taken into consideration when a legal examination on how an IP-address should legally be determined is carried out.

Regulators have reached precisely the opposite conclusion on almost identical facts; a third party which is in possession of an IP-address is considered as holding personal data even if there is no possibility in identifying the person, thereby taking account of all the means available. This indicates that by considering data in possession of a stakeholder not being able to reveal the identity as such, has been classified as a stakeholder holding “anonymised” data. There are no means *likely reasonable* to identify the data subject.

It is therefore my opinion that a dynamic IP-address can not be evaluated in a different manner than it was done with respect to key-coded clinical data. The entity holding the key coded data is not in the possession to “de-anonymise” key-coded data without transferring the key-coded data back to the controller who is in possession of the data as such. It is only the controller who can “de-anonymise” key-coded data.

Even if the legal impact an European Directive possess compared to the legal impact the Safe Harbor Agreement contains of, the outcome of this evaluation can not be underestimated and should therefore be taken into consideration when elaborating on the classification of an IP-address.

The result out of this examination is from my point of view that the term *likely reasonable* has to be evaluated differently, adapted to each concrete situation which is being assessed, thereby always taking account of the proportionality maxim. The term should not unduly be overstretched which I personally think some stakeholders involved do in order to achieve an appliance of data protection law at any price.

---

<sup>92</sup> ”Frequently Asked Question 14”, op.cit.

### 2.2.4.3 Relevant agent of identification

After having elaborated on the possible and legal approach in the identification process, thereby assessing it in relation to recital 26, a further conclusion can be drawn on the aspect on who should be considered as the relevant entity of identification. There are two different approaches which have to be evaluated.

It could be possible to include all stakeholders involved operating in a business sector where the possibility of processing personal data exists. This means that all stakeholders which possibly can get in touch with information on a data subject are taken into consideration. The outcome of this analysis will depend on the interpretation undertaken on the definition on personal data and its recital 26.

It is possible to evaluate a broad interpretation like several stakeholders involved do.<sup>93</sup>

If such an assessment is approached it will be possible to assemble various stakeholders involved who possibly could process personal data. This again means that it is possible to gather and analyse a lot more information stored by various stakeholders, thereby paying full attention to several interests which are touched upon by data protection law.

This approach would also follow up with the situation which is visible in the market place today. The development of the information society is characterized by the introduction of new electronic services. These services are available through different types of technologies and services which create new possibilities for the user, governments and business to interact with each other. At the same time though, electronic services have large capacities and offer new possibilities for the processing of personal data.<sup>94</sup>

The aim of data protection law is inter alia to protect the individual and to give him an adequate insight in this enormous business sector, especially being aware of what exactly is happening to his personal data. The data subject will eventually get more insight the more stakeholders are being considered as a relevant entity in achieving identification.

But there is a drawback on this evaluation as well. The amount of information the data subject could get insight to could hinder one of the core intentions of the Directive.

---

<sup>93</sup> Pp 21 above

<sup>94</sup> Debussere, 2005, p. 70

The aim in protecting the individual in his fundamental rights and the establishment and functioning of the Internal Market set out in recital 3 of the Directive with its free flow of personal data from one Member State to another shall be desirable. It is impossible to achieve this goal if each data in question would have to undergo an evaluation process. A general duty of informing an individual, thereby interpreting the definition of personal data in a broad sense could also be a burden for the data subject, as long as he can not properly determine its proportions.<sup>95</sup>

After having elaborated on an interpretation of the term *likely reasonable* in depth, thereby evaluating several conditions which should be taken into consideration, it is my opinion that a certain line has to be drawn when assessing the different stakeholders involved. Some stakeholders do not have to put a lot of effort in combining certain information to a certain data subject, thereby enabling identification. Some stakeholders will never be able or it will require a lot of effort to combine a certain kind of information they possess with a certain individual.<sup>96</sup>

Recital 26 of the Directive does not give a clear answer on this evaluation. It states that: “...account should be taken of all the means likely reasonable to be used either by the controller or by any other person to identify the said person...”

The first evaluation which can be taken out of this recital is that the controller is not the only entity involved who is responsible for achieving identification.

The further question is though what the recital wants to indicate by referring to *any other person*. Again, should it encompass everyone involved in the business sector gathering information on a data subject?

After having elaborated on the legal qualification of an IP-address in the hands of a Content Provider the answer is from my point of view visible. It is my opinion that a line has to be drawn where an entity legally or illegally acquires or tries to acquire personal data.

This view is supported with the wording the Directive uses in its recital 26 stating “...the means likely reasonable to be used...”

---

<sup>95</sup> Coll, op.cit., p. 59

<sup>96</sup> Coll, op.cit.

Illegal methods in acquiring data should not be encompassed. A hacker who gets access into a system, thereby extracting personal data should not be taken into consideration as an entity responsible for the evaluation of information<sup>97</sup>.

A further interpretation of the term *means* should also take account of a situation where it is unlikely that identification will occur.<sup>98</sup>

A Content Operator like Google should therefore not be considered as the entity responsible in the identification process. Either due to the fact that an illegal handover would occur or due to the situation that is highly unlikely that an Access Provider would forward additional data stored in their database. This examination will of course differ if a Content Provider operates with “cookies”. Here personal data is indirectly collected due to the fact that the data subject does not actively forward data to the Content Provider. In this situation the Content Provider does not have to get in touch with a third party storing additional information on a certain user. This means that an illegal approach in acquiring additional data is not involved, therefore the scope of protection of the Directive will apply. It is my opinion that this interpretation is in light of the Directive, giving it a flexible appliance. This interpretation also takes account of the always expanding possibilities technology offers.

#### 2.2.4.4 Auxiliary elements contained in the identification process

In all academic literature and jurisprudence I was working with elaborating on this topic, there were hardly statements to be found on how auxiliary information should be included in the identification process. The term *identifiable* and its inclusion of several aspects like shown above, which have to be taken account of while analysing the identification process, would seem to open up for the use of auxiliary material. It would lead to a truncation of

---

<sup>97</sup> Schartum & Bygrave, 2004, p. 116

<sup>98</sup> This view is supported by the english version of its recital 26, when operating with the term “means likely reasonable *to be used*”, meaning that methods which are unlikely to achieve identification will not be taken into consideration.

data protection law if auxiliary material would not be taken into account. The thesis will not elaborate on the topic of what exactly should be encompassed by the term *auxiliary*. Important for the aim of the thesis is to analyse how such auxiliary material is being gathered and if this gathering allows for incorporation in the identification process. In a situation where it is unlikely that auxiliary information is being forwarded after request, no account should be taken of this data. This finally embraces the circle discussed throughout the thesis. The mere theoretical possibility in acquiring auxiliary material should not be encompassed of the term personal data. This would lead to a hypothetical and uncertain situation. It is important that data protection law takes account of several situations and evaluates each situation with its concrete facts. It is my opinion that theoretical approaches should not be taken account of in the evaluation process. In an argument given by the WP I found a reference on this which in my view again highlights the weakness of their argumentation.

In their document 136<sup>99</sup> the WP states that *“Internet Access Providers and managers of local area networks can, using reasonable means, identify Internet users to whom they have attributed IP-addresses as they normally systematically log in a file the date, time, duration and dynamic IP-address given to the Internet user. The same can be said about Internet Service Providers that keep a logbook on the HTTP server. In these cases there is no doubt about the fact that one can talk about personal data in the sense of Article 2 a) of the Directive...”*

It is important to notice that the Working Party refers to a logbook in which, as shown above<sup>100</sup>, several tokens are stored. By referring to a logbook, thereby concluding that a dynamic IP-address should be considered as personal data, they combine several token to an entity. This step is a conform approach when analysing the term personal data because auxiliary material in the hands of an entity should be taken into consideration when analysing if an identification of a data subject is possible. The important statement is though that they refer to data contained in the sphere of the certain operator.

---

<sup>99</sup> WP document 136, op.cit, p. 16

<sup>100</sup> Table on logbook, p. 3

It is though decisive to define a barrier where auxiliary material containing certain information should not be taken into consideration in the evaluation process.

The legal examination of auxiliary material and its result has to be put in line with the result achieved in considering the legal relevant entity in the identification process.

Auxiliary data should be taken into consideration if they are collected directly or indirectly by an operator, thereby achieving identification.

However, a theoretical or even an illegal theoretical assessment in acquiring auxiliary material from a third party overstretches the term *indirectly identified* contained in Article 2 (a) of the Directive. It is my opinion that the term *indirectly* has to be interpreted in a way that an operator collects data which the data subject has not actively forwarded to the operator, for instance with the help of technical devices. An illegal theoretical approach can though not be taken into consideration, thereby depriving this auxiliary material from being considered as personal data.

## 2.2.5 Evaluation concerning a static IP-address

### 2.2.5.1 A legal approach

As can be seen from the definition of a static IP-address<sup>101</sup> there exist a technical difference when operating with this certain device. In most cases institutions and companies who possess a durable network connection are assigned a static IP-address. There exists a difference in comparing a dynamic with a static IP-address in light of retracing a certain computer.

A dynamic IP-address is being distributed each time a computer connects to the Internet. This means that a retracing of a computer assigned with a dynamic IP-address is more difficult. In being able to determine a specific computer no consideration has to be made in elaborating on the exact time the computer was connected to the Internet due to the fact

---

<sup>101</sup> Introduction p. 8

that a static IP-address always remains the same. The possibility to ascertain a specific user is much easier operating with static IP-addresses.

Given this technical difference the question is though again, considered from the viewpoint of a Website Provider, if a static IP-address as such can be considered as personal data in light of the Directive. A fixed IP-address will always be able to identify the same specific computer on the Internet. The static IP-address, irrelevant if stored for a longer period than a dynamic IP-address, will still just consist out of a numeric combination. However, to get in touch with additional information the Content Provider is at first sight again dependent on the Access Provider in forwarding certain specific information. But a relevant distinction has to be made in determining such a reconstruction, thereby revealing the identity of the user. A static IP-address might, considered in the abstract not be regarded as being relevant in light of data protection law. As indicated above it is an indirect non-clear identification asset. A static IP-address will though be of high relevance to data protection law when operating with “cookies”, which are specifically addressed to a certain computer.<sup>102</sup>

Due to the fact that those static IP-addresses are assigned for a longer period of time there is a big possibility in setting up a profile of a data subject. There is no doubt that this treatment raises deep data protection issues.

A personal reference can though already be missing for external entities if the computer assigned with a static IP-address is being used by several and constantly substituting persons.<sup>103</sup> Still it is possible to advocate that even the possibility of a multiplicity of persons sharing a machine with the same address just registered under the name of one person is unlikely to disqualify that machine from being treated as personal data.<sup>104</sup>

In elaborating on the topic of several people operating under the same static IP-address it is also advocated that this situation is unlikely to qualify a static IP-address as personal data. This indicates that there exist a big tension in how to legally qualify a static IP-address as well. However, a conclusion could be reached if the use of a certain machine is dependent

---

<sup>102</sup> Hoeren & Sieber, op.cit., Rn. 32

<sup>103</sup> Hoeren & Sieber, op.cit., Rn. 3

<sup>104</sup> Bygrave, op.cit., p. 317

on a password for each user, thereby being able of linking click stream data to exactly this password to a certain period of time.<sup>105</sup> The criterion is though again that additional data has to be taken into account in determining if a static IP-address can be considered as personal data. It is my opinion that as long as the Website Provider has to consult a third party to finally be able to determine a specific user, a static IP-address should not be considered as personal data as such, thereby being aware of the representatives arguing in the opposite direction due to the fact that illegal means in conducting additional information is being considered as part of the term *likely reasonable* in recital 26 of the Directive.

Again a further analogy shall be made taking account of key-coded clinical data.<sup>106</sup> This data can even more be compared to the asset of a static IP-address than a dynamic one, due to the fact that this key coded clinical data is a permanent data stored. As already taken in consideration above,<sup>107</sup> a pharmaceutical company contracts with an investigator to conduct a clinical trial. The investigator who is a neutral person is the one who is responsible for key-coding data. The individual's name and other means which could enable identification is replaced with a unique numerical or alphanumeric code. This means that the company only receives the key-coded data. This indicates that all means which could enable the company to identify the individual are removed. To assure that the health of the trial subject is been taken sufficiently care of the company will provide the physician with a relevant code, thereby enabling the physician to conduct the necessary follow-up process.

A big debate arose between different national Data Protection Authorities whether the key-coded data sent to a pharmaceutical company should be treated as personal data. Some Data Protection Authorities stated that those key-coded data should necessarily be considered as personal data because they classified this data as a “reversibly anonymised”

---

<sup>105</sup> Bygrave, op.cit., p. 318

<sup>106</sup> The following analysis is based on “Covington and Burling”, op.cit

<sup>107</sup> See p. 32

one. Even if the pharmaceutical company is not in the possession to “reversibly anonymise” the key-coded data themselves, it should although be treated as personal data. Due to the fact that the investigator holds the data which enables him to reveal the identity of the data subject, personal data shall be involved. This approach is identically with the approach taken when analysing an IP-address. The theoretical possibility is being seen as sufficient.

The opposite view proclaims that there is no personal data involved when considering the possibilities of the pharmaceutical company. The key-coded data is considered as “anonymous”, therefore the Directive cannot apply.

Unlike the debate on how to legally classify an IP-address, the Commission and Member States positioned themselves in favour of the last view represented.

As already mentioned above<sup>108</sup>, a transfer of key-coded data from an EU-investigator to a US pharmaceutical company does not constitute a transfer of personal data. This means that the Safe Harbour Agreement is not applicable. Here the view is shared that the key-coded data does not reveal data belonging to a data subject.

By drawing a line to the discussion on how an IP-address should legally be treated, it is my opinion after having elaborated on key-coded clinical data, that a static IP-address can not be considered as personal data as well if a third party has to be conducted. This is in line with my legal result in analysing the legal classification of a dynamic IP-address.

Personally it is my opinion that key-coded clinical data “hides” even more sensitive data compared with a static IP-address. This again leads *a fortiori* in not classifying a static IP-address as fulfilling the criteria of personal data in light of a theoretical context.

#### 2.2.5.2 Auxiliary elements contained in the identification process

Auxiliary data has to be taken into consideration while interpreting the term *indirectly identifiable* contained in Article 2 (a) of the Directive. As indicated above<sup>109</sup>, Google refrains from considering IP-addresses in their hands as personal data, due to the fact that

---

<sup>108</sup> See p. 33

<sup>109</sup> See p. 20

they just consider it in the abstract without paying attention to an interpretation of Article 2 (a) and recital 26 of the Directive. It is my opinion that a possibility in attaining auxiliary data which finally enables identification has to be taken into account if assembled in a legal manner.

If viewed in the abstract and not paying attention to auxiliary material the view can be supported that a static IP-address has to be treated in the same way like key-coded clinic data is being treated, which is anonymised before forwarded to a third party.

There is just a theoretical possibility in revealing identification. It will be analysed if this view should be supported.

A fixed IP-address will always be able to identify the same specific computer on the Internet. Google does not pay attention to technical devices which exist when addressing this topic. A closer look will be taken on those technical devices and if they have to be taken into account in the identification process. An example is *ARIN's WHOIS* service provider, a mechanism for finding contact and registration information for resources registered with *ARIN*. *ARIN's* database contains IP addresses, autonomous system numbers, organizations or customers that are associated with these resources.<sup>110</sup>

Another example is *Whois Source*<sup>111</sup>, allowing for an identification of a specific computer.<sup>112</sup>

It is my opinion that the search outcome gained by those technical devices has to be treated as auxiliary information which is encompassed by the term *indirectly identifiable*.

Taken thereby account of *all the means likely reasonable* in the identification process, neither too much time, cost and manpower is spent in revealing identification, nor does an illegal handover of a third party occur. The question is though what exactly a data base like *ARIN* or *Whois Source* does reveal. Here a differentiation has to be made with a static IP-address used by a single user who has signed a contract with an Access Provider and with several users sharing a computer, though operating with different passwords in order to connect to the Internet. I will present the second mentioned method first using my static IP-

---

<sup>110</sup> <http://ws.arin.net/whois/>

<sup>111</sup> Cojocarasu on '*Whois Databases*' for a deeper examination, 2009

<sup>112</sup> <http://whois.domaintools.com/>

address as an example, which is devoted to the University of Oslo. The IP-address of my computer is 129.240.179.20.

If entering this device into the *Whois Source* the information will appear which I have made visible in the Annex Table 1. Most important is that the table indicates the IP location, resolve host, IP-address and the blacklist status. A further comment is made that this address has been further assigned to users in the RIPE NCC region. No indication or further information therein reveals information which could directly be connected to my person. The only information contained therein, allowing for a further assessment of my person, is the indication that this IP-address has been devoted to a computer located at the University of Oslo. This again leads to the fact that to acquire more information about my person a Content Provider has to get in touch with the person responsible at the University, who is mentioned in the *Whois Record* as the responsible host master, to finally attain direct identification. The responsible person could reveal the person “hiding” behind the password which is used in order to eventually get connected to the Internet with this certain IP-address. A forwarding of information would though again constitute illegal operations by a third party as analysed above.<sup>113</sup>

The second step is to analyse if a different result is achieved when entering a static IP-address belonging to a single user who has signed a contract with an Access Provider into the search bar (see table 2 in the Annex). Here I will use the static IP-address 81.217.55.109. By entering the IP-address of that specific user into the search bar, just the Access Provider is indicated who is responsible in distributing static IP-addresses to its customers. As can be seen out of table 2 it is the Austrian Access Provider Kabelsignal AG. In addition the host master responsible for Kabelsignal AG is listed in the index. No further indications are contained in the listing which could reveal the actual users identity.

This again means that a Content Provider like Google would have to consult Kabelsignal AG in order to get auxiliary information of the user to whom that IP-address has been distributed in order to achieve identification, which again would constitute an illegal attainment of auxiliary information.

---

<sup>113</sup> P. 28ff above

Given this examination it is my opinion that an IP-address should be treated as anonymised data as it was done with the key-coded clinical trial data presented above.

The Access Provider can be compared with the EU-investigator who is in possession of data which do enable identification. The Content Provider, even if using a technical device like a *Whois Record* will just be in possession of key-coded data in a broad sense, here the IP-address, like the US pharmaceutical company does.

### 2.2.5.3 Intermediate result

The Directive's context-based appreciation leaves room for a both narrow and broad understanding.

The same can be said about the difficulty to actually identify the person about whom information may only relate indirectly to. It is obvious that the consequence of this observation in light of data protection law is significant if one Member State treats certain information as personal data and another Member State not, leading to the fact that in one Member State data protection law will apply and in the other Member State not.<sup>114</sup>

In the context of this research it is my opinion that a dynamic and static IP-address can not be treated as personal data in light of the Directive if illegal means are being taken into consideration in the identification process. The evaluation will differ as long as an operator acquires data which he has lawfully attained himself by operating with a "cookie" device revealing auxiliary data. In such a circumstance it is necessary to interpret the term *all the means likely reasonable* contained in recital 26 of the Directive in a different legal manner due to its legal range of coverage.

The theoretical assessment valuation in acquiring auxiliary data from a third party does in contrast not pass the *likely reasonable* test, as it did not in the US clinical-trial case either. It is my perception that the Commission and Member States should adhere to their evaluation in order to achieve a stringent approach, thereby assisting in achieving a harmony throughout Europe.

---

<sup>114</sup> Reidenberg & Schwartz, op.cit., p. 125

It is important that the European legislator eventually comes up with a clarification if illegal means shall be incorporated in the identification process, which contains inter alia of elements covering the legal relevant entity responsible for identification and the evaluation of auxiliary data, due to the above described gap it creates which is obvious in the ongoing debate between Google and the WP.

### 3 Search engine as a “controller”

#### 3.1 Problem statement

This Chapter is going to analyse if Google Inc., although established in the U.S., is bound by European legislation. The discussion will focus on Article 4 (1.a) of the Directive which determines that *“each Member State shall apply the national provisions it adopts pursuant to this Directive to the processing of personal data where the processing is carried out in the context of the activities of an establishment of the controller on the territory of the Member State....”*

After having presented the debate a practical and legal examination will follow.

##### 3.1.1 Working Party’s perspective

The WP states *“that a particular processing operation of personal data should be taken as the starting point, though when applied to a particular search engine whose headquarter are located outside the EEA, the question needs to be answered whether the processing of user data involves establishments on the territory of a Member State.”*<sup>115</sup> The WP continues their demonstration in trying to define an establishment by stating that *“the existence of an establishment implies the effective and real exercise of activity through stable arrangements....”*

The WP defines the meaning of the term activity by indicating that an establishment should play a relevant role in the particular processing operation. This relevant role is important for the following legal examination because due to the WP’s opinion a role is inter alia clearly at hand when *“an establishment is responsible for relations with users of the search engine in a particular jurisdiction and a search engine provider establishes an office in a*

---

<sup>115</sup> The following extracts are taken from WP document 148, op.cit

*Member State that is involved in the selling of targeted advertisements to the inhabitants of that state...*”

The WP does not go into further details, thereby leaving an open legal gap on several juridical interpretations. It more or less seems like the WP refrains itself from interpreting Article 4 (1.a). This gets obvious when the WP at the end of its explanation states that *“it is the search engine service provider that is responsible for clarifying the degree of involvement of establishments on the territory of Member States when processing personal data.”* Terms like *establishment* and *activity* are merely touched upon without giving guidelines on their dimension. No further guidelines are given on relevant terms like *relations with users* or *selling of targeted advertisement* either.

Does the term involve a manual operation or is a mere automatic process being considered as sufficient? Does it indicate that an active step in terms of getting involved with a user has to occur or is a mere automatic algorithmic process sufficient in detecting and further on selling targeted advertisement? To come up with an answer on that question a legal interpretation of the term controller will be dealt with in the following.

### 3.1.2 Google’s perspective

Google indicates that *“Google as a search engine is not a content provider in a legal sense of that term; they do not publish or republish content, but provides access to information created and published by others.”*<sup>116</sup>

They address the term of a controller by stating that *“Google does not carry out any further processing operation in terms of manipulating the information for presentation in any particular way other than as a list of search results; in other words Google will be merely operating automatically on the instructions of a user...therefore Google cannot be regarded as the controller of any personal data included in the search result.”*

There are further statements found on Google’s webpage “Help for web search”.<sup>117</sup>

---

<sup>116</sup> The following extracts are taken from “Google’s Response to WP”, op.cit.

<sup>117</sup> ”Google’s hjelp for nettsøk”,

[http://www.google.com/support/websearch/bin/static.py?page=faq.html&hl=no#anchor\\_link\\_18](http://www.google.com/support/websearch/bin/static.py?page=faq.html&hl=no#anchor_link_18)

Those following statements will be of further importance that is why I choose to summarize them. Google is supposed to be a completely automated search engine which uses a program called “spider”. This program “crawls” through the Internet in periodical intervals, thereby following hyperlinks from page to page in order to detect web pages which can be put to the index. Google points out that no manually approach is adopted in this operation process. The actual updating process is fully automated due to the fact that vast amount of information is gathered in the index. Google further points out that the result of an index page can change because a certain webpage can attain public attention which raises the importance of that webpage. The ranging program, which classifies the “Page Rank”-value through clicks, determines its deposition on a certain index page. Their main argument is that due to merely operating automatically they more or less act as a processor of information carried out by a set of certain criteria, foremost the acting behaviour of a user.

### 3.1.3 Statement on the issue

Before analysing the term of a controller and the requirements it should fulfill, some technical background is essential for being able to elaborate on that topic in a legal sense. Google has repeatedly stated that all the ranking and indexing process occurs by mere automatically means. So what exactly is meant by a “ranking” process?<sup>118</sup>

Due to the enormous dimension the “WWW” offers, it is vital to be in possession of a good ranking process in order to attain and maintain users when operating a search engine.

It is not merely sufficient to be able to screen relevant pages from not relevant pages.

The sequence is decisive in presenting the hits. Google is successful because they are aware of the fact that just if the first dozen hits contain adjuvant search results the search engine will be of value. Before it is possible to operate with a ranking system it is decisive to conduct which pages as such shall be considered as a hit. Google operates with a certain

---

<sup>118</sup> The following technical explanations are based on Schoech, 2001,  
[http://inf.fuberlin.de/lehre/WS01/netbasedIS/uebungen/PageRank\\_vortrag\\_schoech.pdf](http://inf.fuberlin.de/lehre/WS01/netbasedIS/uebungen/PageRank_vortrag_schoech.pdf)

benchmark which enables the system a priori to consider on how exactly a hit matches with a certain search request. Google operates with a benchmark that just considers web pages which contain all search terms. Pages containing the search terms in the same order will be stronger evaluated. Pages containing search terms in a geographical nearness will be stronger evaluated too. The link-structure of the “WWW” can be treated as a graph which contains valuable and objective information’s of the importance of a single webpage. With the help of the “Page-rank” system a number will be calculated out of that information. This number will eventually be used for the sorting of a search result. For each page contained in the database a global importance account will be assessed out of the structure of their reference.

So the “Page-rank” system is based on following assumption:

- the more links which refer to a special page indicate the importance of each page.
- the less links a page contains the more important is each link.
- the more important a page is, the more important are the links contained on it.
- the more important the link which refers to another page, the more important is the referred page.

This assessment might indicate that the “page-ranking” system just works on a calculated algorithmic program. This probably would lead to a circumstance in which Google could not be assessed as a controller. The Directive does not give a clear definition on how the term controller has to be interpreted. Article 2 (d) in conjunction with recital 19 of the Directive has to be taken into consideration.

Article 2 (d) states that a *“controller shall mean the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purpose and means of the processing of personal data.”*

The term controller always has to be put in relation to where the establishment resides. Recital 19 states *“whereas establishment on the territory of a Member State implies the effective and real exercise of activity through stable arrangements...”*

An interpretation of the terms “*determining the purpose and means*” and “*real exercise of activity*” indicate that a mere passive operation is not sufficient in being considered as a “*controller established*” in a Member State.

The Working Party already positioned themselves in 2002 when a closer look was taken of the term controller.<sup>119</sup> They stated that “*it seems necessary that the processing takes place in the course of an activity, which falls within the scope of Community law and thus under the directive.*” The term *activity* is not being defined in detail. It is my opinion that it would construe a too broad evaluation of the term if the mere composition of a program by a human being which in the following is being applied would constitute an *activity*. This again could speak in favour of Google’s response in that a mere automatic operating system deprives Google of being considered as a controller of any personal data included in the search results.

This furthermore leads to the analysis if the “indexing” and “page-ranking” operation is merely operated automatically.

### 3.1.3.1 A practical approach

To anticipate the result of my research first, I determined that Google’s claim does not really comply with their operation system in a practical sense. Instead of just operating with an algorithmic system there exist hints that humans make several adjustments in the sense of collecting data. After having made those adjustments a further step of action is being taken in how to present that extracted data. During the indexing period search engines use and display third party description of a website in the search results. This can consist of personal data. This could be a description of a name or an address belonging to a data subject. When those pages are indexed, Google might decide to exclude web pages from their index for reasons ranging from violations of quasi-objective search engine technical requirements to simple capriciousness.<sup>120</sup> This indicates that the indexing operation is being handled with human beings involved.

---

<sup>119</sup> WP document WP 56, p. 6

<sup>120</sup> Goldman, 2006, 188ff

The same can be conducted in the ranking process. Google indicates in their response that the operating system is based on an automatic process. This means that the ranking algorithm do not require humans to make those ranking decisions, thereby denying their status as a controller. This conclusion is from my point of view not right. Instead the choice of which factors are included in the ranking algorithm reflects the search engine operator's editorial judgement about what exactly makes an adjustment valuable.<sup>121</sup> There are even indications that Google manually assess and determine which ranking position a certain web page should possess. The Internet coach Hank von Ess came up with an analysis which is not conform to the statements given by Google in not modifying algorithmically-generated search results. In his evaluation it is manifest that Google operates with humans in the ranking process. Google Inc. even recruits testers in advertising:

*“Quality Rater-(Spanish, Dutch, Italian, French) this is a temporary role offered through Kelly Services. Google Inc. is recruiting part time, temporary, home-based workers to help with work on a search quality evaluation on a project basis. Candidates will evaluate search results and rate their relevance. All candidates must be web savvy and analytic, have excellent web search skills and a broad range of interest. Legal eligibility to work in the Netherlands, Italy, Spain or France. The job involves frequent written communication with fellow Quality Raters.”*<sup>122</sup>

It is obvious that this notice does not indicate the real operating system; still it is an indication which at least creates confusion on how to legally qualify Google.

It seems that Google regards themselves as a neutral operator due to their application of the technology mentioned. These manual interventions can be an exception to their otherwise used automatic process operation. Still indications reinforce that Google play an active role in forming their indexing operation to accomplish their economic and editorial goal.<sup>123</sup> This would justify a treatment of Google's European data-centers as controllers, due to their

---

<sup>121</sup> Goldmann, op.cit.

<sup>122</sup> Eric van Ess, "Google Secret Lab, Prelude", <http://www.searchbistro.com/index.php?/archives/19-Google-Secret-Lab,-Prelude.html>

<sup>123</sup> Eric Goldman, op.cit, p. 192

possibility in controlling, scanning and placing of, in their view “important”, web pages on a higher index position.

### 3.1.3.2 A legal approach

I want to conduct an additional legal approach due to the just described possibility approach. As designed above search engines browse the Internet with the help of “crawlers”, thereby attaching an index of the captured web pages. After this process has been carried out the user is offered, after having searched on a certain term, via a link access to a certain web page. It is not manifest in the Directive if such an operational process is covered by the term controller in light of an active operation method involved. There has been some Norwegian jurisprudence dealing with a manual or mere automatic storing of data. The Data Inspectorate allowed a video surveillance in the entrance areal of a train but not in the train cabin. The reason for this ruling was that the recording of personal data and the replay of it would just occur if a criminal episode had occurred. If there is no suspiciousness of a criminal act no human being will take a look of the recording.<sup>124</sup>

In another case the Data Inspectorate ruled in favor of permitting the video surveillance recording on a bus. Again they stated that a human being would just record the video if a criminal episode had occurred on the bus.<sup>125</sup> This shows that a mere automated stored operation without further human insight is permitted.

The difference in respect to the operation process of Google is that a third party is involved. The question has to be asked if this ruling can be transferred on Google’s operation process. From my point of view Google’s operation method would more refer to an “intermediation” process. Therefore the argument is represented that not the search engine operator but the user who has attached information to the Internet is the controller.<sup>126</sup> It is my opinion that this is just an objective assessment not reflecting the business operation

---

<sup>124</sup> “Personvernemnda-Vedtak PVN-2005-13”, <http://www.personvernemnda.no/vedtak/index.htm>

<sup>125</sup> “Personvernemnda-Vedtak PVN-2005-13”, <http://www.personvernemnda.no/vedtak/index.htm>

<sup>126</sup> Ott, 2009, p. 158ff

involved. The question still has to be asked why Google should refrain from not interacting in the indexing operation scheme. The results of a search are of vital importance viewed from a user perspective and from Google itself. Taken Google's perspective it is vital on how each single user perceives the outcome of a single search. This is the reason why search engines manipulate their ranking algorithms in order to satisfy the majority interests.<sup>127</sup>

Besides those just indicated problems there is a possibility that a search engine provider carries out further intern manipulation in order to raise their revenue through online advertising. Due to the fact that search engine operators earn their money predominantly with online advertisement<sup>128</sup>, they have a big interest in companies taking out advertisements on their web pages. A company needs an advertisement if it is not represented high enough on the index page. There is an endangerment that a search engine operator consciously degrades certain popular web page hits in order to agitate a company to buy key word hits for being able to advertise with a high index position.<sup>129</sup> This means that the search engine operator has the possibility in evaluating the importance of certain web pages, thereby enabling them to remove or to displace web pages which contain personal data. It is difficult to determine such behaviour due to the fact that there are a number of different criteria's in "sliding down" on the index list. Those criteria's are difficult to separate from intern manipulation handled by a search engine provider. This is even more the case if a search engine operator justifies their operation method with legal criteria's.

Google justifies inter alia their storing of a data log with the following remark: *"Log data is essential to prevent and investigate threats to our users. Data from search logs is also one tool we use to fight web spam and return cleaner and more relevant results"*.<sup>130</sup>

They further continue to explain their operation by stating that *"web spam is the junk you see in search results when websites successfully cheat their way into higher positions in*

---

<sup>127</sup> Introna & Nissenbaum, 2000, p. 169

<sup>128</sup> In 2007 Google generated 99 % of their total volume through online advertisement

<sup>129</sup> Kuehling & Gauss, 2007, p. 881

<sup>130</sup> Google's Response to the WP, op.cit.

*search results or otherwise violate search engine quality guidelines...all of this log data - IP addresses and cookie information- makes your search results cleaner and more relevant... ”*

It must be detained until now that all considerations still just indicate a probability assessment which still is vague in legally regarding a search engine operator as a controller. Recital 47 states that “*whereas where a message containing personal data is transmitted by means of a telecommunications or electronic mail service, the sole purpose of which is the transmission of such messages, the controller in respect of the personal data contained in the message will normally be considered to be the person from whom the message originates, rather than the person offering the transmission service... ”*

Recital 47 has to be read in conjunction with Article 2 (d) stating “*controller shall mean...determines the purpose and means of the processing of personal data... ”*

This indicates that the definition exists of a term *by means* and of a term *sole purpose*.

It must be analysed if a search engine operator can codetermine those terms in a cumulative way.

Viewed from an objective perspective, Google does not allocate determinations on a specific purpose. They appear in a neutral position concerning those stored information’s. They can neither determine the search terms the individual user enters in the search tool bar nor on which specific web pages personal data is being stored.<sup>131</sup> It is my opinion that search engine operators do not exclude, more or less are aware of the purpose that personal data is being searched for.

This again leads to a further analysis in how to cope with the aim of a requested *active* step. A mere sufferance of illegal actions committed by a third party cannot in itself constitute an active action. A further *corpus delicti* has to be fulfilled. A mere sufferance of an action carried out by a third party can construe an active deed if the person is legally obliged to act. This is the case if a person or entity possesses an affirmative obligation to act in order to protect others. It is my legal perception that Google and their European data-centers are creating an origin of danger in light of the possibility of personal data being processed.

---

<sup>131</sup> The following discussion is based on Ott, op.cit.

This legal specification of an affirmative obligation to act assigns Google the duty to take active measures in preventing a breach of law.

This interpretation leads to the legal assessment that Google determines the *purpose*.

It is in dispute if Google also determines the by law demanded *means of the processing of personal data*.

Data Protection Law is not construed in a sense of a three-person-proportion. In the search engine sector there are several stakeholders involved. A distinction has to be made in terms of the Webmaster- the Search Engine- and a third person.<sup>132</sup>

The crucial point is that web pages captured by the search engine are not transferred directly to the third person, so that the search engine operator is not aware of the fact if the third person actually receives the desired data. It is difficult to construe a situation which can be compared with the position of an Access-Provider because the Access Provider can control the forwarding of personal data. If viewed from an objective point of view a search engine is therefore more akin to an “intermediator”. They cannot control the means of data which are stored on a linked webpage. The “intermediator” role is limited to *where* a third person is able to find requested information but not exactly *what* can be obtained on that page.

There are other technical operations in place which could lead to another interpretation of those required terms. “Cache”, “Snippets” or “Thumbnails” are technical means by which personal data are shown directly on the index. Here an interpretation has to deal with the same legal terms which are though not part of the thesis.

In this analysed context it is my opinion that the Directive is not applicable if a mere “intermediator” due to their neutrality is involved. It is the webmaster that finally transmits information stored on that page after a request has been carried out via a search engine.

Due to the missing *means of processing personal data* and the complexity of determining practical internal operations it is my opinion that Google cannot be regarded as a controller in light of Article 4 (1a) of the Directive, thereby not including those technical devices mentioned above which could lead to another evaluation.

---

<sup>132</sup> Ott, op.cit.

## 4 Conclusion

The thesis aimed to analyse the classification of a static and dynamic IP-address as personal data under the Directive. Instead of choosing to work with additional technical devices like a “cookie” inserted on the user’s computer my aim was to analyse the IP-address in the abstract. This required to look on how an IP-address as such consists of and how it is technically inserted in the transmission context.

Therefore Chapter 2 started by describing the operation and transmission process on the Internet as such. It was shown how the “Transmission Control Protocol/Internet Protocol” is characterised, passing from there on into the composition of an IP-address and its functioning within the “TCP/IP” operational system.

It subsequently described what a log file consists of, thereby indicating the storing of the IP-address in a log file.

Further on the thesis described the crucial issue of personal data being involved when a dynamic or static IP-address is stored in a log file, thereby never losing sight of the ongoing debate between the WP and Google. An in depth analysis was carried out with European jurisprudence and academic literature. It was shown that it is not possible due to different interpretation possibilities of the term *likely reasonable* to come up with a “black or white” answer. A personal approach was though undertaken.

A further analysis was undertaken if auxiliary material should be encompassed of the term *likely reasonable* and if illegal obtained data should be taken into consideration in the evaluation process as well.

In addition it was analysed who should be the relevant entity in the identification process. Chapter 2 highlighted in its final part that the Directive’s context based appreciation leaves room for a both narrow and broad understanding, the result being a big gap throughout the Member States.

Chapter 3 analysed if a search engine operator can be treated as a controller under the Directive, thereby focusing on the terms *means* and *purpose*. Most importantly this Chapter

identifies differences in approaching from a practical, economic –more speculative-, and legal analysis.

I do believe that search engines operate in a grey area. Due to the different technical possibilities the legal assessment can change and vary, making it again difficult to give an answer in either direction.

Nowadays a situation has come to hand in which the legislation, which I believe is a strong mechanism, and technology do not operate hand in hand. The reason for this development is not the technology as such. It is up to the human being while working with a conformity approach to realize that the huge technical possibilities have to be brought in line with legislation. Still legislation just forms one part. A self-regulatory and co-regulated area could also be a trigger in approaching the big complexity the topic offers.

Further kids in school must be reached and be aware of all the possibilities which occur on the Internet. This generation grows up with the Internet as an everyday tool. It is important that while it is a useful tool in their hands, they in addition understand the danger it bears. I further proclaim that an autonomous control body should be in place which deals with the accusation of manipulating the operating system in the search engine sector.

This would not just be a helpful tool for the user but also for the search engine operator in “discharging” themselves from accusations.

## **Selective bibliography**

### **A Statutes and legislation referred to in this thesis**

“Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 *on the protection of individuals with regard to the processing of personal data and on the free movement of such data*”, Official Journal L 281, 23.11.1995 P. 0031-0050.

“Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 *on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market*”, Official Journal L.178 of 17.06.2000.

“Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 *on a common regulatory framework for electronic communications networks and services (Framework Directive)*”, Official Journal L 108/33, 24.02.2002.

“Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 *concerning the processing of personal data and the protection of privacy in the electronic communications sector*”, Official Journal L 201, 31.07.2002 P. 0037-0047.

“Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 *on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications network and amending Directive 2002/58/EC*”, Official Journal of the European Union (105).

European Court of Justice: ‘Case-101/01 Lindqvist (2003)’, ECR I – 12971.

German Law (2004) “*Das Telekommunikationsgesetz (TKG)*”, available at <http://ladisch.de/dr/tkg.html>.

German Law (1987) “*Die Strafprozessordnung (StPO)*”, available at [www.gesetze-im-internet.de/bundesrecht/stpo/gesamt.pdf](http://www.gesetze-im-internet.de/bundesrecht/stpo/gesamt.pdf).

German Law (2009) “*Das Strafgesetzbuch (StGB)*”, available at [www.gesetze-im-internet.de/bundesrecht/stpo/gesamt.pdf](http://www.gesetze-im-internet.de/bundesrecht/stpo/gesamt.pdf).

Norwegian Official Report (1997): ‘*Norges Offentlige Utredninger – Et bedre personvern – forslag til lov om behandling av personopplysninger*’, NOU 1997:19, (Statens Forvaltningstjeneste, Statens Trykning, Oslo 1997).

Norwegian Data Protection Tribunal (2005): ‘*Klage på Datatilsynets vedtak om begrensninger i adgangen til fjernsynsovervåking av publikumsområder i tog*’, Personvernemndas avgjørelse av 2. august 2006, Datatilsynets referanse: 2004/894 BSD/-

Norwegian Data Protection Tribunal (2005): ‘*Klage på vedtak om begrensninger i adgangen til kameraovervåking av bussenes publikumsområder*’, Personvernemndas avgjørelse av 2. august 2006, Datatilsynets referanse: 2004/915 SVE.

Swedish Law (2006): ‘*Frågan om viss uppgift skall vara anse som personuppgift enligt personuppgiftslagen*’ (1998:204); PUL, Dom 27.12.2006, Meddelad i Stockholm, Mål nr.: 15646-05, Rotel 441.

## B Books

Bygrave, Lee. A.: '*Data Protection Law- Approaching its Rationale, Logic and Limits*', (Kluwer Law International, 2002).

Cojocarasu, D. Irina: '*Anti-Spam Legislation – Between Privacy and Commercial Interest; An Overview of the European Union Legislation Regarding The E-Mail Spam*', Complex 1/06 Norwegian Research for Computers and Law (Unipubskriftserier, Institutt for Rettsinformatik, 2006).

Cojocarasu, D. Irina: '*Legal Issues Regarding Whois Databases*', Complex 2/09 Norwegian Research Center for Computers and Law (Unipubskriftserier, Institutt for Rettsinformatik, 2009).

Coll, Line: '*Innsyn I Personopplysninger I Elektroniske Markedsplasser*', (Complex 3/2002, Institutt For Rettsinformatik).

Hoeren Thomas: '*Internet- und Kommunikationsrecht-Praxislehrbuch*', (Verlag Dr.Otto Schmidt, 2008).

Hoeren, Thomas & Ulrich, Sieber: '*Handbuch Multimedia Recht, 19. Ergaenzungslieferung 2008*', (Beck C.H., 2008).

Jahnel, Dietma & Schramm, Alfred & Staudegger, Elisabeth: '*Informatikrecht*', (SpringerWienNewYork, 2000).

Schaar, Peter: '*Datenschutz im Internet – Die Grundlagen*', (Beck C.H., 2002).

Schartum, Dag Wiese & Bygrave Lee A.: '*Personvern i informasjonssamfunnet – en innføring i vern av personopplysninger*', (Fagbokforlaget Vigmostad & Bjørke AS, 2004).

Vogenauer, Stefan: *'Die Auslegung von Gesetzen in England und auf dem Kontinent'*, Band I, 2000, pp.400-425, Max-Planck-Institut fuer auslaendisches und internationles Privatrecht- Beitraege zum auslaendischen und internationalen Privatrecht.

Widmaier, Gunter: *'Muenchener Anwaltshandbuch Strafverteidigung'*, (Beck C.H., 2006).

C Journal articles

C1 off –line articles and journals

Church, Peter & Kon, Georgina: *'Data Protection and search engines-Google at the heart of a data protection storm'*, in Computer Law & Security Report, 2008, pp 461-465.

Debussere, Frederic: *'The EU E-Privacy Directive: A Monstrous Attempt to Starve the Cookie Monster?'*, in International Journal of Law and Technology, Volume 13 No.1, Oxford University Press, 2005, pp 70-97.

Foss, M & Bygrave L.A.: *'International Consumer Purchases through the Internet: Jurisdictional Issues pursuant to European Law'*, in International Journal of Law and Technology, Volume 8 No.2, Oxford University Press, 2000, pp 99-138

Goldman, Eric: *'Search Engine Bias And The Demise Of Search Engine Utopianism'*, in Yale Journal of Law and Technology, 2006, pp 188-200.

Holznagel, Bernd: *'Domainnamen- und IP-Nummern-Vergabe- eine Aufgabe der Regulierungsbehoerde?'*, in MultiMedia und Recht, issue 4, 2003, pp 219-222.

Introna, Lucas and Nissenbaum, Helen: *'Shaping the Web: Why the Politics of Search Engines Matters'*, in *I/S: A Journal of Law and Policy For The Information Society*, July-September 2000, 16 (3).

Koecher, Jan: *'Anmerkung zum Urteil des Landgerichts Berlin in Speicherung von IP-Adressen'*, in *MultiMedia und Recht*, issue 12, pp 800-801.

Kuehling, Juergen & Gauss, Nicole: *'Suchmaschine-eine Gefahr fuer den Informationszugang und die Informationsvielfalt?'*, in *Zeitschrift fuer Urheber- und Medienrecht*, 2007, pp 881-889.

Landgericht Berlin: *'Speicherung von IP-Adressen bei Nutzung des Portals bmj.bund.de'*, in *MultiMedia und Recht*, issue 12, 2007, pp 799-801.

Landgericht Darmstadt: *'Datenspeicherung bei Flatrate-Tarif'*, in *MultiMedia und Recht*, issue 5, 2006, pp 330-336.

Meyerdierks, Per: *'Sind IP-Adressen personenbezogene Daten?'*, in *MultiMedia und Recht*, issue 1, 2009, pp 8-13.

Ott, Stepan: *'Das Internet vergisst nicht – Rechtsschutz fuer Suchobjekte?'*, in *MultiMedia und Recht*, issue 3, 2009, pp 158-163.

Pahlen-Brandt, Ingrid: *'Datenschutz braucht scharfe Instrumente - Beitrag zur Diskussion um personenbezogene Daten'*, in *Datenschutz und Datensicherheit (DuD)*, 2008, issue 1, pp 34-37.

## C2 On line journals

Amtsgericht Muenchen: '*data protection; storing of an IP-address*', ruling 30.09.2008, file number 133 C 5677/08, available at: <http://beck-online.de>

Fleischer, Peter: '*Can a website identify a user based on a IP-address?*', in Peter Fleischer's blogspot 02/2008, available at: <http://peterfleischer.blogspot.com/2008/02/can-website-identify-user-based-on-ip.html>

Fleischer, Peter: '*Are IP-addresses "Personal Data"?*', in Peter Fleischer's blogspot 02/2007, available at: <http://peterfleischer.blogspot.com/2007/02/are-ip-address-personal-data.html>

*'Google Response to the Article 29 Working Party Opinion On Data Protection Issues Related to Search Engines 09/2008'*, available at: <http://www.scribd.com/doc/5625427/google-ogb-article29-response>

Landau, Markus: '*Ueberblick ueber Erfolgsmaße für das CRM im Internet*' im Rahmen des Seminars innerhalb des Schwerpunkts Wertschoepfungsmanagement mit dem Thema "*Customer Relationship Management (CRM) im Internet*", Wintersemester 2001/2002, available at <http://www.ecommerce.wiwi.uni-frankfurt.de>

Pouillet, Yves: '*Report on the Application of Data Protection Principles to worldwide telecommunication networks*' for the Council of Europe's T-PD Committee, point 2.3.1, T-PD 2004, available at <http://www.search.coe.int/taxis/search>

Schoech, Voker C.: '*Die Suchmaschine Google, Seminar: Algorithmen fuer das WWW*', 2001, available at: [http://www.inf.fuberlin.de/lehre/WS01/netbasedIS/uebungen/PageRank\\_vortrag\\_schoech.pdf](http://www.inf.fuberlin.de/lehre/WS01/netbasedIS/uebungen/PageRank_vortrag_schoech.pdf)

Reidenberg, J.R. & Schwartz, P.M.: '*Data Protection Law and On-line Services: Regulatory Responses*', European Commission 1998, available at

[http://europe.eu.int/comm/internal\\_market/en/media/dataprot/studies/regul.pdf](http://europe.eu.int/comm/internal_market/en/media/dataprot/studies/regul.pdf)

Tanase, Matthew : *'IP Spoofing-an Introduction'*, in Security Focus, available at <http://www.securityfocus.com/infocus/1674>

The Official Google Blog.: *'Making ads more interesting'* posted by Susan Wojcicki 11/2009, available at: <http://google.blogspot.com/2009/03/making-ads-more-interesting.html>

Tielemans, Henriette representing Covington & Burling: *'Comments on Implementations and Application of the 1995 Data Protection Directive'*, available at [http://ec.europa.eu/justice\\_home/fsj/privacy/docs/lawreport/paper/covington-burling\\_en.pdf](http://ec.europa.eu/justice_home/fsj/privacy/docs/lawreport/paper/covington-burling_en.pdf)

## D Reports and other documents

### D1 From the Art.29 Data Protection Working Party

*"Opinion 5/2002 on the Use of Public Directories for Reverse or Multi-criteria Searching Services (Reverse directories)"*, 5058/00/EN/FINAL WP 33.

*"Opinion 4/2007 on the concept of personal data"*, 01248/07/EN WP 136.

*"Opinion 1/2008 on data protection issues related to search engines"*, 00737/EN/ WP 148.

*"Working document on determining the international application of EU data protection law to personal data processing on the Internet by non-EU based web sites"*, 5035/01/EN/Final WP 56.

*"Working document on Privacy on the Internet –An integrated EU Approach to On-line Data Protection-*", 5063/00/EN/FINAL WP 37.

## D2 from the European Commission

Commission of the European Communities: “*On-line services and data protection and privacy – Regulatory responses*”, Volume II, 1998.

Commission of the European Communities: “*On-line services and data protection and the protection of privacy; Part one – Description of the general situation; Part two – Case studies*”, Volume I, 1999.

## E From other sources

‘Centre for Democracy & Technology’: “*A Primer on Behavioral Advertising*”, July 2008, available at: <http://cdt.org/publications/policyposts/2008/12>

‘FAQ 14-Pharmaceutical and Medical Products’, available at <http://www.export.gov/safeharbor/FAQ14PharmaFINAL.htm>.

Eric van Ess’s Search Bistro, “*Google Secret Lab, Prelude*”, available at <http://www.searchbistro.com/index.php?/archives/19-Google-Secret-Lab,-Prelude.html>

‘28th International Data Protection and Privacy Commissioners' Conference London’, 2 and 3 November 2006 on “*Privacy Protection and Search Engines*”, available at [http://ec.europa.eu/justice\\_home/fsj/privacy/news/docs/pr\\_google\\_annex\\_16\\_05\\_07\\_en.pdf](http://ec.europa.eu/justice_home/fsj/privacy/news/docs/pr_google_annex_16_05_07_en.pdf)

The Internet Engineering Task Force: ‘*Hypertext-Transfer-Protocol*’, available at <http://www.ietf.org/rfc/rfc2616.txt>

Internet Protocol- ‘*Darpa Internet Program Protocol Specification*’, September 1981, prepared for Defense Advanced Research Projects Agency by Information Sciences Institute, available at <http://tools.ietf.org/html/rfc791>

Press Release: '*Nielsen//Netratings*', available at  
[http://www.nielsenonline.com/press.jsp?section=ne\\_press\\_release&nav=1](http://www.nielsenonline.com/press.jsp?section=ne_press_release&nav=1)

'Princeton University's Wordnet Search 3.0',  
available at <http://wordnetweb.princeton.edu/perl/webwn>

RFC Sourcebook – 'IP,Internet Protocol', Available at  
<http://www.networksorcery.com/enp/protocol/ip.htm>

## Annex

Table 1:

Information on 129.240.179.20 (static IP-address assigned to the University of Oslo)

<b>IP Location:</b>	 Norway Oslo University Of Oslo Norway
<b>Resolve Host:</b>	jusiri59.uio.no
<b>IP Address:</b>	129.240.179.20 
<b>Blacklist Status:</b>	Clear

## Whois Record

```
OrgName: RIPE Network Coordination Centre
OrgID: RIPE
Address: P.O. Box 10096
City: Amsterdam
StateProv:
PostalCode: 1001EB
Country: NL

ReferralServer: whois://whois.ripe.net:43

NetRange: 129.240.0.0 - 129.242.255.255
CIDR: 129.240.0.0/15, 129.242.0.0/16
NetName: RN-ERX-129-240-0-0
NetHandle: NET-129-240-0-0-1
Parent: NET-129-0-0-0-0
NetType: Early Registrations, Transferred to RIPE NCC
```

Comment: These addresses have been further assigned to users in  
Comment: the RIPE NCC region. Contact information can be found in  
Comment: the RIPE database at <http://www.ripe.net/whois>  
RegDate: 2003-01-10  
Updated: 2003-06-18

== Additional Information From whois://whois.ripe.net:43 ==

inetnum: 129.240.0.0 - 129.240.255.255  
netname: UNIONET  
descr: University of Oslo, Norway  
country: NO  
admin-c: LO20-RIPE  
tech-c: KB100-RIPE  
tech-c: UH22-RIPE  
status: EARLY-REGISTRATION  
rev-srv: ifi.uio.no nissen.uio.no nn.uninett.no  
mnt-by: UNINETT-MNT  
mnt-lower: UNINETT-MNT  
mnt-irt: IRT-UNINETT-CERT  
source: RIPE # Filtered

irt: IRT-UNINETT-CERT  
address: UNINETT CERT  
address: Abels gate 5  
address: N7465  
address: Trondheim  
address: Norway  
phone: +47 73557900  
fax-no: +47 73557901  
e-mail:  
signature: PGPKEY-E26339A9  
encryption: PGPKEY-E26339A9  
admin-c: TI123-RIPE  
tech-c: TI123-RIPE  
auth: PGPKEY-E26339A9  
auth: PGPKEY-9ADF790B  
auth: PGPKEY-2508D151  
remarks: Emergency telephonenumber +47 73557961 (GMT+1/GMT+2 with DST)  
remarks: <http://www.trusted-introducer.org/teams/uninettcert.html>  
remarks: This is an accredited IRT (level 2)  
irt-nfy:

mnt-by: TRUSTED-INTRODUCER-MNT  
source: RIPE # Filtered

person: Lars Oftedal  
address: Universitets Senter for Informasjonsteknologi  
address: Universitetet i Oslo  
address: Postboks 1059, Blindern  
address: N-0316 Oslo  
address: Norway  
phone: +47 22 85 24 70  
phone: +47 22 85 25 20  
fax-no: +47 22 85 27 30  
e-mail:  
nic-hdl: LO20-RIPE  
source: RIPE # Filtered

person: Knut Borge  
address: USIT/UiO  
address: Gaustadalleen 23, Blindern  
address: Postboks 1059 Blindern  
address: N-0316 Oslo  
address: NORWAY  
phone: +47 22 85 25 19  
fax-no: +47 22 85 27 30  
e-mail:  
nic-hdl: KB100-RIPE  
source: RIPE # Filtered

person: UiO.no Hostmaster  
address: c/o Knut Borge  
address: Universitets Senter for Informasjonsteknologi  
address: Universitetet i Oslo  
address: Postboks 1059 - Blindern  
address: N-0316 Oslo  
address: Norway  
phone: +47 22 85 24 70  
phone: +47 22 85 25 19  
fax-no: +47 22 85 27 30  
e-mail:  
nic-hdl: UH22-RIPE  
source: RIPE # Filtered

Table 2:

Information on 81.217.55.109 (static IP-address assigned to a customer by an Access Provider in a contractual relationship)

```
81.217.55.109 - Whois Information

% This is the RIPE Database query service.
% The objects are in RPSL format.
%
% The RIPE Database is subject to Terms and Conditions.
% See http://www.ripe.net/db/support/db-terms-conditions.pdf

% Note: This output has been filtered.
%       To receive output for a database update, use the "-B" flag.

% Information related to '81.217.0.0 - 81.217.159.255'

inetnum:        81.217.0.0 - 81.217.159.255
netname:        AT-KABELSIGNAL-MCNS-2-NET
descr:          Kabelsignal AG
remarks:        -----
remarks:        Abuse/Spam notifications to: abuse@kabsi.at
remarks:        -----
country:        AT
admin-c:        TE439-RIPE
tech-c:         KABS1-RIPE
status:         ASSIGNED PA
mnt-by:         AS8339-MNT
mnt-lower:      AS8339-MNT
source:         RIPE # Filtered

role:           KABS1.AT Hostmaster Role Account
address:        KABELSIGNAL AG
address:        Suedstadtzentrum 4
address:        A-2346 Maria Enzersdorf
phone:          +43 2236 45564 0
fax-no:         +43 2236 45564 2030
e-mail:         hostmaster@kabsi.at
remarks:        -----
remarks:        Abuse/Spam notifications to: abuse@kabsi.at
remarks:        -----
admin-c:        TE439-RIPE
```

```
tech-c:      TE439-RIPE
tech-c:      CS60-RIPE
tech-c:      PH708-RIPE
tech-c:      FZ110-RIPE
tech-c:      MG8573-RIPE
nic-hdl:     KABS1-RIPE
mnt-by:      AS8339-MNT
source:      RIPE # Filtered
```