FINAL FINDINGS

Dutch Data Protection Authority investigation into the collection of Wifi data by Google using Street View cars

z2010-00582

7 December 2010

PUBLIC VERSION

TABLE OF CONTENTS

1. Introduction

Pursuant to Article 60 of the Wet bescherming persoonsgegevens (Wbp) [Dutch Personal Data Protection Act], the Dutch Data Protection Authority (Dutch DPA) [College bescherming persoonsgegevens (Cbp)] has initiated an investigation ex officio into the collection of data with respect to Wifi networks in the Netherlands (hereinafter: Wifi data) by Google Inc. (hereinafter: Google) using Street View cars.

The company collected the Wifi data when making photos for its *Street View* service. Google uses this service to publish photos of houses, streets and squares on the Internet. The data were collected using Street View cars. The cars are equipped with cameras to make panoramic photos. From March 2008 to May 2010, the cars in the Netherlands were also equipped with an antenna and related equipment and software to pick up signals from wireless Internet traffic.

Google announced on 27 April 2010 that it was collecting wireless network information worldwide using the Street View cars.[1] At the time, the company stated that it only collected data *on* (the existence of) Wifi routers. The company stated that it did not collect data that originated *from* wireless Internet networks. On 14 May 2010, Google announced on its 'Official Google Blog' that, during the collection of data on Wifi networks, it had also recorded and stored the content of communication.[2]

On the basis of the above statements, the Dutch DPA, pursuant to its supervisory role, initiated an investigation. The investigation focused on the following five questions:

- What sort of Wifi data did Google collect precisely, both *on* and *originating from* Wifi networks? Are these personal data as defined in Article 1, header and under a, Wbp?
- Do the collected Wifi data also include *special* categories of personal data, as defined in Article 16 Wbp?
- For what purpose did Google collect the Wifi data? Are these purposes legitimate, within the meaning of Article 7 Wbp?
- Did Google inform the data subjects of the various purposes for which the personal data were collected and processed, as prescribed in Articles 33 and 34 Wbp?
- Did Google notify the Dutch DPA of the processing of the Wifi data, within the meaning of Article 27, in conjunction with 28 Wbp?

The investigation focuses on verification against Articles 7 (legitimate purposes), 8 (legitimate ground for processing: consent or necessity), 16 in conjunction with 23 (prohibition on the processing of *special categories of* personal data, unless with *explicit* consent), 27 in conjunction with 28 (notification to the Dutch DPA), 33, 34 (obligation to provide information) and 6 Wbp (fair and careful data processing).

[1] Source: Google, 'Data collected by Google cars', 27 April 2010, URL: http://googlepolicyeurope.blogspot.com/2010/04/data-collected-by-google-cars.html.
[2] Source: Google, 'WiFi data collection: An update', 14 May 2010, with updates from 17 May and 9 June 2010, URL: http://googleblog.blogspot.com/2010/05/Wifi-data-collection-update.html

2. Proceedings

2.1 Course of the proceedings

The Dutch DPA requested by letter dated 21 May 2010 that written information be provided by Google. Google was requested to reply within ten calendar days. Google requested, by email dated 27 May 2010, that the Dutch DPA grant it a postponement of the date to reply. By letter dated 27 May 2010 the Dutch DPA granted a postponement of the date to reply until 8 June 2010 at the latest. The Dutch DPA received part of the requested information by letter dated 8 June 2010.[3]

The Dutch DPA announced, by letter dated 17 June 2010, that it would conduct an onsite investigation on 28 June 2010. The Dutch DPA requested Google to cooperate in the aforementioned investigation by granting insight into and creating two readable forensic copies, namely of data on a representative hard disk with Wifi data recorded in the Netherlands, and of the file with all communication content data that originated from wireless networks in the Netherlands, hereinafter: *payload data*. Google sent in a written reasoned request to postpone the investigation by one day to Tuesday 29 June 2010. This request was granted.

Google indicated, by letter dated 24 June 2010, that it would not cooperate in creating the requested forensic copies. The Dutch DPA demanded cooperation, by letter of 28 June 2010.

The Dutch DPA commenced the onsite investigation on 29 June 2010. During the investigation, Google did not cooperate in creating the requested, readable, forensic copies. At a later stage Google did provide the requested copy of all *payload data* collected in the Netherlands after all, see below. On that day, the Dutch DPA's investigation focused on examining the data on the hard disk, by means of *remote access* to a computer in Belgium containing a copy of this hard disk.

During the aforementioned investigation, the Dutch DPA obtained verbal information from two relevant technical employees and a Google company lawyer authorised by the director. The Dutch DPA took away several (copies of) documents in the form of three hardcopies of 1 A4. A copy of the documents drawn up and copied during the investigation was provided to Google.

During the aforementioned investigation, the Dutch DPA requested Google to provide information – by 2 July 2010 at the latest – in response to the remaining two questions stated in the letter of 21 May 2010, which information was still lacking. Google failed to provide this information in time. In connection with the factual questions that arose during the aforementioned investigation, the Dutch DPA sent an additional request for information. Google provided part of the requested (additional) information on 6 July 2010, but not the information requested on 21 May 2010.

By letter dated 8 July 2010, the Dutch DPA demanded a full answer to the questions posed by 5 pm on 12 July 2010 at the latest. The Dutch DPA also requested new information and access to the business data and documents, by letter dated 8 July 2010, with respect to the *payload data* captured in the Netherlands, with the request the provide this information at the latest by 5 pm on 15 July 2010. Google provided the two answers to the questions of 21 May 2010 that had still been lacking by means of two emails dated 12 July 2010, and indicated that it needed more time to be able to provide further information. On 15 July 2010 Google provided part of the requested

[3] 2 of the 11 questions had not been answered.

information, by email sent at 4:48 pm. Google did not provide the requested hardcopies with respect to the *payload data*. Instead, Google offered a copy that had been rendered legible, containing a file with all *payload data* collected in the Netherlands.

On 16 July 2010, the Dutch DPA sent a reminder of its demand by letter dated 8 July 2010. Google was demanded, by means of this letter, to provide a copy – that had been rendered legible – of the file containing all payload data collected in the Netherlands, by Monday 19 July 2010 at the latest. The Dutch DPA also demanded additional information to clarify uncertainties in previously provided information, with the instruction to provide this information at the latest on 21 July 2010. By email dated 19 July 2010, Google submitted a reasoned written request for a postponement with respect to the provision of the demanded copy of the file containing all *payload data* collected in the Netherlands. By letter dated 20 July 2010, the Dutch DPA granted a postponement, until 22 July 2010 at the latest. The Dutch DPA received the hard disk containing the *payload data* on 21 July 2010 and drew up an official record of the transfer. Google received a copy thereof. By fax dated 21 July 2010, the company Stroz Friedberg, acting on Google's behalf, provided the password to access the data on the hard disk. By email dated 22 July 2010, Google provided the additional information that was demanded by letter dated 16 July 2010.

The Dutch DPA sent its report of provisional findings to Google by letter dated 21 September 2010, with the request to respond within two weeks, thus by 5 October 2010. By letter dated 22 September 2010 Google submitted a reasoned request for postponement until 2 November 2010. By letter dated 8 October 2010 Google requested permission to submit its opinion of the report of provisional findings in English. The Dutch DPA indicated, by letter dated 11 October 2010, that the opinion must be written in the Dutch language. Google then requested postponement of 1 week, by letter dated 12 October, in order to be able to provide its response in the Dutch language. By letter dated 12 October 2010 the Dutch DPA granted another week's postponement - until 26 October 2010. Google submitted its opinion to the Dutch DPA by letter dated 26 October 2010.

By letter dated 9 November 2010 (further specified by letter dated 11 November 2010), the Dutch DPA requested Google to indicate, within ten calendar days, which of the information provided was considered confidential by the company. Google responded by letter dated 19 November 2010. The Dutch DPA sent Google the final Findings by letter dated 23 November 2010. This also included a draft public version of the Findings, requesting Google to indicate within five calendar days whether the public version still contained any information that, in the opinion of Google, was considered confidential by the company. In response, Google twice requested postponement of its answer. Both requests were denied. By email sent on the evening of Friday 26 November 2010, Google responded substantively to the request to (again) check for company confidential information.

3. Facts

3.1 General Google working method as regards Wifi data

Google Inc., incorporated on 4 September 1998, with its headquarters in Mountain View, California in the United States, has as its purpose to "*to make all information worldwide accessible and usable*".[4] For this purpose, Google not only offers a search engine, but a whole range of online services, varying from webmail to the sale of online advertisements, and from online maps (Maps) to a browser (Chrome). In the Netherlands, Google Inc. operates as Google Netherlands BV, with its registered office in Amsterdam and registered with the Chamber of Commerce since 27 November 2003 under number 34198589. The company description of Google Netherlands BV is as follows: "*To conduct a business in the field of an Internet search engine and to provide services and information and advice in the field of searching and retrieving information from Internet, Intranet and other (electronic) communication.*"

Google collected Wifi data when making photos for the *Street View* service the company offers. As part of this service Google publishes photos on the Internet of houses, streets and squares all over the world.[5] The data were collected worldwide in 30 countries with the aid of Street View cars.[6] The cars are equipped with cameras to make panoramic photos. From March 2008 to May 2010, the cars in the Netherlands were also equipped with an antenna and related hardware and software to pick up and store signals from wireless Internet networks.

Google captured the Wifi data with the aid of a generally available antenna.[7] The data recorded were captured with open source software (Kismet), which is also generally available. The data obtained using Kismet were subsequently processed for storing by means of a Google proprietary application (gstumbler and gslite as executable). The software was set to change channels five times per second to make recordings of 0.2 seconds per channel.

The cars were equipped with [CONFIDENTIAL: description of the technical equipment of the cars and technical method of determination of the location].

The Wifi signals were stored on a hard disk in the Street View car. Google delivered these hard disks to a warehouse in Belgium each week by courier. The content of the hard disk was sent to a Google data centre in Belgium from this warehouse, and from there on to Google servers in the United States via the Internet. Google declared that it did not archive or create a backup of the hard disks, because the loss of a single hard disk would cost less than setting up an archiving system or making backups.[8]

Google recorded two different types of data by means of the Street View cars driving through the neighbourhoods: communication content data (*payload data*) originating from unprotected

---

[4] Source: Google company information, URL: http://www.google.nl/intl/nl/corporate/facts.html
[5] The photos can be accessed at http://maps.google.com. This is a website containing maps and satellite images. Users of this online service can zoom in down to street level to view Street View photos.
[6] Source: Google, 'Where is Street View available?', URL: http://maps.google.com/intl/en_us/help/maps/streetview/where-is-street-view.html
[7] Maxrad BMMG24005 omnidirectional antenna. Source: Google, 'Data collected by Google cars', 27 April 2010, URL: http://googlepolicyeurope.blogspot.com/2010/04/data-collected-by-google-cars.html.
[8] Statement by Google during the onsite investigation on 29 June 2010.

wireless Wifi networks in the Netherlands and data on (the existence and location of) Wifi routers and other communication equipment.

3.2 Payload data

On 9 June 2010 Google published a report by Stroz Friedberg LLC., an American digital forensics firm.[9] Commissioned by Google, the firm reviewed the (source code of the) software that Google used to collect the Wifi data. The report confirms that the software was set to capture and store payload data from unencrypted Wifi networks.

The Stroz Friedberg analysis of the source code shows that the software offered the possibility to store or not store communication content data of unencrypted Wifi routers and other Wifi-capable devices.[10] The programmer was able to choose between *true* (on) and *false* (off) to store the data from unencrypted networks and selected *true.* This choice means that Google did store the *payload data* of unencrypted networks, but not that of encrypted networks.[11]

During the onsite investigation, the Dutch DPA examined the hard disk containing Wifi data with respect to routers in an area south of Rotterdam, from the period of 28 April to 6 May 2010. See **Annex I** for a map of where these Wifi data were collected. The Dutch DPA had requested a disk containing data from an area that would be representative of the Netherlands both in terms of population density and Internet use.

Google has indicated that there were two other hard disks at the warehouse in Belgium and that the hard disk shown was the most representative hard disk available. It later turned out that the area was a relatively sparsely populated area south of Rotterdam.[12]

The data on the hard disks from the Street View cars were stored in *binary* format. During the onsite investigation, Google translated the alpha numerical data into (legible) *ascii* format using its own software, for the benefit of the Dutch DPA. In doing so, the non-alpha numerical signs were converted into *octal* format.[13]

The nature of the access to the data provided by Google to the Dutch DPA (remote access via a command line) has put considerable technical constraints on the possible questions that could be posed.

---

[9] Source: Google, 'WiFi data collection: An update', 14 May 2010, hyperlink in the update of 9 June 2010 to: Stroz Friedberg, 'Source Code Analysis of gstumbler, 3 June 2010, URL: http://www.google.com/googleblogs/pdfs/friedberg_sourcecode_analysis_060910.pdf

[10] A WEP-bit is sent with nearly every packet, both in a number of management frames and in the control frames and the data frames. This bit only provides a reliable indication of whether the payload has been encrypted in the data frame, where the payload data are stored.

[11] There were, as regards writing data to the hard disk, a total of four options for standard settings. Source: Stroz Friedberg report, page 11, paragraph 56.

[12] By letter dated 2 July 2010, the Dutch DPA received a map from Google of the area south of Rotterdam where the recordings on the hard disk examined by the Dutch DPA were made.

[13] The software Google used to render the protocol buffer format readable is called codex. The protocol buffer is described by Google in a Developers Guide at http://code.google.com/intl/nl-NL/apis/protocolbuffers/docs/overview.html. In derogation from this description, Google has applied its own proprietary sequence when writing the Wifi data to the hard disk.

During the investigation, the Dutch DPA compared the source code used by Google with the source code described in the Stroz Friedberg report. It was established that with the exception of two deviations, the software used by Google was in conformity with the source code described.[14]

In view of the technical restrictions referred to above, the Dutch DPA was unable to form an opinion on the data collected by Google *on* the Wifi routers, nor was it able to form a comprehensive image of the substance of the *payload data*. For this reason, the Dutch DPA demanded further cooperation in the creation of a forensic copy of the file stored in the United States containing all *payload data* collected in the Netherlands and rendered legible.

The file obtained consists of *packets*, i.e. information that is structured to be transported across the packet-switched Internet as intelligently as possible. A packet consists of two types of information. Control information and user data, comparable with a letter (*payload*) in an envelope (information such as the sender and destination).

In the file containing all *payload data* collected in the Netherlands, the first packet dates from 7 March 2008. The final packet dates from 29 April 2010. Therefore Google had collected data in the Netherlands during a total period of 451 days.

It has now been established that at that time Google collected nearly 30 gigabytes of *payload data* (translated into legible format) in the Netherlands, which originated from Wifi networks and which was recorded on 714 hard disks containing a total of 4,774 gstumbler log files.[15]

The qualitative analysis for calculating the various percentages is based on these 4,774 log files.[16]

The recorded data are not meaningless fragments. It is factually possible to capture 1 to over 2,500 packets per individual user in 0.2 seconds. Moreover, the car may have captured a signal from a single Wifi router several times, for example if the car stopped for a red traffic light, or if the car drove around a building.

As each *packet* contains address details, such as (internal) IP addresses and / or email addresses, it is possible to link several *packets* from 1 Internet user to each other, and in doing so construct an accurate picture of the communication of an often identifiable user.

The investigation showed that Google not only stored data from the field 'raw data' in the file containing *payload data*, but also 12 additional fields, including the latlon of the car at the moment the packet was captured. It is therefore accurately known for each packet that was captured where GPS-reception was possible – what the position of the car was at the time the packet was captured and stored.

---

[14] The source code of the software used by Google contained two additional fields not described in Annex C of the Stroz Friedberg report. The Dutch DPA has established that, in any case, these two fields do not contain payload data.

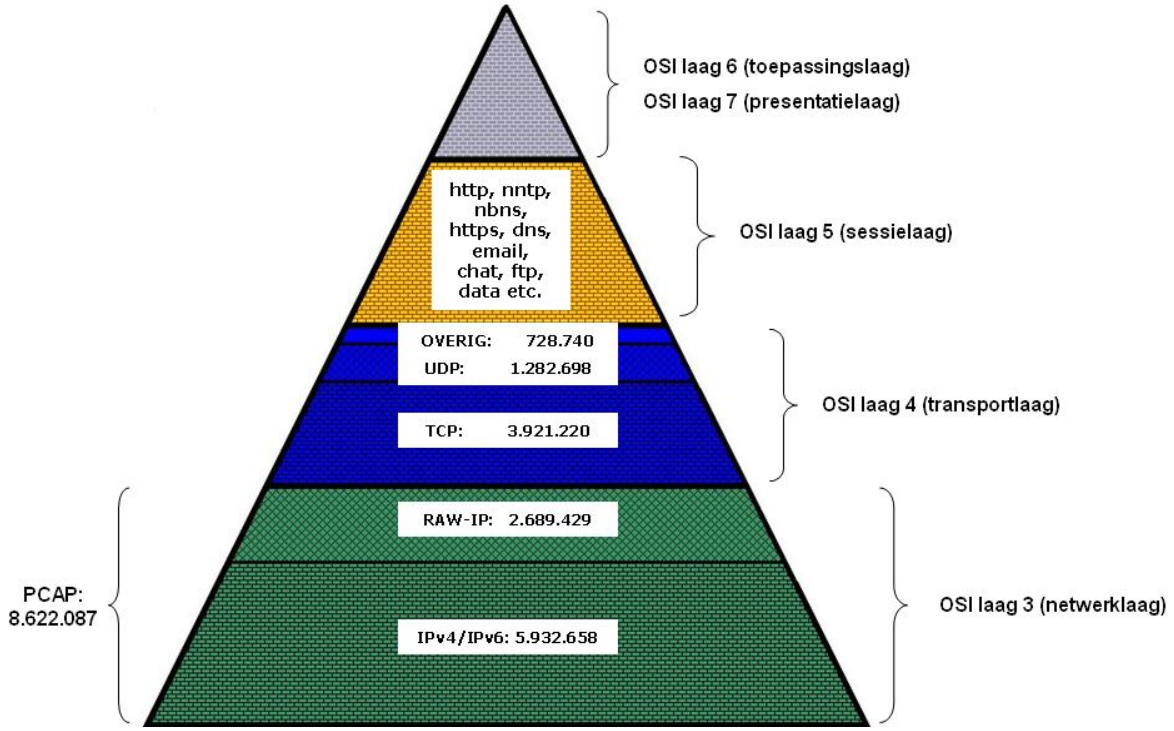
[15] In its view of the provisional findings of 26 October 2010, Google added that the original file in binary form contained only 6.75 gigabytes. The file received by the Dutch DPA from Google contained 29.8 gigabytes of payload data that had been rendered readable. The Dutch DPA's investigation focused on the RAW fields present in this file.

[16] In its opinion Google states that the payload data only concerned a small part of the material on the hard disks, including photos for the Street View service. The Dutch DPA examined the content of the copy of the payload data collected by Google and made available to the Dutch DPA.

The Dutch DPA converted the *payload data*, as translated by Google from proprietary binary format to octal format and ascii for the benefit of the Dutch DPA, back to binary format, in order to be able to systematically examine the data. Over 8.8 million Wifi packets were found in the binary pcap file of 4.7 gigabytes thus obtained.[17]

The *payload data* can be described on the basis of the various layers of the OSI model. This model is a standardised structuring of network traffic in 7 layers. The layers are, from high to low, Physical Layer, Data Link Layer (LLC), Network Layer (routing of *packets* via IPv4 and IPv6), Transport Layer (TCP and UDP), Session Layer (maintaining and ending sessions between two communicating hosts), Presentation Layer and Application Layer (email, web traffic, IRC etc.).

GRAPH I: structure of *packets* according to the OSI model



[*OSI layer 6 (Application Layer)*
OSI layer 7 (Presentation Layer)
OSI layer 5 (Session Layer)
OSI layer 4 (Transport Layer)
OSI layer 3 (Network Layer)]

The Dutch DPA removed the LCC part of the *packets* from the second layer of the OSI model in order to be able to use tools generally available on the Internet for further analysis in layers three and further. This leaves over 8.6 million packets in pcap format that are in principle legible. A part of these packets cannot be examined further, as it only concerns a fragment.

[17] This report consistently speaks of packets, also where it concerns other layers of the OSI model. The official term is for example 'frame' for packets in the Datalink Layer and 'data' for packets in the outer Application Layer.

The type of Internet traffic cannot be established on the basis of such fragments. 2.7 million fragments of the 8.6 million packets concern raw IP (31.2%).

The Dutch DPA's further examination was limited to 68.8 percent of the legible data, totalling 5,932,658 packets that are recognisable as IPv4 and IPv6 in layer three of the OSI model.

These *packets* consist for 66% of TCP traffic, for 22% of UDP traffic (used, for example, for DNS queries) and for 12% of other traffic.[18]

All these *packets* contained at least two IP addresses: of the sender and the receiver. In this context, the receiver is taken to mean a device connected to a Wifi router. The public IP addresses of these receivers were not stored, but rather the internal, non-unique addresses in the ranges 10.0.0.1 to 10.255.255.254, 172.16.0.1 to 172.31.255.254 and 192.168.0.1 to 192.168.255.254. In the case of a website visit, a packet also contains a public IP address (of the website).

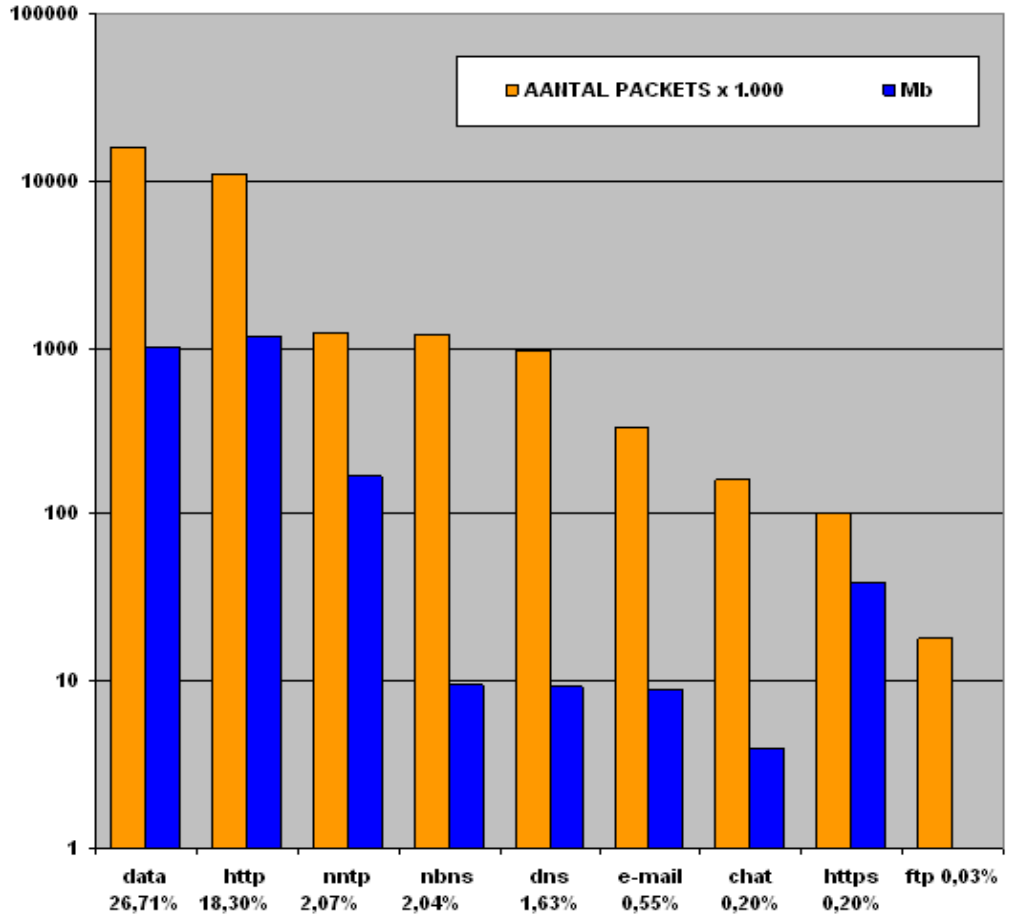
The *packets* were further analysed with the aid of programmes such as Wireshack, capinfos, tshark, ngrep, perl scripts and foremost in order to be able to make a qualitative analysis of the types of Internet traffic (such as email, web traffic etc.) in the TCP and IP layer of the traffic. See **Annex II** for a flowchart of the analysis method.

The average size of each packet captured is 564.58 bytes, i.e. over 500 numbers, letters or signs, includes spaces). The number of packets Google would have been able to capture every 0.2 seconds depends on the processing time of the Wifi network. The maximum speed of the 802.11g standard is 54 Mbit per second. This means that a maximum of 2,507 packets can be sent (in all layers of the OSI model) and therefore captured. At a speed of 5.5 Mbits per second, a maximum of 255 packets can be sent in 0.2 seconds, and a maximum of 46 packets at a speed of 1 Mbit.

Packets were often transmitted and stored by Google more than once, for example because a receiver asks for a repeat if the packet is not captured in full or if the sending computer does not receive confirmation of receipt. This inherent to the way in which Internet works.

[18] Other traffic consists for example of ICMP (used for example for ping), IGMP (used for multicasting games and video streaming) and L2TP (used for example for Virtual Private Lan).

GRAPH II: Number of packets per traffic stream and megabytes, in a logarithmic scale.



[*number of packets x 1,000*]

Data can contain all manner of Internet traffic. It is not possible to automatically analyse what protocol was used, because several fragments are lacking. These are legible, however, and therefore suitable for further analysis.

DNS is the abbreviation of domain name system. DNS servers translate domain names into IP addresses on the Internet.

NNTP is the protocol used in reading and posting contributions in Usenet newsgroups.

NBNS stands for NetBIOS Name Service. This traffic originates from computers using Windows as operating system in which the NetBIOS protocol is applied within a TCP-IP network.

Email contains POP traffic, SMTP and IMAP.

Chat contains IRC, Microsoft Messenger, Yahoo Messenger and Yabber.

3.2.1 E-Mail

98 unique passwords were found in the legible email traffic. A combination of login name, password and the IP address of the mail server was found 56 times in the captured email traffic. This concerns mail servers of well-known providers in the Netherlands, such as Caiway, Chello, KPN, Solcon, Tiscali, Versatel, Wanadoo, XS4ALL and Ziggo. In addition, login data of mail servers that can be used by multiple domains, which can only be used to log in after some searching, were also found.

In the incoming email traffic, the actual content of correspondence was found. The nature of this correspondence is highly diverse. The fact that email can be intercepted in legible form by third parties is caused by the fact that few providers offer a secure version of the Internet protocols to retrieve email from the server (secure POP) or to send it (secure SMTP).

The Dutch DPA found, for example, two consecutive incoming emails from a bank. The bank confirms a securities transaction performed by a client. The emails contain the initials, last name, email address and the bank account number of the relevant person, including the amount to be paid.

The payload data contains 32,862 packets of email traffic.[19] As many packets contain both the email address of the sender and/or receiver as well as the (non-public) IP address of the receiver, it is possible to link several packets – pertaining to one person – to each other, and therefore on the basis of several packets related to the same person, a picture can be constructed of the parties with whom he or she corresponds, and at what times and about which subjects the correspondence took place.

Another example from the payload data concerns the inbox of a client of a webmail provider. Based on the timing of the emails, the addresses of the senders and, in particular the subject line, it is possible to reconstruct an accurate picture of moments in the life of this data subject, his interests and his career development.

Email addresses

In total, the payload data contains 16,640 email addresses, part of which were captured in fragmented state.[20] The number of complete email addresses captured is 10,195. The majority thereof, 5,890 (57.8%), fall under the top level domain .com.

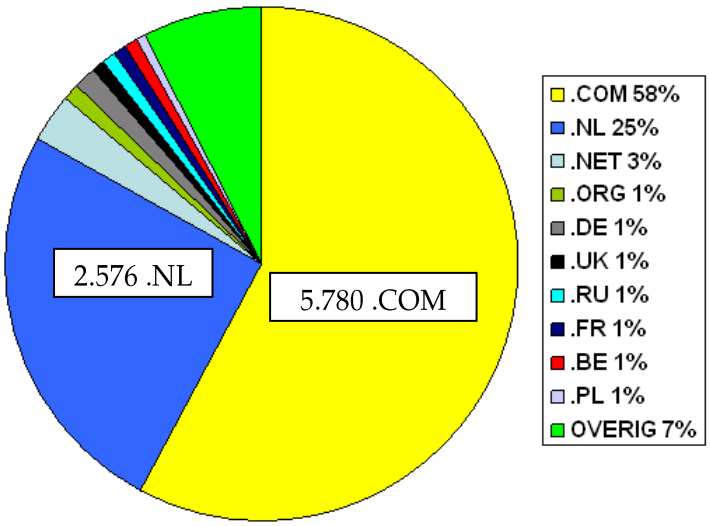
---

[19] The Dutch DPA counted the number of email *packets* as follows:
tshark -r tt-BIG-chunks_00000_20100810114607.pcap -R "pop || imap || tcp.port == 25 || tcp.port == 110" -w smtp_pop_imap_000.pcap && \
tshark -r tt-BIG-chunks_00001_20100810121956.pcap -R "pop || imap || tcp.port == 25 || tcp.port == 110" -w smtp_pop_imap_001.pcap && \
tshark -r tt-BIG-chunks_00002_20100810125505.pcap -R "pop || imap || tcp.port == 25 || tcp.port == 110" -w smtp_pop_imap_002.pcap && \
tshark -r tt-BIG-chunks_00003_20100810135058.pcap -R "pop || imap || tcp.port == 25 || tcp.port == 110" -w smtp_pop_imap_003.pcap && \
mergecap -w smtp_pop_imap.pcap smtp_pop_imap_000.pcap smtp_pop_imap_001.pcap smtp_pop_imap_002.pcap smtp_pop_imap_003.pcap && \
capinfos smtp_pop_imap.pcap.

[20] The search used is: "ngrep -I tt-fileset-1to20.pcap -q "[a-zA-Z0-9]+@[a-zA-Z0-9]+\....." | grep -Eirho '[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,4}' | sort | uniq"

Of the email addresses belonging to .com, over 55% were issued by hotmail.com (Microsoft), 9% by gmail.com (Google), 5% by yahoo.com and 3% by msn.com (Microsoft). Of the .nl addresses, 25% were issued by hyves.nl, 19% by Live.nl (Microsoft) and 3% by chat.spelpunt.nl.

GRAPH III: Number of email addresses per domain



3.2.2 Web traffic

In the payload data, all kinds of web traffic were also found . The web traffic contains, for example, 137 recognisable searches performed with the Google search engine.[21] A number of searches concerns sensitive data.

"green+poo+baby"
"how_many+weeks+pregnant"
"homecare+store+TWELLO"
"tinto+brass+posta+kutusu+18+porno"
"wedding dress+for+rent+in+amsterdam"
"<first name>+<last name>+<last name>"

Multiple packets were captured from each of the persons who performed these searches. 944 packets were recorded of a person who searched for a combination of a PO Box address and a porn movie by director Tinto Brass. It is evident from that traffic that this person visited a number of specific film sites and weblogs and viewed a number of erotic movies via those sites.

The last search concerns a unique combination of the first name, maiden name and last name.[22] This combination leads to a single person. It is possible that this person 'searched herself' on the Internet, a so-called vanity search.

[21] The database was subjected to a targeted search for the term www.google.nl/search? The results were assessed starting from "?q=". The double entries were then removed.
[22] Names anonymised by the Dutch DPA. The original is available in the Dutch DPA's investigation file.

Referrers
The web traffic also contained many so-called referrers, which are the URL's of the last web page a person visited. A total of 3,354 different addresses were found.[23] The referrers contain a large number of pages with titles that suggest pornography, dating and/or a sexual nature. Several pages with medical data and pages focusing on persons with a certain religion and/or personal belief were also found. Two pages were found specifically aimed at online job applications, my.monsterboard.nl and solliciteren.werkenbijdelandmacht.nl.

The investigation shows that the web traffic contains 983 URL's of profile pages of the social network site Hyves. Members of Hyves have the option of restricting access to their profile page to selected friends only. Members can do this to prevent third parties from establishing the existence of their page via external or internal search engines. As the users may have changed their setting in the course of time, it cannot be established which part of these URL's were restricted when Google captured the contents of the Internet traffic.

Other web traffic
Thirteen mobile phone numbers were found in the web traffic.

Two computers infected with malware were also found in the web traffic, which broadcasted their SSID and PSK Key in an unencrypted manner. The combination of the SSID and the PSK key allows malicious third parties to enter the security key of an up-to-date Wifi router (encrypted with WPA2) and subsequently read along with all traffic. The fact that malware was installed is also evident from the port used to send traffic: 34444.

3.2.3 NBNS

Word documents were found in the payload data in the category 'NBNS'. Part of the NetBIOS traffic is SMB traffic (Server Message Block), the network protocol used by Microsoft Windows to enable file exchange between multiple computers. If this file exchange takes place via a (unencrypted) Wifi router, the files can be captured by third parties in a legible format.

The same Word document was found twice in the NBNS traffic. This document contains a comprehensive psychological report concerning a minor, a boy with a serious reading disability. The report contains his first name, his last name, his address details and his date of birth. The report was marked as 'confidential'.[24]

3.2.4 Chat / MSN

Chat traffic was also found in the payload data. Six mobile phone numbers and one VOIP number were found in the chat traffic. 187 private messages were found sent between persons instead of messages that are readable for all persons participating in the chat. 112 private messages were sent to participants of the website Spelpunt.nl.

In one chat message, a person states "I am no longer working, declared unfit for work". Another message states: "finger severely bruised". There are messages that speak of (head) ache. "yeah

[23] A total of 3,354 different addresses were found23. The search used was: "ngrep -I tt-merged-1to20.pcap -q 'erer' | grep -Ehrio '...erer.+' | grep -Ehrio '\/\/.+\/' ".
[24] The Dutch DPA searched the NBNS traffic on the basis of keywords such as 'confidential' and 'sick'.

you know how that shit works heh:P", followed by (from the same person): "yeah, but I woke up this morning with a terrible headache". Three other persons wrote: "it is very good but it gives you such a headache :P"; "pain! Therefore sssssst." and "headache etc.".

3.2.5 Other files

In the payload data, 16,456 video and audio files were found, such as pictures or photos in .gif (8,534) and jpeg format (5,883), but also web files (1,956 htm), 6 audio recordings (wav) and 3 movies (avi and wmw). Further investigation shows that only very few files were recorded in full. The majority of the files concerns fragments, such as the lower half of a photo.

In the DNS traffic (Domain Name System), many names selected by the owner of equipment such as PC's, laptops and iPhones were found. This could be a first name (Sanne's PC), but also a combination of a first and last name, in combination with the MAC address of the device or otherwise. In total, the Dutch DPA found 12.554 unique DNS instructions.[25] This set contains 2.205 times (18%) a (unique) first name or the combination of the first and last name, or a MAC address, in combination or otherwise with, such as PC_VAN_FAM<last name>, IPhone-van-<First name>-<Last name>, MacBook belonging to <first name> <last name>, iPod of <First name>-<Last name,> iMac belonging to <First name> <Last name>, laptop belonging to <first name> <last name> [MAC address], Computer belonging to <First name> <Last name> or iPhone-<MAC address>.[26]

In the FTP traffic 7 login names, 2 unique passwords and 1 combination of a login name, password and the relevant server were found.

And finally, 31 telephone conversations were found via Internet telephony (sip, session initiation protocol), consisting of 199 packets. The packets contain traffic data concerning telephone traffic; when and with whom (public IP address) communication by telephone took place. The content of the conversations was encrypted in all cases (by means of sip-tls).

3.3 Data on Wifi routers

The written information obtained by request shows that Google, in the period of 4 March 2008 to 6 May 2010[27] , collected data on **3,606,579 different Wifi routers.**[28] Google recorded the unique MAC address of each of these routers. Mac addresses are the unique numbers laid down in the hardware of equipment by the manufacturer, such as on memory chips and/or network card in computers, telephones, laptops or routers. MAC stands for Medium Access Control.

---

[25] These are unique queries from which the double entries have been removed. The total number of DNS [queries] with double entries was 97,220. Only the so-called 'dns query' was removed from these packets by means of a search. This results in 96,436 packets. The dataset was reduced to 14,816 packets by sorting and removing the double entries. And finally, those packets containing the word 'malformed' and the expression 'standard query response unknown' were filtered out. This leaves 12,554 packets.

[26] Versions of the first names, last names and MAC addresses anonymised by the Dutch DPA. The originals are available in the Dutch DPA's investigation file.

[27] 6 May 2010 is the most recent date found on the hard disk with data examined by the Dutch DPA. The written information obtained shows that Google *"shortly after 7 May 2010 stopped processing disks it received from Street View cars"*. Source: Letter from Google dated 8 June 2010, page 7.

[28] Email from Google dated 12 July 2010, annex dated 9 July 2010, page 1.

An example of a MAC address is: :<mac-address>. The MAC address of a Wifi router (with an Internet access point) is referred to as BSSID.[29]

Google recorded this unique number of both secure and unsecure Wifi routers. Google also recorded the signal strength of each Wifi router, in many cases the network name, and it calculated a location for each Wifi router.

On 1 January 2009, the Netherlands had 16,485,787 inhabitants, living in 7,312,579 households. There were at time over 900,000 companies in the Netherlands as well. On 1 January 2009, there were 5,73 million broadband connections in the Netherlands, or nearly 35 connections per 100 inhabitants.[30]

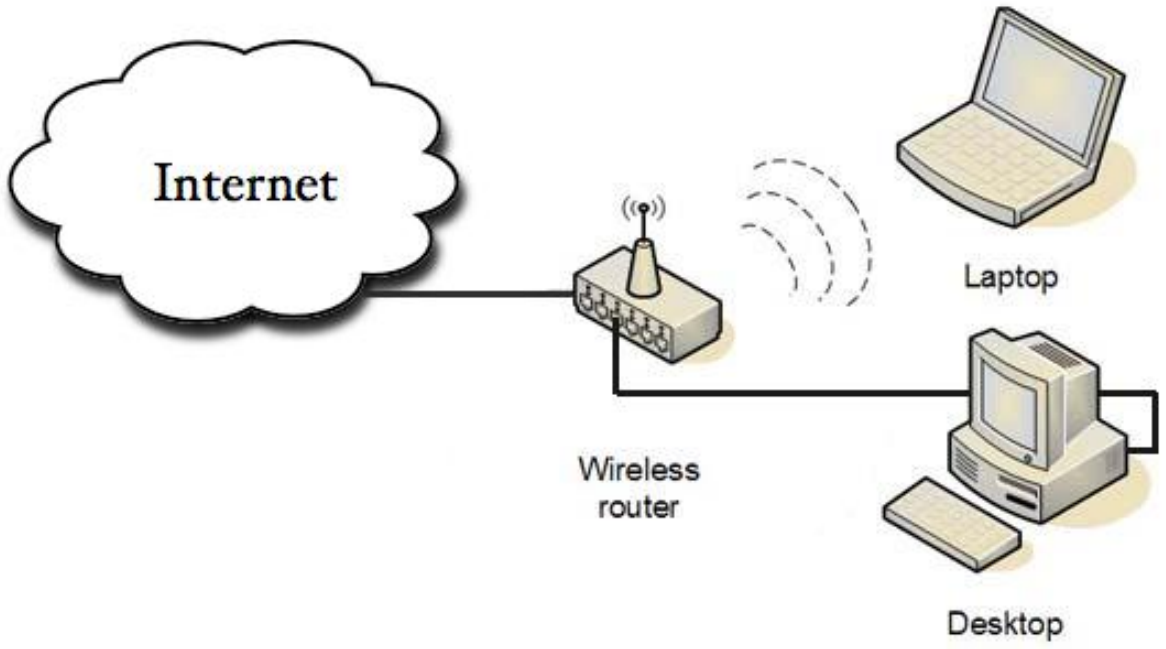
Even if it is assumed for the sake of ease that all households and companies with a broadband connection have a Wifi router,[31] Google still collected data on nearly 63% of the households and companies with a broadband connection in the Netherlands.[32]

[CONFIDENTIAL: numbers of MAC addresses of Wifi routers per province collected by Google]

Google states that it did not keep a record of when (in what month, year) the Wifi data were precisely collected or whether the Street View cars drove through each province more than once.[33]

The very large number of Wifi routers Google was able to register can be explained by the manner in which Wifi routers make their existence known. Most broadband routers also have a Wifi antenna as standard. The standard setting of the routers used most commonly in the Netherlands is 'on' for this Wifi connection, also if the user has only connected his computer(s) to the router using cables and also if the user's internal wireless communication is encrypted with WEP, WPA or WPA2. Comparable to a radio, a Wifi router continuously broadcasts its own network name and MAC address, also if no one is using the connection.

[29] The Stroz Friedberg report shows there are several address fields for storing MAC addresses. The determination of the MAC address that concerns the unique hardware address of the router (and therefor not a different Wifi-capable communication device such as an iPhone, desktop computer or laptop) depends per packet on the settings of the FromDS and the ToDS fields.
[30] Source: OPTA marktmonitor 2009, URL: http://www.opta.nl/nl/download/bijlage/?id=545 and Telecompaper, 'Dutch broadband market to reach six million in 2009', 11 June 2009, URL: http://www.telecompaper.com/news/article.aspx?cid=675788
[31] The Dutch DPA is not aware of any figures or investigations concerning the number of Wifi routers in relation to the number of broadband connections in the Netherlands as of 1 January 2009. The assumption that 100% Wifi router ownership is undoubtedly too high, but compensates for the fact that there are also households and businesses with more than 1 router and that some routers may have been replaced during the two-year period Google collected these data with its Street View cars.
[32] About 62.9% of all households and companies per 1 January 2009.
[33] E-mail from Google dated 12 July 2010, annex, page 2



A user must make a conscious choice to hide the network name of the router, i.e. to turn off the so-called beacon frame. The beacon frame continuously broadcasts short signals into the ether containing the name of the network and the MAC address of the router, the BSSID. Users that hide their SSID by turning off the beacon frame often do so for security considerations, in order to prevent neighbours or passersby from obtaining access to the network. This however does not enable the user to prevent the capturing of the network name, the MAC address and the signal strength of the router by third parties and/or their recording by third parties such as Google. The network name and the BSSID are, after all, not just contained in the beacon frame but also in several other management frames.[34]

Google states: *"As explained in the Stroz Friedberg report: "[t]he gslite program parses and stores the SSID information for all wireless networks, whether the SSID is broadcast or not." (P. 12, paragraph 52.b.). This was done because, also in the case a network is set not to broadcast SSID as part of the beacon frames, the 802.11 standard can still require that the SSID is included in other wireless data frames (that are publicly broadcast and are consequently received and stored by the gslite software), notwithstanding the network setting of "hiding" the SSID".*[35]

It is therefore not sufficient for a user who wishes to prevent third parties from capturing the SSID of his Wifi router to turn off the beacon frame. As soon as communication takes place via the antenna of the Wifi router, a third party can capture the name of the Wifi router and the signal strength from all frames.

[34] The SSID is contained in the management frame, which has various subtypes. One of those subtypes is Beacon. Others include: Association Request, Association Response, Reassociation Request, Reassociation Response, Probe Response and Probe Request. The SSID is always visible in the latter six types.
[35] Letter from Google dated 8 June 2010, annex, page 3.

Written information obtained from Google shows that in respect of each captured Wifi signal on the hard disk in the car, 18 characteristics on the Wifi routers were stored.[36] With respect to these data, 8 of these are further processed by the company.[37] These are:

- Latlon data
- MAC address
- SSID
- time per second
- time per micro second
- signal strength
- error flags
- channel

Latlon data are positioning data from GPS satellites (Global Positioning System) expressed in latitude and longitude. Google stored the GPS latitude and longitude data in 32 bit numbers, the equivalent of 7 decimal places. The latlon data concern the position of the car at the moment a Wifi signal is captured.

SSID stands for Service Set Identifier and indicates the legible network name of the Wifi router. That name can have the standard setting of the manufacturer or router reseller (such as Linksys or SpeedtouchBFCB8F), but users can also give their own self-chosen name to their wireless network, for example 'Piet Smit's network' or 'Jansenstraat 10 2 hoog'.

Signal strength is the range of the signals broadcast by the Wifi routers. The stronger the signal, the further away the signal can be captured. A range of approximately 30-40 meters within the home and 90-95 metres outside the home is maintained for the Wifi routers most commonly sold in the Netherlands, in accordance with the IEEE 802.11b and the 802.11g standard.[38]

Channel is the frequency in Megahertz at which the Wifi router broadcasts signals. Normally speaking, Wifi routers have 14 different channels.[39] A user may elect to set a different channel than the standard channel no. 6, if, for example, interference with other equipment occurs with equipment in the neighbourhood that broadcasts at the same frequency, such as other routers, microwave ovens or wireless telephones.

For each BSSID, multiple latlon data can be stored, depending on the strength of the broadcast signal and the distance of the Street View car to the relevant router. [CONFIDENTIAL: technical operation determination of the position of the Wifi routers]

[36] E-mail from Google dated 22 July 2010, annex, page 9-10. See also Stroz Friedberg, page 21, overview of the Kismet metadata.
[37] Google states that it does not use the following characteristics: carrier encoding, data rate, adapter ID, signal quality, noise level, header length, drone version, data length and capture length.
[38] Sources: About.com, „What Is the Typical Range of a Wi-Fi LAN?, URL: http://compnetworking.about.com/cs/wirelessproducts/f/wifirange.htm, Fon Wikibeta, Range/nl, URL: http://wiki.fon.com/wiki/Range/nl (last visited 27 August 2010).
[39] The Wifi standard has been defined in IEEE 802.11. In the Netherlands, most routers make use of the 802.11b and 802.11g standard, at the 2.4 Ghz range. These standards divide the spectrum into 14 channels. In Europe, channels 1 to 13 are allowed.

Google stored the data on the Wifi routers that were collected by the Street View cars on servers in the United States. These servers contain the BSSIDs (router MAC addresses), in many cases the SSID (network name) of the router, RSSI (signal strength) and the broadcast channel used by the router. In the Google geolocation server (hereinafter: Google GLS) the MAC address is stored in combination with the location of the Wifi router that had been calculated as accurately as possible. [CONFIDENTIAL: data used by Google to determine the location of the Wifi routers]

Google did not just record the MAC addresses of the Wifi routers (the BSSIDs) on the hard disks of the Street View cars, but also all other MAC addresses of equipment that communicated with the Wifi router at the moment the car drove by, such as desktop computers, laptops and other equipment suitable for Wifi communication such as iPhones. The investigation shows that Google stored all these other MAC addresses. This is evident from a calculation provided by Google of the number of MAC addresses that are BSSIDs of Wifi routers, taking into account the nature of the MAC address, either the recipient or the sender (see footnote 29). The collection and processing of these 'other' MAC addresses falls outside the scope of this investigation.

Although Google ceased the collection of Wifi data by means of the Street View cars shortly after 7 May 2010, the company still collects new BSSIDs each day via the numerous Internet users that use the Google GLS by means of Wifi-enabled devices. When using the service, users of modern smartphones often transmit the exact GPS data of their own location. This is evident from specific Google privacy statements intended for users of its geolocation services. See paragraph 3.4 of this report. This ongoing collection of MAC addresses of Wifi routers and refining of the geolocation of the Wifi routers falls outside the scope of this investigation.

Google offers third parties free access to the data in the Google GLS via the so-called *Google Geo Location API*. A third party such as Twitter uses this according to Google. Twitter users can automatically add a precise location, using Google GLS, to their messages in order to provide them with context.[40] The browsers Firefox and Chrome have also set the Google GLS as the default geolocation server, so that third parties can retrieve the location of the user via the browser.

Google argues that no detailed information is provided via the network through the Google GLS and the API, but only a location: *"The GLS and the location-based Google services do not disclose the SSID and MAC addresses or other network information related to a Wifi access point (a user device can, however, detect this information independently) (...)*[41]

Anyone can use this geolocation service, either by installing software made available free of charge by Google or by typing in a MAC address on the website of a third party that has installed this software on its server. Google GLS then shows a calculated location for the relevant Wifi router.

The Dutch DPA searched the *payload data* for MAC addresses and found 75 of them probably related to print servers.[42] Such print servers form part of the Wifi network, and are used, for example, if two persons within a family want to use the same printer via the wireless network.

---

[40] Google, 'Data collected by Google cars', 27 April 2010, URL: http://googlepolicyeurope.blogspot.com/2010/04/data-collected-by-google-cars.html

[41] Letter from Google dated 8 June 2010, annex, page 4.

[42] The search used is: "ngrep -q -I tt-merged-1to20.pcap 'mac=' > tt-mac-1.out $ cat tt-mac-1.out | grep -Erhio 'x-hp-mac=.....' | sort | uniq"

The print server transmits the same signals as a Wifi router. That is why such print servers are also stored in the Google GLS as Internet access points.

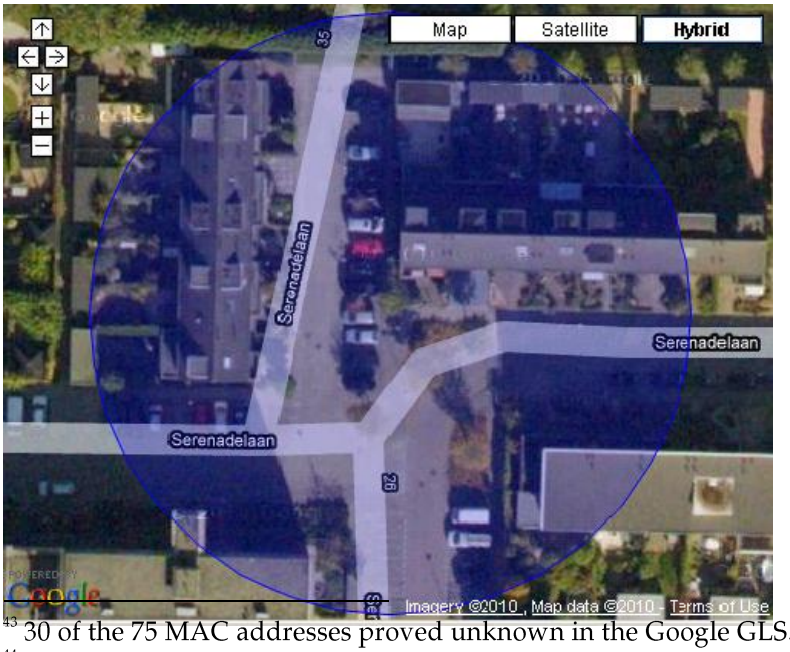
The Dutch DPA entered these 75 MAC addresses in the Google GLS and in response received calculated locations for 45 of these MAC addresses.[43] The Dutch DPA looked closely at each of these 45 locations at satellite level (via Google Maps) and examined how many home addresses could be involved. This varies from 1 farm in Overijssel to at the most 65 apartments in a residential area in Helmond. The average number of houses or apartments per MAC address is 8. The examined MAC addresses came from all over the country, from rural areas as well as cities. This set is therefore representative for the accuracy of the geolocation of the Wifi routers. See **Annex III** for a map of the examined MAC addresses.

Data obtained for MAC address <x[44]> via Google Maps API

"latitude":51.8579504
"longitude":4.5237314
"country":"The Netherlands"
"country_code":"NL"
"region":"South Holland"
"city":"Barendrecht"
"street":"Serenadelaan"
"street_number":"<x[41]>"
"postal_code":"2992"
"accuracy":36.0

According to the Google GLS, the user of the print server with this MAC address lives at the Serenadelaan in Barendrecht, between house numbers <x[41]>.

The location of this MAC address has an accuracy of 36 metres, according to the GLS.



[43] 30 of the 75 MAC addresses proved unknown in the Google GLS.
[44] The MAC addresses and house numbers have been anonymised by the Dutch DPA. The original is available in the Dutch DPA's investigation file.

The latitude and longitude can be displayed graphically via Google Maps. Zooming in on the combination of map and satellite data shows that it concerns approximately two rows of houses in a Barendrecht residential area.[45] It concerns at most 9 different self-contained houses.

Other MAC addresses in the set examined by the Dutch DPA have an accuracy of 24 metres according to the GLS.[46]

Data obtained for MAC address <x[47]> via Google Maps API

"latitude": <x[44]>
"longitude": <x[44]>
"country":"The Netherlands"
"country_code":"NL"
"region":"Overijssel"
"city":"Losser"
"street":"<x[44]>"
"street_number":"<x[44]>"
"postal_code":"7587"
"accuracy":24.0

Zooming in on the latitude and longitude in Google Maps it proves to be a single farm with stable in a rural area of Overijssel.

<Image anonymised by the Dutch DPA>

[45] The Dutch DPA used the following website: http://samy.pl/mapxss/.
[46] 13 of the 45 MAC addresses have an accuracy of 24 metres, the others have an accuracy of 36 metres.
[47] Anonymised by the Dutch DPA. The original is available in the Dutch DPA's investigation file.

Data obtained for MAC address <x[48]> via Google Maps API

"latitude":52.2108823
"longitude":5.1808856
"country":"The Netherlands"
"country_code":"NL"
"region":"North Holland"
"city":"Hilversum"
"street":"Diependaalselaan"
"street_number":"<x[45]>"
"postal_code":"1213"
"accuracy":36.0

This example shows that the accuracy of the location is not limited to rural areas. In Hilversum, there appear to be 3 houses as well that can be identified as the address of the relevant Internet access point.

The same level of accuracy, down to three houses, applies to a MAC address in Arnhem.

Data obtained for MAC address <x[49]> via Google Maps API

"latitude":51.9477717
"longitude":5.848366
"country":"The Netherlands"
"country_code":"NL"
"region":"Gelderland"
"city":"Arnhem"
"street":"<x[46]>"
"street_number":"<x[46]>"
"postal_code":"6846"
"accuracy":24.0

[48] Idem.
[49] Idem.

Based on the signal strength of the Wifi router, the precise address where the Wifi router is located can subsequently be determined. Google can do this from its own offices partly with the aid of other data held by Google, including new data on Wifi routers that Google continuously captures from the users of its geolocation services.

Third parties are also able to trace the exact address by walking around with a smartphone or laptop in the neighbourhood where the Wifi router is approximately located according to the Google GLS (an average of 8 houses). There are free applications for smartphones and there is all sorts of free software for laptops that can be used to graphically display the signal strength in the form of bars or in the shape of a radar screen. The precise device can be easily recognised and located, because the user sees the signal strength in combination with the MAC address and the network name. The closer the user gets to the precise house where the Wifi router is located, the stronger the signal strength will be.

Scanning with these apps/software is simple, consistent and reproducible.[50]

3.4 Information

On 22 April 2010, two German personal data protection authorities (the Hamburg personal data protection authority and the federal German personal data protection authority) announced that Google collected Wifi data with its Street View cars.[51] On 27 April 2010, Google announced that

[50] See the whitepaper „Converting Signal Strength Percentage to dBm Values" van het Amerikaanse bedrijf WildPackets, Inc. from 2002, URL: http://www.wildpackets.com/elements/whitepapers/Converting_Signal_Strength.pdf and the article „Modeling Signal Attenuation in IEEE 802.11Wireless LANs - Vol. 1" (July 2005) of information science student Daniel B. Faria of the American Stanford University.

[51] Bundesbeauftragte für den Datenschutz und die Informationsfreiheit, 'Google-Street-View-Fahrten werden auch zum Scannen von WLAN-Netzen genutzt', press release 22 April 2010, URL: http://www.bfdi.bund.de/cln_136/DE/Oeffentlichkeitsarbeit/Pressemitteilungen/2010/GoogleWLANScanning.html

data on Wifi networks were collected with the Street View cars.[52] At the time, the company stated that it only collected data *on* (the existence) of Wifi networks. The collection of data with respect to Wifi networks was allegedly limited to MAC addresses and SSID's (network names). The company explicitly stated that it did not collect payload data, that is, communication content data from Wifi routers. "*However, all data payload from data frames are discarded, so Google never collects the content of any communications.*"[53]

On 14 May 2010, Google announced that it had indeed recorded and stored payload data when collecting data on Wifi networks.[54] The Hamburg personal data protection authority requested access to the data collected on 5 May 2010, and Google declared that it had discovered at that precise time that content data had been stored after all. "*Nine days ago the data protection authority (DPA) in Hamburg, Germany asked to audit the WiFi data that our Street View cars collect* (...). [We said] *we did not collect payload data (information sent over the network). But it's now clear that we have been mistakenly collecting samples of payload data from open (i.e. non-password-protected) WiFi networks, even though we never used that data in any Google products.*"[55] Google furthermore stated that the company had stopped collecting (all types of) Wifi data with its Street View cars. "*In addition, given the concerns raised, we have decided that it's best to stop our Street View cars collecting WiFi network data entirely.*" Google failed to report in that connection that the company continuously collects data on Wifi routers via the users of its geolocation services.

In response to a request from the Irish personal data protection authority to destroy all *payload data* collected in Ireland, Google instructed an American security company to monitor the separation of the *payload data* from the other data on Wifi networks. The *payload data* were subsequently stored in a separate folder per country on servers in California and North Carolina in the United States. Google published its statement to that effect on its weblog on 17 May 2010.[56]

The Google web pages on Street View, including a privacy statement, and the Google general privacy statement do not contain specific information about the processing of data *via* Wifi routers.[57] In its general privacy statement Google does not inform users about the fact that it collected data on Wifi routers from the general public. Google states that telephone data of *users* such as GPS and the cell data of mobile operators are only processed when they *use* location services.

---

[52] Source: Google, 'Data collected by Google cars', 27 April 2010, URL: http://googlepolicyeurope.blogspot.com/2010/04/data-collected-by-google-cars.html. Google writes: "*This blog also addresses concerns raised by data protection authorities in Germany.*"
[53] Source: Google, 'Copy of Google's submission today to several national data protection authorities on vehicle-based collection of Wifi data for use in Google location based services', 27 April 2010, page 2-3, URL: http://www.google.com/googleblogs/pdfs/google_submission_dpas_Wifi_collection.pdf
[54] Source: Google, 'WiFi data collection: An update', 14 May 2010, URL: http://googleblog.blogspot.com/2010/05/Wifi-data-collection-update.html
[55] Idem.
[56] Source: Google, 'WiFi data collection: An update', update of 17 May 2010 with a hyperlink to http://www.google.com/press/pdf/ISEC_Letter.pdf. This is a declaration of American security company iSec Partners Inc. that the data were separated on 15 and 16 May.
[57] Source: Google, Dutch privacy statement with respect to Street View, URL: http://maps.google.com/ intl/nl/help/maps/streetview/privacy.html. General Google privacy statement in Dutch (last amended on 11 March 2009), URL: http://www.google.nl/intl/nl/privacypolicy.html

Location data – Google offers location-enabled services, such as Google Maps for mobile. If you use those services, Google may receive information about your actual location (such as GPS signals sent by a mobile device) or information that can be used to approximate a location (such as a cell ID).

In addition, Google uses a large number of separate privacy statements for various services. These clauses are also aimed at specific purchasers of certain Google products or services and not at the general public of which Google collected the data on Wifi routers. The privacy statement for Google's proprietary browser Chrome has included the following section on geolocation services since 9 May 2010:

If you use Google Chrome's location feature which allows you to share your location with any site, Google Chrome will send local network information to Google Location Services to get an estimated location. The local network information includes, depending on the capabilities of your device, information about the Wifi routers closest to you, cell ids of the cell towers closest to you, the strength of your Wifi or cell signal, and information about your device such as your device's IP address. We use the information to process the location request and to operate, support, and improve the overall quality of Google Chrome and Google Location Services. Information collected above will be anonymised and aggregated before being used by Google to develop new features or products and services, or to improve the overall quality of any of Google's other products and services.[58]

Google also has a separate privacy statement concerning the use of geolocation services in the Firefox browser. In this statement, the use of Wifi routers is referred to as follows:

Information we collect
• If you allow a website to get your location via this service, we will collect, depending on the capabilities of your device, information about the Wifi routers closest to you, cell ids of the cell towers closest to you, and the strength of your Wifi or cell signal. We use this information to return an estimated location to the Firefox browser and the Firefox browser sends the estimated location to the requesting website. For each request sent to our service, we also collect IP address, user agent information, and unique identifier of your client. We use this information to distinguish requests, not to identify you.

Uses
• We use all the information above to process the Firefox geolocation requests. We also use all the information above to operate, support, and improve the overall quality of the Google Location Service[59]*.*

Google confirms in this information, aimed at the specific target group of users of these browsers in so far as they use geolocation services, that the company has a database with the locations of Wifi routers. The company moreover indicates that it continuously processes new MAC addresses combined with signal strength.

Google's notification to the Dutch DPA on 4 February 2010 with respect to the Street View service under number m1426715 only describes the processing of photos. Google Inc., established in Mountain View, California, in the United States, was mentioned as the responsible party (as data controller).

[58] Source: Google, Chrome Privacy Notice, 4 May 2010, URL: http://www.google.com/chrome/intl/en/privacy.html. Note! This information was not yet present in the privacy statement of 9 February 2010.
[59] Source: Google, Google Location Service in Mozilla Firefox Privacy Policy, 24 April 2009, URL: http://www.google.com/privacy-lsf.html

The notification does not contain information on the processing of Wifi data. Google did not submit any other notifications to the Dutch DPA.

3.5 Purposes of processing

3.5.1. Purposes of processing payload data

According to Google, the software was accidentally set to capture the content of communication over unencrypted Wifi networks.[60] Google declares: *"Google did not use payload data in any Google product or service and did not use Wifi information collected by Street View cars for profiling activities aimed at individual end users".*[61]

3.5.2 Purposes of processing data on Wifi routers

Written information obtained shows that Google used data *on* (the existence of) Wifi routers to offer geolocation services via a geolocation server to users of mobile communication equipment (laptops or mobile phones) and for route navigation purposes. Google states that it does not intend to publish a database of MAC addresses, SSIDs and their calculated location.[62]

4. Assessment

4.1 Data Controller

Pursuant to Article 1(d) of the Wbp the data controller is *the natural person, legal entity or any other person or administrative body who, individually or jointly with others, establishes the purpose of and the means for processing personal data.*

Article 4(3) Wbp provides: *A data controller, within the meaning of the second paragraph, is prohibited from processing personal data, unless it designates a person or body in the Netherlands who/that acts on its behalf in accordance with the provisions of this act. Such a person will be deemed the data controller responsible for the application of this act and the provisions based on it.*

Google Inc., having its registered office in Mountain View, California, United States of America, established the purposes of the processing of data on Wifi routers. Google Inc. also established the means to collect these data. The written information obtained shows that Google Inc. concluded contracts with the Dutch branch of an international temping agency to hire drivers for the Street View cars. It has also become evident that Google Inc. is responsible for the equipment and software to collect and further process the data. Google Netherlands B.V. acted as local representative of Google Inc., in accordance with Article 4(3) Wbp. Consequently, Google Inc. is the responsible data controller within the meaning of Article 1(d) Wbp, as locally represented by Google Netherlands B.V.

---

[60] Google writes: *So how did this happen? Quite simply, it was a mistake. In 2006 an engineer working on an experimental WiFi project wrote a piece of code that sampled all categories of publicly broadcast WiFi data. A year later, when our mobile team started a project to collect basic WiFi network data like SSID information and MAC addresses using Google's Street View cars, they included that code in their software—although the project leaders did not want, and had no intention of using, payload data.* Source: http://googleblog.blogspot.com/2010/05/Wifi-data-collection-update.html (original blog posting van 14 May 2010).

[61] Letter from Google dated 8 June 2010, annex, page 8.

[62] Idem. Google writes: *"Google does not publish the Wifi network information it collected with the Street View cars, nor does it intend to do so."*

In its opinion on the provisional findings of the Dutch DPA, Google states that it is not the data controller, within the meaning of Article 1(d) Wbp, responsible for the collection of data *on* and *from* Wifi routers. Google has put forward a different argument for the two forms of processing.

Google argues that it is not responsible for collecting data *on* Wifi routers, because these do not concern personal data. This argument cannot be accepted, because, as will be explained below, data concerning Wifi routers are personal data.

Google furthermore states in its opinion that it also cannot be held responsible for the collection of personal data in the payload data, because the company was not aware of, and had no purpose for the collection of these data until May 2010. *"Google takes the position that it is impossible for an organisation to be responsible if said organisation does not have any purpose for processing the personal data."*[63]

Google seems to imply that a party is not responsible as long as said party does not set a clear purpose for collecting data. This would render the applicability of the law dependent on awareness or intentions. This is not the case. The legal concept of 'data controller' (as responsible party) should be interpreted objectively, not subjectively.

The issue in this context does not concern a non-recurring incident, but the systematic capturing and recording of confidential communication data from unencrypted Wifi routers during a two-year period in the Netherlands (and elsewhere). 'Not being aware' that data are collected on such a large scale and for such a long period of time is a consequence of the way in which Google has set up its operations, i.e. has set the goal and method of its data processing. By setting up its operations in this way Google knowingly and willingly accepted the possibility that data are collected without a clearly defined purpose.

Furthermore, the personal data were collected and processed during the organised activity of driving around in the Street View cars, Google hired the personnel that drove the cars and Google fitted the cars with the technical equipment to facilitate the collection and recording of *payload data*. Google set the purpose and provided the means for this activity.

4.2 Jurisdiction

Article 4(2) of the WBP provides:

2. This act applies to the processing of personal data, by or on behalf of a data controller who does not have a branch office in the European Union, using automated or other means located in the Netherlands, unless these means are only used for the transmission of personal data.

It is an established fact that Google Inc. does not have a (designated) establishment in the European Union. On Dutch territory Google Inc. collected data with respect to Wifi networks by making use of cars driving around in the Netherlands, and other means, including equipment. This concerns "automated or non-automated means" that can be used to process personal data, within the meaning of Article 4(2) Wbp. The Wbp therefore applies to the processing of personal data *on* and *from* Wifi routers by Google Inc. in the United States, as represented locally by Google Netherlands B.V.

[63] Google position, paragraph 77, page 26.

4.3 Personal data

According to Article 1, header and under a, Wbp, personal data are defined as: *any data concerning an identified or identifiable natural person.* 'Processing' personal data has been defined in Article 1(b) Wbp and includes the capture, collection and storage of personal data.

According to the explanatory memorandum to the Wbp personal data are defined as: *"all data that can provide information about an identifiable natural person"*.

Data that do not concern a specific person, but, for example, a product or process, can sometimes also provide information about a specific person. The explanatory memorandum[64] mentions in this connection telephone numbers, car number plates and postal codes with house numbers.

It is not required that all possibilities to use data related to persons are excluded. If this is a theoretical possibility, but it is inconceivable that it will actually happen, it may be assumed that the data are not regarded as personal data. In this connection the legislator explicitly states: *"If it is possible, however, to use the data to investigate fraud, these data will be regarded as personal data. Whether or not there is an intention to use the data for this purpose is not relevant. Data are deemed personal if they can be used for such person-oriented purposes."*[65] The fact that data can be personal if its processing is not intended to identify the data subjects, is also confirmed by the joint European personal data protection authorities united in the Article 29 Working Party.

Furthermore, the person must have been identified or, at any rate, must be directly or indirectly identifiable.

Data are *directly* identifying if they apply to a person whose identity can be established without much effort.[66]

Indirectly identifiable data form another category. Indirectly identifying data concern anonymous data that can be traced back to a specific person by combining them with other data. The Explanatory Memorandum provides in this connection: *"it must be assumed that the data controller is reasonably well equipped. In specific cases the special expertise, technical facilities and the like of the data controller should, however, be taken into account.*[67]

Data that, at a certain point, is not yet "reasonably identifiable" because identification would require a disproportionate use of money and means, and therefore cannot be considered personal data, may become personal data as a result of the continued development of information technology, because identification has then become easy. The turning point will depend on the assessment of the technical possibilities in any specific case.[68]

The law provides specifically with respect to the continued development of technology*: "As regards the continued development of information technology, it should be taken into account that where in the past it may have concerned a disproportionate effort (making the data non-personal, because non-identifiable), this effort decreasing as new technologies become available. Therefore, what can reasonably be*

[64] Explanatory memorandum to the Wbp, page 46-47.
[65] Idem, page 47. WP29 Opinion 4/2007 on the concept of personal data, page 18. See also: District Court of The Hague, 2 April 2010, LJN BM1481.
[66] Explanatory memorandum to the Wbp, page 48.
[67] Idem, page 49. See also: the Dutch DPA 30 June 2003, z2003-0477.
[68] Idem, page 1-2.

considered data that cannot be traced back to a person, at a certain state-of-the-art level, may as yet become personal data as technology advances, because of the increased possibilities of personal identification."[69]

4.3.1 Data on Wifi routers

It is an established fact that Google collected the following data on Wifi routers: the worldwide unique identifying number of the hardware (BSSID), the calculated location, the signal strength and, in many cases, the SSID (the network name). The combination of the MAC address and the calculated location was stored in the Google GLS.

The collected MAC addresses of Wifi routers, combined with the calculated location, are, in this context, personal data within the meaning of Article 1(a) Wbp, because they are unique identifying numbers that can be used to identify individual owners. It is important in this connection that MAC addresses distinguish unique hardware[70] and that said hardware, within the context of geolocation services, are inextricably linked to the location of its owner.[71]

After all, not only 3,606,579 MAC addresses were included in the specific application of the Google GLS, but also a calculated approximation of the location of each router. This report contains examples of the location of MAC addresses in the Google GLS. The accuracy of the approximation varies between 24 and 36 metres in the examined MAC addresses, with a range of on average 8 apartments or self-contained houses.

It is an established fact that Google processes the signal strength of Wifi routers, as is evident from the Stroz Friedberg report. It is evident from the specific privacy statements for users of the geolocation services in the Chrome and Firefox browsers that Google continuously captures new information on new and already located Wifi routers via the users of its geolocation services, including the BSSID and the signal strength. As a result, the location approximations in the Google GLS are becoming increasingly accurate.

Google itself, pre-eminently, has the means to identify the individual owners of the Wifi routers. The effort Google would have to make to identify the homeowners with the aid of the data it already holds, and the continuous measurements, cannot be considered to be disproportionate, in particular in view of the fact that it is precisely Google itself that has access to an enormous potential of capable technicians and computer scientists. Google can perform this identification from its own offices on the basis of data the company already holds and receives on a daily basis.

Parties other than Google can also determine where a Wifi router is located with the aid of Google GLS. Once the location has been roughly determined using Google GLS (an average of 8 houses), anyone can easily trace the specific address of the Wifi router with the MAC address concerned, by measuring the signal strength of the available MAC addresses in the vicinity of the

[69] Idem, page 49.
[70] The MAC address has been burned into the hardware and can never be changed. It is possible, however, to emulate (imitate) a different MAC address using certain software. This not easy. The process is called cloning or spoofing.
[71] There is case law in the Netherlands concerning a specific Internet user who was found during a police investigation on the basis of his MAC address. See: District Court of The Hague, 2 April 2010, LJN BM1481, URL: http://jure.nl/bm1481. A stalker could stand in front of an apartment block or in the part of the street where the MAC address has been localised and will be able to easily identify his or her target.

indicated houses with a laptop or Wifi-enabled mobile phone using freely available software or applications, until the correct owner has been found.

Google nevertheless states in its response to the provisional findings that it would be "impossible" to identify individual Wifi routers, because the location approximation of a Wifi router in the GLS concerns several addresses in a single street, and not a specific house address. *"It is impossible for anyone who only has an SSID, BSSID or wireless MAC address combined with the location information calculated by Google to identify the relevant wireless device (or, as regards the SSID, devices)."*

As pointed out above, this is factually incorrect. Google furthermore fails to acknowledge the fact that, pursuant to the Explanatory Memorandum, the traceability criterion must be based on an "*adequately equipped data controller*". *The special expertise, technical facilities and the like of the data controller should be taken into account in specific cases, however.*"[72]

The circumstance that in some cases a MAC address combined with the calculated location cannot be traced back to an individual, or cannot be traced back without a disproportionate effort, does not alter the above.

Google furthermore states: *"Even if a specific house can be identified, as was recently illustrated by a recent decision of the Rotterdam District Court in B/Schoonderwoerd, it is not necessarily the case that, using information with which a specific house has been identified, the person or persons who live in said house has been identified: only the identity of the building has been traced and that does not constitute personal data."*[73]

The court gave a summary consideration in the abovementioned case B/Schoonderwoerd, perhaps superfluously, that the value history of a house cannot be considered, in principle, to be personal data. The case does not concern a similar situation. The applicant was not admissible in the main action, because he had notified the wrong private company. In this summary consideration, the Court did not deal with the question of whether the relevant person can be identified *indirectly*. Moreover, the legislator had already determined that the combination of a postal code and a house number constitutes personal data in many cases. *"Telephone numbers, number plates of cars and postal codes with house numbers should, in certain circumstances, be designated as personal data."*[74]

It has been established, with respect to network names, that equipment can have a name chosen by its owner, that can be directly identifying (first and last name in combination with the calculated location) or indirectly identifying (the combination of a network name with the unique MAC address combined with the calculated location). Standard network names such as "SpeedtouchBFCB8F" are therefore personal data as well within the meaning of Article 1(a) Wbp in the combination with MAC addresses and calculated locations processed by Google.

4.3.2 Payload data

It is an established fact that Google collected communication content data from unencrypted Wifi routers throughout the Netherlands on a large scale. The payload data contains nearly 6 million

[72] Explanatory memorandum to the Wbp, page 48. See also: the Dutch DPA 30 June 2003, z2003-0477.
[73] Google position, paragraph 107, page 34.
[74] Explanatory memorandum to the Wbp, page 46-47.

analysable packets. Many packets contain personal data as defined in Article 1(a) Wbp, such as names, email addresses, mobile phone numbers and names of computers, laptops and iPhones which include first and/or last names. Persons can be easily identified on the basis of these data by looking up the name in a telephone book or by calling or sending an email.

Google collected a total of 16,640 email addresses in the Netherlands. Email addresses are in most cases personal data because they pertain and can be traced back to a natural person, also if this person uses a business email address, such as Jan@companyname.nl.

In the DNS traffic (requests sent to servers that translate domain names into IP addresses) 2,205 instances of a (unique) first name, a unique combination of a first name and last name, in combination with a MAC address or otherwise, were found. These data are also personal data, in particular in combination with the fact that Google stored the exact location of the car with respect to each *packet.* It should be taken into account here as well that Google is pre-eminently able to identify these persons.

The circle of possible persons with the same name is very small and those involved are therefore identifiable without a disproportionate effort, because the exact location data of the car were stored in respect of each *packet*.

Often, several *packets* of a single user were captured. These can be linked to each other because they contain an (internal) IP address and/or an email address. In particular in view of its broad range of online services and its enormous potential of highly capable technicians and computer scientists, Google is pre-eminently able to identify a person on the basis of such data without a disproportionate effort. In their opinion on personal data, the joint European personal data protection authorities, united in the "Article 29 Working Party", have established that personal data concerns identifiable persons if the person can be "reasonably" indentified using "all means". *"If, taking into account "all means in respect of which it may be assumed that they can reasonably be deployed by the person responsible for the processing or by any other person", that possibility does not exist or is negligible, the person may not be considered "identifiable" and the information will not be seen as "personal data".*[75]

Google states that the payload data were not personal data at the time Google collected and recorded the data. *"These data do not constitute personal data if they are stored on the Street View hard disks, because they are stored in a proprietary binary format and individual persons cannot be identified, directly or indirectly, on the basis of these data. If the payload data were to be converted into a format that is legible for persons, they may contain small amounts of fragmented personal data in exceptional cases."*[76]

The fact that Google stored the data in binary format does not mean that these data are consequently unreadable. Binary means that the information is stored very efficiently in the form of zeroes and ones. It is an established fact that Google has the means to translate this binary information into a human-readable format, and that Google also used these means (the codex software) from the first moment that data on Wifi networks were collected with the cars in order to be able to process the data on the Wifi networks. This means that the binary format concerns a form of pseudo anonymisation, which does not prejudice the Dutch DPA's assessment.

[75] WP29 Advice 4/2007 on the concept of personal data, adopted on 20 June 2007, URL Dutch version http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_nl.pdf
[76] Google position, paragraph 7, ii, at a, page 5.

The issue of whether Google has ever used the data or would ever wish to use the data for a specific purpose, and whether such processing could cause damage to the data subject, is not relevant in this case. Pursuant to the European Court of Human Rights in the case of Amann[77] , the mere collection and storage of personal data constitutes a violation of the privacy and correspondence protected by Article 8 of the European Convention on Human Rights.

The Court held as follows:

"The Court reiterates that the storing by a public authority of information relating to an individual's private life amounts to an interference within the meaning of Article 8. The subsequent use of the stored information has no bearing on that finding (see, mutatis mutandis, the Leander judgment cited above, p. 22, § 48, and the Kopp judgment cited above, p. 540, § 53)."[78]

Special personal data

The data collected by Google also contain *special* personal data as defined in Article 16 Wbp, such as a comprehensive psychological report on a boy with a serious reading disability. The report contains his first name, his last name, his address details and his date of birth.

Communication data

The investigation shows that Google captured a large quantity of communication content data, such as the combination of login name, password and email/FTP server, payment data with respect to investments or an overview of email messages received.

Although the Wbp does not apply a separate definition of *sensitive* communication data, there is no doubt in society that such personal data should be treated with the strictest confidentiality.

Google disputes this. *"Google does not accept that unencrypted communication data that are sent throughout the Netherlands publicly via unencrypted Wifi networks, are 'sensitive communication' data, as the Dutch DPA states in its report."*[79] Google states that the data are not sensitive because radio signals in the ether do not enjoy protection under criminal law: *"radio signals that can be received in a simple manner with a lawfully obtained reception device are specifically excluded in the article that prohibits the interception of personal electronic messages."*[80]

The confidentiality of communication arises directly from international conventions such as the European Convention on Human Rights and the Charter of Fundamental Rights of the European Union. Case law of the European Court of Human Rights (ECtHR) shows that both the content of, and information on electronic communication, enjoy protection on the ground of Article 8(1) European Convention on Human Rights (ECHR).[81]

Although the term 'correspondence' in the ECHR does not explicitly apply to electronic communication, the ECtHR deemed the protection to apply successively to telephone

[77] European Court of Human Rights, February 2000, (*Amann*).
[78] Idem, paragraph 70, see also 68-69.
[79] Google position, paragraph 70, page 24.
[80] Idem.
[81] In HR 9 January 1987 (*Mother receiving benefits who was spied on*), the Supreme Court explicitly acknowledged that Article 8 ECHR has horizontal effect, which means that it is a right citizens can invoke against other citizens and companies.

conversations (Klass), telephone numbers (Malone) and secret recording of conversations with bugging equipment (P.G. & J.H.).[82]

In the Huvig case, the Court considered that bugging telephone conversations constitutes a serious violation of privacy and correspondence: "*Tapping and other forms of interception of telephone conversations represent a serious interference with private life and correspondence and must accordingly be based on a "law" that is particularly precise. It is essential to have clear, detailed rules on the subject, especially as the technology available for use is continually becoming more sophisticated.*"[83]

The Court explicitly extended the protection to email correspondence and Internet use in the Copland case.[84] The Court considered: "*Accordingly, the Court considers that the collection and storage of personal information relating to the applicant's telephone, as well as to her e-mail and Internet usage, without her knowledge, amounted to an interference with her right to respect for her private life and correspondence within the meaning of Article 8.*"

It is quite revealing that in the Charter of Fundamental Rights of the European Union the term 'correspondence' has been replaced with the much broader term of 'communication' in Article 7.

The fact that communication data are generally considered to be highly confidential data is also evident from the fact that the European directive *on the obligation to retain telecommunication traffic* has explicitly exempted communication content data, such as the subject line of email and web traffic, from the obligation to retain. Preventatively storing these data of all citizens for the purposes of investigation would constitute an unacceptable violation of the fundamental right to protection of private and family life and correspondence (Article 8 ECHR).

The overview of emails captured shows that traffic data on telecommunications can paint a surprisingly detailed picture of an individual's private life. Based on the timing of the emails, the addresses of the senders and, in particular the subject line, it is possible to reconstruct an accurate picture of moments in the life of the data subject, his interests and his career development.

Another example of such sensitive communication data are some searches performed with the Google search engine. Searches such as "how many weeks pregnant", "green baby poo" and "home care store TWELLO" refer to urgent medical questions. The fact that someone apparently intends to rent a wedding dress provides an indication of wealth.

A search does not necessarily need to relate to a need of the person who performed it. The data subject could also search for information on behalf of a family member, acquaintance or for example in connection with a school assignment. Searches nevertheless do *relate to* a specific data subject, because they provide a clear indication of this person's interests and needs. The aim of *behavioural advertising* is, for example, to show targeted advertising to the data subjects on the basis of their Internet behaviour. It is not relevant in this connection whether the data subject purchased the product displayed or service recommended for himself or for a third party.

That same sensitivity applies to the websites most recently visited, the *referrers*.

---

[82] ECtHR 6 September 1978 (*Klass*), ECtHR 2 August 1984 (*Malone*), NJ 1988, 534, ECtHR 25 September 2001 *(P.G. and J.H. v. United Kingdom),* no. 44787/98.
[83] ECtHR 24 April 1990, NJ 1991, 523 (*Huvig*) § 32.
[84] ECtHR 3 April 2007 *(Copland v. United Kingdom).*

A search for porn can seriously embarrass the data subject. It was clear in the case examined that a person visited several website with unmistakable pornographic content and to have viewed specific files there. This data subject is also identifiable, because a location of the Wifi router can be calculated in respect of each search.

The combination of the login name and password for a email server means that the content of all incoming and outgoing email of a data subject, who can be identified at that time, can be read. Google stored 56 such combinations. In theory, this means that Google is able to obtain insight into copies of sent emails on the server, to the extent a data subject has stored them, and read and copy these email messages. The same applies to contact details in an address book.

4.4 Information

Article 33 Wbp provides that the data controller must inform the data subject of its identity and the purposes of the processing *before the data are obtained*. In this connection, the data controller is required to provide further information, pursuant to Article 33(3) Wbp, *to the extent such is necessary, in view of the nature of the data, the circumstances in which they are obtained or the use that is made of them, to guarantee proper and careful processing towards the data subject.* (underlining by the Dutch DPA).

Article 34 Wbp provides: *If personal data are obtained in a manner other than meant in Article 33, the data controller will inform the data subject as referred to in the second and third paragraph* (i.e. of the identity, purpose and further information within the meaning of Article 33 Wbp, explanation Dutch DPA), *unless it is already aware of this information: a. at the time the data pertaining to him are recorded, or b. if the data are intended to be provided to a third party, at the latest at the moment the data are first provided.*

In this connection, the controller is not required to provide information, pursuant to Article 34(4) Wbp, *if providing the information to the data subject proves to be impossible or requires a disproportionate effort. In such cases, the data controller will record the origin of the data.*

Article 27(1) Wbp provides: *a fully or partially automated processing of personal data intended for the realisation of an purpose or several related purposes, is notified to the Dutch DPA or the official before processing is commenced.*

4.4.1 Data on Wifi routers

Google used its Street View cars to collect data on (the existence of) more than 3.6 million Wifi routers in the Netherlands. It is an established fact that Google did not inform the citizens of the Netherlands who own a Wifi router in advance or during the collection of the processing of these personal data. Google has not yet incorporated the processing of Wifi data in its general privacy statement nor in the specific explanations to Maps and Street View. The only information made available by Google concerns the continuous collection of Wifi data and is aimed at the specific target group of users of the Chrome and Firefox users to the extent they make use of the geolocation services. This information is not aimed at data subjects whose Wifi routers have been mapped by Google.

It is also an established fact that Google did not notify the Dutch DPA of the processing of personal data. Google notified the Dutch DPA on 4 February 2010 of the processing of

photographic data for the purpose of the Street View service, but this notification does not contain any information on the collection of data on Wifi routers.

The legislator intended that the obligation to notify the personal data protection authority of the intended processing of personal data should serve to make the processing of personal data transparent, so that data subjects are able to check the purposes of certain instances of data processing.

"Notification is intended to promote transparency of data processing. The freedom of action of each person to process data for purposes that are in themselves legitimate, is restricted by the freedom of the data subject not to be subjected unnecessarily to processing of personal data that pertain to him, which freedom is guaranteed by the constitution. This leads, on the one hand, to the substantive rule that the interests of the data controller and the data subject have to be weighed against each other, and, on the other hand, to the procedural rule that the consideration has to be verifiable."[85]

In its opinion concerning the provisional findings, Google indicated that Article 33 Wbp does not apply to the collection of data, but rather Article 34 Wbp. At the same time, the company invokes the exception contained in Article 34(4) Wbp of not having to inform owners of Wifi routers. "Apart from the fact that Google is of the opinion that Wifi network information does not constitute personal data, Google does not acknowledge that it violated a statutory obligation to provide information. (..) [Google allegedly] *would have to make a disproportionate effort to notify all data subjects (whoever they may be) that are linked to the millions of SSIDs, BSSIDs and MAC addresses that were transmitted publicly via Wifi networks in every region of the Netherlands, of the fact that Google was collecting these data. Google was allegedly exempted from the obligation to provide information pursuant to Article 34(4) Wbp."*[86]

Google collected the data on Wifi routers itself, at the front door of the data subjects, by means of the Street View cars driving through the neighbourhoods. There is *de facto* secret collection of data, because Google never made it known that the Street View cars also collected data on Wifi routers in addition to photos of houses. Google consequently did not allow the data subjects to adjust their behaviour, for example by switching off their router at the time Street View cars passed by or to otherwise object to data pertaining to them being included in the Google GLS.

The Explanatory Memorandum to the Wbp gives the example of camera supervision. Article 34 Wbp applies if such occurs secretly.

The collection of data with the aid of video cameras can be classified under Article 33 if it does not concern secret observation. If the data subject is aware of the presence of cameras and he also knows for which purpose these are used, he will have the option of avoiding this. If he does not do so, it can be argued that he consciously made his personal data available for the relevant purpose.[87]

The fact that Google secretly collected the data, does not release the company from the obligation, pursuant to Article 34(1)(a) Wbp, to inform the data subjects at the moment the data are recorded.

The Dutch DPA does not share the qualification that informing the data subjects would require a disproportionate effort. Nor does the Dutch DPA understand how this effort would be different

[85] Explanatory memorandum to the Wbp, page 132-133.
[86] Google position, paragraph 124, page 40.
[87] Explanatory memorandum to the Wbp, page 155-157.

from the manner in which Google for example informed data subjects in Germany of its intention to make photos of houses at certain times for its Street View service, later followed by information on the intention to publish these photos, and the possibilities of the data subjects to oppose publication of data pertaining to them.

There were sufficient means available in the Netherlands as well to inform the data subjects of the data collection, as announced in routes via websites, press releases, targeted advertisements and identifiability of the relevant cars. There are also sufficient means available to as yet inform the data subjects adequately of the data collection. In view of the enormous extent of the data collection, the interest of the data subjects in knowing that data pertaining to them are being collected and for which purpose, and how they can avoid this, and in view of the possibilities available to Google in this context to provide the public with information, such an implementation of the obligation to provide information would certainly not be disproportionate in this connection.

Google acted contrary to Article 34 Wbp by failing to comply with its obligation to inform data subjects in the Netherlands of the collection and processing of Wifi router data. Google acted contrary to Article 27 Wbp in conjunction with Article 28 Wbp by not notifying the Dutch DPA of the intended Processing.

4.4.2 Payload data

Google collected nearly 30 gigabyte of communication content data in the Netherlands during a period of nearly two years. It is an established fact that Google informed neither the data subjects nor the Dutch DPA in advance of the collection of this *payload data*. After an initial denial on 27 April 2010, the company later announced on 14 May 2010 that *payload data* had been collected by accident. Google failed to comply with the obligation to provide information pursuant to Article 34 Wbp at any rate before 27 April 2010. Google has not informed the Dutch public after that date either of the processing of data pertaining to it, nor did it comply with the obligation to notify the Dutch DPA of the intended processing.

4.5 Purpose and ground

In accordance with Article 7 Wbp, the data controller must have *well-defined, explicitly described and legitimate purposes* for collecting the data.

Pursuant to Article 8 Wbp, a data controller must have a ground for processing data. In the present case, two grounds for justification apply: consent and necessity.

Article 8, header and under a and f respectively, Wbp provides: *Personal data may only be processed if: (...)*

a. the data subject has granted its explicit consent for the processing;

f. the processing of data is necessary to look after the legitimate interest of the data controller or a third party to whom the data are provided, unless the interest or the fundamental rights and liberties of the data subject, in particular the right to the protection of the private and family life, prevails.

4.5.1 Data on Wifi routers

As regards the recording of data *on* Wifi routers and other communication equipment, Google did not announce the purpose of the processing until after commotion had arisen in the media concerning the fact that Google was collecting data on Wifi networks with the Street View cars. Google states that it needs the MAC addresses and the calculated location to be able offer a location service on the Internet, via the Google GLS.

Google did not mention a well-defined or explicitly described purpose for registering the SSID (the network name of the Wifi routers). This could be a standard name assigned by the manufacturer, but also a combination of a first and last name or address with a house number assigned by the owner. The combination of the SSID, BSSID and the calculated location constitutes personal data. The lack of information on the precise purposes for processing the SSID's means that Google is acting contrary to Article 7 Wbp.

Google did not request approval from the data subject for collecting the BSSIDs, the SSIDs and the signal strength. The only other applicable ground for this data processing is contained in Article 8(f) Wbp. As regards this ground for justification, the own justified interest must be weighed against the rights and freedoms of the data subjects, in particular the protection of their private and family life. A data controller who wishes to invoke this foundation has to demonstrate that the processing is necessary, and that the purpose cannot be achieved otherwise or with less drastic means. This consideration of proportionality and subsidiarity is followed by a second consideration, whereby the interests of the data subject will have an independent value when compared with the interests of the data controller. In the event the interests of the data subject in protection of his private and family life are decisive, the data controller is obliged to refrain from proceeding with the processing of data.

As regards the collection of MAC addresses of Wifi routers combined with the calculated location data and the signal strength for the purpose of being able to offer a new product, the GLS, for which there is a demand in the market, Article 8(f) Wpb can be invoked, the need to collect these data for a legitimate interest, while Google provides for a number of guarantees to prevent the interest of the data subjects in the protection of their private and family life from prevailing.

In its opinion, Google states that applicability of the Wbp would cause an obstruction to the development of new cartographical services. *"If the Dutch DPA and/or the Article 29 Working Party designates Wifi network information commonly collected by players on the digital maps market as personal data (combined with GPS coordinates or otherwise), this could have far-reaching consequences for this large and fast-growing line of business. (...) Such a classification would obstruct the future growth of the market for digital maps in EMEA and prevent the development of new map products."*[88] Google does not substantiate this argument. Google does not indicate which potential obstructions or impediments could be caused by the personal data qualification.

The first and foremost consideration is that the data controllers can have legitimate interests in the development of new and innovative services on the Internet for which there is a demand. In accordance with Article 8(f) Wbp, the impact of these services on the private and family life of the data subjects should be taken into account in this connection. Data controllers should embed guarantees to prevent disproportionate damage. Embedding guarantees (privacy by design) does

[88] Google position, paragraph 117, page 37-38.

not constitute a potential obstruction to the future growth or prevention of the development of new map products.

In the present case, Google did not inform the data subjects of the processing of data on Wifi routers. Careful processing of personal data requires that data subjects are actively informed of the recording of their personal data and the specific purpose for which these data are collected and processed.

Google should take into account the risk of serious violations of private and family life as a result of the traceability of Internet users via the MAC address of their Wifi router combined with the localisation that is becoming ever more accurate. In special cases, for example if there is a risk of *stalking* if a data subject moves to a new place of residence, but can be traced there by means of the same MAC address, a possibility of objection should be provided for, as referred to in Article 40 Wbp. It is important in that connection that the guarantees are continuously adjusted to the advancing technical possibilities to identify the data subjects.

The method of data collection Google has so far applied does not sufficiently show a balanced consideration of interests (principle of proportionality). It should also be taken into account in this connection that if the justified interest of the data controller can be served otherwise or with less drastic means, the processing of data will not be permitted (principle of subsidiarity). Google can leave the SSIDs (network names) out of the Google GLS and limit itself to the combination of the BSSIDs and their calculated locations to offer location approximation services. The lack of a need to process the SSIDs means that there is no justified purpose for their collection.

Google acted contrary to Article 7 (purpose) and Article 8 (ground) Wbp by collecting data on Wifi routers. A well-defined, explicitly described and justified purpose for collecting the SSID's is lacking. A ground for collecting the SSIDs, in the present case consent or a justified interest, is also lacking.

4.5.2 Payload data

As regards the collection of the payload data, Google answered the questions of the Dutch DPA and publicly acknowledged that it does not have a (justified) interest. After discovery, Google put a worldwide stop to the collection and processing of communication content data.

In view of the lack of a justified purpose on the part of Google for collecting confidential communication data and the confidential nature of some collected personal data, Google cannot rely on a justified interest in collecting the data, i.e. a ground contained in Article 8(f) Wbp. Consent was not given either, let alone the *explicit* consent required on the ground of Article 16 Wbp for breaching the prohibition on the processing of *special* personal data. Google therefore lacks any ground for processing the data.

Google has acted contrary to the obligation contained in Article 7 of having a well-defined, explicitly described purpose for the data collection, by collecting the payload data. Google acted contrary to Article 8 Wbp as it does not have a ground, in the present case consent or a justified interest. When collecting the *special* personal data, Google also acted contrary to Article 16 Wbp: the prohibition on processing *special* personal data.

4.6 Due care

Article 16 Wbp provides that the personal data must be processed in accordance with the law and in a proper and careful manner.

All instances of wrongfulness identified in the present context have an effect on the assessment on the ground of Article 6 Wbp.

As regards the collection *on* Wifi routers, it is an established fact that Google mapped over 3.6 million unique Wifi routers in the Netherlands. Google collected these personal data secretly, without providing any information to the data subjects concerning the purpose of the collection, neither directly nor indirectly, by means of a notification to the Dutch DPA. This lack of information means that Google is unable to guarantee the proper and careful processing of personal data.

Applying privacy by design is part of a proper and careful approach in order to prevent data from being processed that are not needed for the (justified) purpose. Google should have taken measures to guarantee that, as a result of its business operation or without complying with the requirements of the law, personal data are not collected on such a scale.

The fact that Google does not have a justified purpose nor a ground for collecting and processing *payload data* means that Google acted contrary to the law. It is, at the very least, evidence of carelessness that Google drove around the Netherlands for two years without – as indicated by the company – noticing that so much *payload data* was being stored. By setting the software differently (*false or true*) in advance, Google could have prevented any *payload data* from having been collected at all. Google states that the collection of the *payload data* concerned a mistake of an individual programmer: *"Quite simply, it was a mistake. In 2006 an engineer working on an experimental WiFi project wrote a piece of code that sampled all categories of publicly broadcast WiFi data."*

In its opinion, Google argues: *"Google is busy implementing considerable remedial measures and is improving its working methods in order to prevent any future reoccurrences of incidents whereby payload data are accidentally collected."*[89]

Whatever the case may be, in the present case Google clearly failed to implement any, or, at any rate, sufficient, measures to prevent decisions from being taken and implemented at employee level, intentionally or otherwise, that entail major privacy risks.

As a consequence of the manner in which Google operated as a company in the present case, the privacy rights of large groups of data subjects were violated. In particular, in view of the nature of Google's business activities and the specific expertise that is available in this connection, it may be expected of Google that it operates with due observance of the requirements set by the Wbp and that Google takes account, prior to the performance of activities, of the risks and consequences of the company's mode of operation with respect to the rights of the individual data subjects. The wrongfulness of Google's approach with respect to the collection of Wifi data cannot be attributed to an individual programmer, but rather to the company that takes this risk.

[89] Google position, paragraph 133, page 42.

This careless mode of operation means that Google acted contrary to Article 6 Wbp.

5. Conclusion

Google drove around the Netherlands in Street View cars from March 2008 to May 2010. These cars collected data *on* all Wifi routers that were switched on at the moment the car passed by and collected communication content *from* unencrypted Wifi routers.

Google literally mapped the Wifi router of 63% of the households and businesses in the Netherlands with a broadband connection, in the sense that the unique identifying MAC address of the router was recorded in combination with the calculated location. Even if the owner of the Wifi router had encrypted his communication, Google nevertheless collected and stored information on this router. The combination of a MAC address of a Wifi router with the calculated location constitutes personal data in the present context.

By collecting data on Wifi routers, Google acted contrary to Article 7 (purpose) and Article 8 Wbp (ground). There is no well-defined, explicitly described and justified purpose for collecting SSID's (network names). A legitimate ground for collecting the SSID's, in the present case consent or a justified interest, is also lacking.

Google acted contrary to Article 34 Wbp by ignoring its obligation to inform data subjects in the Netherlands of the collection and processing of Wifi router data. Google acted contrary to Article 27 Wbp in conjunction with Article 28 Wbp by not notifying the Dutch DPA of the intended processing.

Google also collected data *from* unencrypted Wifi networks, the so-called payload data, on a large scale in the Netherlands. During a two-year period, the company stored nearly 30 gigabytes of information, 70% of which, as it turned out, could be analysed. The investigation shows that these data are not meaningless fragments. As many as 2 to 2,500 packets of a data subject may have been captured. These packets can be linked to each other and this can be used to create an detailed picture of a data subject, such as the fact that someone purchases a large number of shares, including the person's name, email address and bank account number.

Sometimes, even 1 packet can contain a deluge of *special* personal data, such as in the case described above of a comprehensive psychological report on a minor, a boy with a serious reading disability. The report contains his first name, his last name, his address details and his date of birth.

Countless personal data were found in the analysed traffic, varying from the first and last name to the email address, and from Google searches to visiting pornographic websites.

Google has acted contrary to the obligation contained in Article 7 of having a well-defined, explicitly described and justified purpose for the data collection, by collecting the *payload data*.

Google acted contrary to Article 8 Wbp as it does not have a ground, consent in the present case, or a justified interest. When collecting the *special* personal data, Google also acted contrary to Article 16 Wbp: the prohibition on the processing of *special* personal data.

It is careless of Google that it drove around the Netherlands with its Street View cars for a period of two years without noticing that so much payload data was being stored. Google could have prevented any payload data from being stored at all by making different choices when setting up the software (*false* instead of *true*). Google can be blamed for the fact that the company took considerable privacy risks by failing to implement any, or, at any rate, sufficient, measures to prevent this large-scale violation of privacy rights. The lack of a ground, and its careless mode of operation, means that Google acted contrary to Article 6 Wbp as regards the *payload data*.