Improving privacy protection in the area of behavioural targeting

Zuiderveen Borgesius, F.J.

Publication date
2014
Document Version
Final published version

5 Data protection law, material scope

Whether data protection law applies at all to behavioural targeting is hotly debated. Many firms using behavioural targeting say they only process "anonymous" data and that data protection law, therefore, doesn't apply. For instance, the Interactive Advertising Bureau Europe states on a website on which it provides information about behavioural targeting:

> The information collected and used for this type of advertising is not personal, in that it does not identify you – the user – in the real world. No personal information, such as your name, address or email address, is used. Data about your browsing activity is collected and analysed anonymously.[815]

According to the Article 29 Working Party, however, firms usually process personal data when they use behavioural targeting. The Working Party also views behavioural targeting as personal data processing if a firm can't tie a name to the data it holds about an individual. Moreover, it's often fairly easy for a firm or another party to attach a name to the data. This chapter argues that the Working Party's view is correct. Data protection law should apply to behavioural targeting.

The chapter is structured as follows. Section 5.1 concerns the difference in scope of data protection law and the legal right to private life. Section 5.2 shows that the Working Party views behavioural targeting as personal data processing, due to the

---

[815] Interactive Advertising Bureau Europe – Youronlinechoices.

fact that a firm can use such data to "single out" a person, also when the firm can't tie a name to the data it has on an individual. Section 5.3 shows that the firm doing behavioural targeting or another party can often tie a name to data about an individual. Section 5.4 is more normative than the rest of the chapter and argues that data protection law should generally apply to behavioural targeting. Section 5.5 concerns discussions about lighter rules for pseudonymous data, triggered by the proposal for a Data Protection Regulation. Section 5.6 shows that behavioural targeting often entails the processing of special categories of data, such as data regarding health or political opinions. Section 5.7 concludes.

5.1 Difference in scope of data protection law and privacy

The scope of data protection law is both broader and narrower than the right to privacy as protected by article 8 of the European Convention on Human Rights. Data protection law has a broader scope because it applies to all personal data – any information relating to an identifiable person. The scope of data protection law isn't limited to information that is sensitive or private. Hence, data protection law applies regardless of whether there's an interference with privacy.

On the other hand, the scope of data protection law is narrower than the right to privacy in article 8 of the Convention. For instance, when somebody uses binoculars to spy on a neighbour in the bathroom, there's an interference with privacy. But data protection law doesn't apply, as the spy doesn't "process" personal data.[816] Moreover, many judgments regarding the right to private life have nothing to do with personal data processing.[817]

---

[816] See for the definition of processing article 2(b) of the Data Protection Directive, and see on non-automated processing recital 15 of the Data Protection Directive. See also Pool 2014, p. 178; p. 298. Some data processing activities are outside the Directive's scope see chapter 4, section 2.

[817] See e.g. ECtHR, X and Y v. The Netherlands, No. 8978/80, 26 March 1985, on the impossibility of instituting criminal proceedings against the perpetrator of sexual assault on a mentally handicapped girl of sixteen years old. See on the difference between the right to privacy life and the right to data protection also Opinion AG (12 December 2013), C-293/12 and C-594/12, Digital Rights Ireland Ltd, par. 61; González Fuster 2014.

The European Court of Human Rights hasn't extended the protection of article 8 of the Convention to all personal data. In other words, certain data processing activities don't infringe upon privacy according to the Court.[818] If personal data processing concerns data regarding people's private life, or if data processing is extensive, the Court is likely to find that privacy is affected.[819]

Figure 5.1 illustrates the scope of the legal right to private life (article 8 of the European Convention on Human Rights) and data protection law. The scope of the right to privacy and the right to data protection partly overlap. In many cases, data protection law and the right to privacy both apply. For instance, if a firm processes personal data about a person's private life, both legal regimes apply.

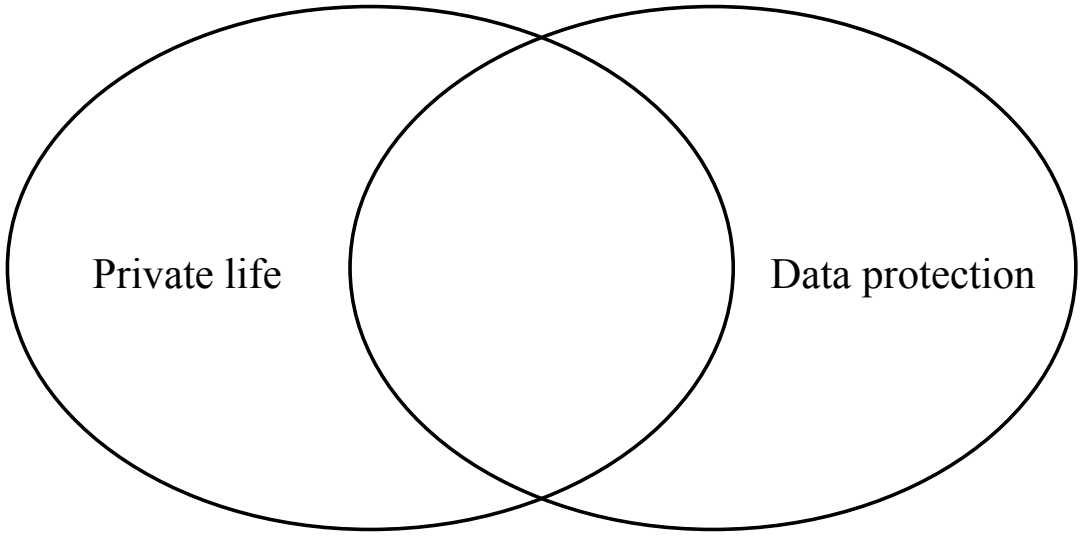
Some situations are covered by the right to private life, but not by data protection law (see the left part of the figure). Somebody who spies on his or her neighbour doesn't necessarily process personal data. There may be a privacy infringement, while data protection law doesn't apply. In certain situations data protection law applies, where the right to private life doesn't (see the right section of the figure). For instance, data protection law applies to an electronic phonebook, because it includes people's names and phone numbers, which are personal data. In this instance, the right to private life doesn't necessarily apply; being listed in the phonebook doesn't have to interfere with privacy.[820]

---

[818] See De Hert & Gutwirth 2009, p. 24; Kranenborg 2007, chapter 4 (in Dutch) and p. 311-312 (in English).
[819] See Gellert & Gutwirth 2013, p. 526.
[820] The phonebook is merely an example. There's a special regime for subscriber directories (article 12 of the e-Privacy Directive).

Figure 5.1

The scope of the legal right to privacy (article 8 of the European Convention on Human Rights) and data protection law.



5.2 Data that single out a person

Data protection law only applies if "personal data" are processed. Any operation that is performed upon personal data, such as collection, storage, or analysis, falls within the definition of "processing."[821] But do firms process "personal data" when they use behavioural targeting? The personal data definition in the Data Protection Directive reads as follows:

> "Personal data" shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.[822]

Personal data are therefore not limited to a name and address, but include all kinds of data that relate to an identifiable person. An identifiable person is someone who can be identified, directly or indirectly. The European Court of Justice has confirmed several times that information without a name can constitute personal data.[823]

In 2007 the Article 29 Working Party published a detailed opinion on the concept of personal data. The opinion is structured around four elements of the Data Protection Directive's definition of personal data: (i) any information, (ii) relating to, (iii) an

---

[821] Article 2(b) of the Data Protection Directive defines processing of personal data ("processing") as: "any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction."

[822] Article 2(a) of the Data Protection Directive, capitalisation adapted.

[823] For the Court, personal data are "any information relating to an identified *or identifiable* individual" (CJEU, C-92/09 and C-93/09, 9 November 2010, Volker und Markus Schecke and Eifert; CJEU, C-468/10 and C 469/10, ASNEF, 24 November 2011, par. 27). See also ECJ, C-101/01, Lindqvist, 6 November 2003, par 27: "identifying [people] by name *or by other means*, for instance by giving their telephone number or information regarding their working conditions and hobbies, constitutes 'the processing of personal data (…)'" (emphasis added).

identified or identifiable, and (iv) natural person. The first element is "any information." Data processed for behavioural targeting, such as digital information about a person's web browsing history, fall within the scope of "any information."[824]

The second element is "relating to." Sometimes information relates to a person because it refers to an object, such as a computer or a car. Case law of the European Court of Justice confirms that data that relate to an object can identify a person.[825] With behavioural targeting, a firm often recognises a person's device, such as a computer or a smart phone. The Working Party explains that information may relate to a person because of one of three elements: a content element, a purpose element, or a result element.[826]

Information relates to a person because of its content when it's "about" a person. The Working Party gives the example of a patient's medical file. The information in such a file is clearly about a person, regardless of the purpose or the result of using the information.[827] When a firm holds an individual but nameless profile for behavioural targeting, that information relates to a person because of its content. For the person with ID *xyz* on his or her computer, the firm might have a list of visited websites, or a list of inferred interests. The information tied to ID *xyz* is *about* that person.

Information processed for behavioural targeting may also relate to a person because of a "result" element.[828] If a firm shows an ad to a specific person, the firm treats that person differently from others. If a firm targets an ad based on data about an individual, the data relate to that person because of the "result."

---

[824] See Article 29 Working Party 2007, WP 136, p. 6-9.
[825] See section 5.3 below, on IP addresses, in CJEU, C-70/10, Scarlet v Sabam, 24 November 2011. In Lindqvist, the Court mentions a phone number as an example of information that can identify somebody. Arguably a phone relates to an object rather than to a person (ECJ, C-101/01, Lindqvist, 6 November 2003, par 27).
[826] Article 29 Working Party 2007, WP 136, p. 9-10.
[827] Article 29 Working Party 2007, WP 136, p. 10.
[828] Article 29 Working Party 2007, WP 136, p. 11.

Information can also relate to a person because of a "purpose" element.[829] A purpose element is present if a firm uses data "with the purpose to evaluate, treat in a certain way or influence the status or behaviour of an individual."[830] If an identifier that is used for behavioural targeting is primarily linked to a device, the data attached to that identifier often "relate" to a person. If a firm processes data about an individual for behavioural targeting, the processing purpose is influencing that individual, to make that person click on an ad, or buy products. Ads are targeted to a particular device because the firm hopes that the user of that device buys something. The International Working Group on Data Protection in Telecommunications notes that advertising aims to influence people rather than devices.

> While ads may well be addressed to a machine at the technical level, it is not the machine which in the end buys the proverbial beautiful pair of red shoes – it is an individual. Thus, the claim that the processing of behavioural data for marketing is directed "only" at machines in the first place may well be seen as an attempt to blur our vision as societies on the gravity of the problem, when in reality the individual and not the machine is the only instance that can make all such tracking operations a "success" for its proponents (i.e., when the red shoes are finally being bought).[831]

Some data processing activities for behavioural targeting don't concern personal data. As previously noted, in phase 3 of behavioural targeting, a firm can use data it has to construct a predictive model: *1% of people who visit websites about sports, click on*

---

[829] See also International Working Group on Data Protection in Telecommunications (Berlin Group) 2013, p. 6.
[830] Article 29 Working Party 2007, WP 136, p. 10 (emphasis original).
[831] International Working Group on Data Protection in Telecommunications (Berlin Group) 2013, p. 3. This "Berlin Group" was founded in 1983 and consists of representatives from Data Protection Authorities and other bodies of national public administrations, international organisations and scientists from around the world.

ads for running shoes, while 0.5% of random people clicks on such ads.[832] Such a predictive model doesn't consist of personal data, as it doesn't relate to a specific person.

But as soon as a predictive model is applied to an individual (phase 5), the information relates to that person because of the "purpose."[833] For instance, if a person with the cookie with ID *xyz* on his or her computer visits a website, an ad network may recognise that person (ID *xyz*) as a person who visits a lot of websites about sports. The firm has a predictive model saying that people who visit websites about sports are more likely to click on ads for running shoes. Therefore, the firm shows the person advertising for shoes. At that moment, the firm applies the predictive model to a specific person, with the purpose of influencing that person.[834] Hence, the firm processes personal data. The Working Party concludes: "the information collected in the context of behavioural advertising *relates to,* (*i.e.* is about) a person's characteristics or behaviour and it is used to influence that particular person."[835] In sum, behavioural targeting often entails the processing of "information relating to a natural person."[836]

This brings us to the third element of the personal data definition. Does behavioural targeting entail the processing of data that relate to an "identifiable" person? In other words, does a firm process data that "directly or indirectly identify" a person, if it processes data about a person, and it would be hard for anybody to tie a name to the data?

Many behavioural targeting firms say they only process "anonymous" data when they don't add a name to a person's data. Therefore, the argument goes, they don't process

---

[832] See chapter 2, section 5.
[833] Koops 2008, p. 331.
[834] See Hildebrandt et al. 2008.
[835] Article 29 Working Party 2010, WP 171, p. 9 (emphasis original).
[836] Article 2(a) of the Data Protection Directive.

personal data when using behavioural targeting.[837] The European concept of personal data has a broader scope than the US concept of "personally identifiable information." Although definitions in US statutes differ, the concept typically refers to information such as a name or a social security number.[838] Perhaps some US firms think that only information such as a name or a social security number makes a person identifiable.

Computer scientists would refer to nameless individual profiles that are used for behavioural targeting as pseudonymous data: "a pseudonym is an identifier of a subject other than one of the subject's real names."[839] A handbook on data protection law summarises: "[i]n contrast to anonymised data, pseudonymised data are personal data."[840] The Working Party concurs.[841]

The Directive's personal data definition mentions an "identification number" as an example of information that can identify a person. Cookies with unique identifiers are strings of numbers and letters. There's no reason to exclude such cookies and similar technologies from the category identification numbers.[842] A cookie or another unique identifier allows a firm to follow a person's online behaviour, and to make inferences about that person's interests. As the Interactive Advertising Bureau UK explains, "[c]ookies are used in behavioural advertising to identify users who share a particular interest so that they can be served more relevant adverts."[843]

The Working Party says that a person can be identified without knowing his or her name. In its 2007 Opinion on personal data, the Working Party says that "singling out" an individual implies identifying that individual.[844] A person is identifiable if she

---

[837] See e.g. Interactive Advertising Bureau Europe - Youronlinechoices (about).
[838] See Schwartz & Solove 2011, with references to statutes.
[839] Pfitzmann & Hansen 2010, par. 9.
[840] European Agency for Fundamental Rights 2014, p. 36. "Data are anonymised if all identifying elements have been eliminated from a set of personal data. No element may be left in the information which could, by exercising reasonable effort, serve to re-identify the person(s) concerned. Where data have been successfully anonymised, they are no longer personal data" (internal footnote omitted), p. 45.
[841] Article 29 Working Party 2014, WP 216, p. 20.
[842] See Cuijpers et al. 2007, p. 25. See also Traung 2012, p. 37.
[843] Interactive Advertising Bureau United Kingdom 2009, p. 4. See on cookies chapter 2, section 2.
[844] Article 29 Working Party 2007, WP 136, p. 14.

can be distinguished within a group.[845] A firm that aims to individualise a person wouldn't have a strong case if it argued that its aim was not to identify that person. "In fact, to argue that individuals are not identifiable, where the purpose of the processing is precisely to identify them, would be a sheer contradiction in terms."[846] In later opinions the Working Party says explicitly that cookies and similar files with a unique identifier are personal data, because they "enable data subjects to be 'singled out', even if their real names are not known."[847]

> [B]ehavioural advertising involves the processing of unique identifiers be that achieved through the use of cookies, or any kind of device fingerprinting. The use of such unique identifiers allows for the tracking of users of a specific computer even when IP addresses are deleted or anonymised. In other words, such unique identifiers enable data subjects to be "singled out" for the purpose of tracking user behaviour while browsing on different websites and thus qualify as personal data.[848]

The impact assessment for the proposal for a Data Protection Regulation agrees with the Working Party about online identifiers: "[e]ven without a name or other traditional identifying attribute, it is often possible to effectively identify the individual to whom

---

[845] Article 29 Working Party 2007, WP 136, p. 12.
[846] Article 29 Working Party 2007, WP 136, p. 16. The Working Party doesn't make this remark in the context of behavioural targeting.
[847] Article 29 Working Party 2008, WP 148, p. 9. See also Article 29 Working Party 2010, WP 171, p. 9; Article 29 Working Party 2011, WP 188, p. 8.
[848] Article 29 Working Party 2011, WP 188, p. 8. See along similar lines CNIL 2014 (Google) (Google), p. 11-12; College bescherming persoonsgegevens (Dutch DPA) 2013 (Google), p. 50-57. The Working Party has described singling out as follows: "the possibility to isolate some or all records which identify an individual in the dataset" (Article 29 Working Party 2014, WP 216, p. 11).

the data relates."[849] The data can be used to individuate, isolate, or individualise a person.[850] Many authors agree.[851]

The fourth element of the personal data definition says that the information must relate to a "natural person."[852] This is usually the case with behavioural targeting. However, it's possible to think of situations where behavioural targeting data tied to a unique identifier aren't personal data, because they don't refer to a person. For instance, a computer in an internet café might be used by many people.[853] An ad network that builds a profile based on a cookie placed on that computer, might compile a profile based on the surfing behaviour of a group of people. Arguably such a profile doesn't consist of personal data. Nevertheless, if a firm uses a unique identifier for behavioural targeting, it usually identifies a person.

The Working Party isn't alone in its interpretation that behavioural targeting generally entails personal data processing. The International Working Group on Data Protection in Telecommunications agrees that behavioural targeting usually entails the processing of personal data.[854] Dutch law even contains a legal presumption regarding the use of tracking cookies and similar technologies for behavioural targeting. The use of such cookies is presumed to entail the processing of personal data.[855] In a letter to Google, signed by 27 national Data Protection Authorities, the Working Party says

---

[849] Impact Assessment for the proposal for a Data Protection Regulation (2012), Annex 2, p. 24.
[850] The phrase "individuate" is used by Hildebrandt for instance (Hildebrandt 2008, p. 19). Zwenne speaks of "isolating" (Zwenne 2013, p. 32). It must be noted that Zwenne disagrees with the Working Party about "singling out."
[851] See e.g. De Hert & Gutwirth 2008, p. 289; Leenes 2008; Traung 2010; and more hesitant: Koëter 2009. See also the references in Zwenne 2013, p. 35-36. Zwenne disagrees on this point with the Working Party: see Zwenne 2013, with references.
[852] See on the question of whether privacy rights do – or should – continue after death McCallig 2013; Harbinja 2013; Edwards 2013; Korteweg & Zuiderveen Borgesius 2009. See on the question of legal persons should be protected by data protection law Bygrave 2002, chapter 9-16.
[853] The example is taken from Article 29 Working Party 2007, WP 136, p. 17.
[854] International Working Group on Data Protection in Telecommunications (Berlin Group) 2013.
[855] Article 11.7a of the Dutch Telecommunications Act (See for a translation Zuiderveen Borgesius 2012).

that Google processes personal data about its "passive users."[856] These are users that are tracked by Google's ad network.[857]

Outside Europe, some regulators arrive at similar conclusions. For instance, the Privacy Commissioner of Canada says that behavioural targeting usually entails personal data processing.[858] In the US, the Federal Trade Commission (FTC) released a report in 2012 with recommendations regarding online data processing.[859] The recommendations apply to firms "that collect or use consumer data *that can be reasonably linked to a specific consumer, computer, or other device* (…)."[860] Therefore, the recommendations also apply to firms that gather data about individuals but don't tie a name to the data. However, not all regulators see behavioural targeting as the processing of personal data. For instance, the Office of the Australian Privacy Commissioner states that "[t]he information collected by online advertisers may often not be sufficient to identify you; it might just be general information about your interests and sites you have visited."[861]

In sum, according to the Working Party, a firm uses data to identify a person if the firm uses the data to single out somebody. Apart from that, we will see in the next section that firms using behavioural targeting can often attach names to the individual profiles they hold.

---

[856] Article 29 Working Party 2013 (Google letter).
[857] Article 29 Working Party 2013 (Google letter appendix), p. 2, footnote 2. Passive users are "users who does not directly request a Google service but from whom data is still collected, typically through third party ad platforms, analytics or +1 buttons." Reports by national Data Protection Authorities come to the same conclusion as the Working Party. See CNIL 2014 (Google); College bescherming persoonsgegevens (Dutch DPA) 2013 (Google)
[858] The International Working Group on Data Protection in Telecommunications also has members from outside Europe.
[859] Office of the Privacy Commissioner of Canada 2012, p. 2. See also Office of the Privacy Commissioner of Canada (Google) 2014.
[860] Federal Trade Commission 2012. p. 22 (emphasis added). In a 2014 report, the FTC includes "a persistent identifier, such as a customer number held in a "cookie" or processor serial number" in its personal data definition (Federal Trade Commission 2014, Appendix A, p. A16).
[861] Office of the Australian Privacy Commissioner 2011. The Australian Act speaks of "personal information" – not of personal data.

5.3 Data that identify people by name

It's often relatively easy for a firm that has an individual profile of a person, or for another party, to attach a name to that profile. To structure the analysis, this section distinguishes four situations where a firm processes data about a person.

> (i) A firm processes data about an individual, and it knows the name of the individual.
>
> (ii) A firm processes data about an individual, and it's fairly easy for the firm to tie a name to the data.
>
> (iii) A firm processes data about an individual, and it's difficult for the firm to add a name to the data, but it would be fairly easy for another party to tie a name to the data.
>
> (iv) A firm processes data about an individual, and it would be difficult for anybody to tie a name to the data.

Situation (i)

A firm processes data about an individual, and it has the individual's name. This firm clearly processes data about an identified person. For example, a provider of a social network site that engages in behavioural targeting often has profiles with names. Facebook requires people to register under their own name.[862]

---

[862] Facebook's Name Policy.

Situation (ii)

A firm processes data about an individual, and it's fairly easy for the firm to tie a name to the data. The preamble of the Data Protection Directive says: "to determine whether a person is identifiable, account should be taken of *all the means likely reasonably to be used* either by the controller or by any other person to identify the said person."[863]

The question is thus: what means can a firm that processes data about a person "reasonably likely use" to identify a person?[864] The answer to this question depends, among other things, on the state of science and technology, and on how costly it would be to identify somebody. According to the Working Party, "a mere hypothetical possibility to single out the individual is not enough to consider the person as 'identifiable'."[865]

It's often possible to identify people within an (supposedly) anonymised data set. In 2000 Sweeney found that 87% of the US population is uniquely identified by three attributes: their date of birth, gender, and ZIP code.[866] Techniques to re-identify data subjects continue to improve. Additionally, re-identification may become easier if more data sets that could be coupled with the source set become available, for instance from social network sites. Furthermore, computers keep getting faster, reducing the time needed for complicated calculations. Computer scientists summarise that de-identification of data is an "unattainable goal."[867]

Sometimes the person behind a nameless profile can be found without sophisticated data analysis. In 2006 search engine provider AOL released a data set of individual

---

[863] Recital 26 of the Data Protection Directive (emphasis added).
[864] Following the definition of "data subject" (article 4(1)) of the European Commission proposal for a Data Protection Regulation (2012), this study switches the words "likely reasonably" to "reasonably likely."
[865] Article 29 Working Party 2007, WP 136, p. 15. See also Article 29 Working Party 2014, WP 216, p. 8-9.
[866] Sweeney 2000; Sweeney 2001.
[867] Narayanan & Shmatikov 2010, p. 3. See generally on re-identification research Sweeney 2000; Sweeney 2001; Narayanan & Shmatikov 2008; Ohm 2010; Koot 2012 (chapter 2); Wu 2013; Article 29 Working Party 2014, WP 216.

nameless search profiles, tied to a random number. Within a few days, New York Times journalists had found one of the searchers: "A face is exposed for AOL searcher no. 4417749."[868] The search queries suggested that the searcher was an elderly woman with a dog, living in a specific town. An interview confirmed that the journalists had correctly identified her.

A behavioural targeting firm can often tie a name to the data about an individual it processes, taking into account "the means reasonably likely to be used" by the firm.[869] For instance, some firms offer services directly to consumers. If a firm has a cookie-based profile of a user, and the same firm offers an email service to that person, it can tie the user's email address to the cookie. Most email addresses are personal data.[870] In addition, email addresses and email messages often contain the user's name. The situation is similar if a firm offers a social network site or another service where people log in.

A search engine provider that has a nameless profile, tied to a unique identifier in a cookie, can also attach a name to a profile in many circumstances, as illustrated by the AOL case discussed above. If the firm stores all search queries, tied to the cookie, it holds plenty of information about the user. The firm could identify the person based on his or her searches. If the user sometimes searches for his or her name, this would be even easier.[871] As a Google employee said in a court case, "[t]here are ways in which a search query alone may reveal personally identifying information."[872] In sum, firms that process nameless profiles can often attach a name to the data with relative ease. They process personal data.

---

[868] Barbarom & Zeller 2006. See also Van Hoboken 2012, p 318; Article 29 Working Party 2014, WP 216, p. 11.
[869] Recital 26 of the Data Protection Directive.
[870] An "info@" email address of a company might not constitute personal data, if it doesn't refer to an individual.
[871] See on such "vanity searchers" Soghoian 2007.
[872] Cutts 2006, p. 9.

Situation (iii)

A firm processes data about an individual. It's difficult for the firm to add a name to the data, but it would be fairly easy for *another party* to tie a name to the data. An example might be an ad network that has a cookie-based profile of a person, including an IP address. Let's assume that it's difficult for the ad network to tie a name to the profile. But for the internet access provider of the person, it's fairly easy to tie a name to the IP address. For an online shop this would be easy too, if a person orders a product and provides the shop with a name and address.

Does it matter that only another party can identify the person? According to recital 26 of the Data Protection Directive, the answer is no: "to determine whether a person is identifiable, account should be taken of all the means reasonably likely to be used either by the controller *or by any other person* to identify the said person."[873] The recital's approach is sometimes called the absolute approach. A relative approach would imply only looking at the means at the disposal of the data controller.[874]

While recital 26 suggests an absolute approach, the means at the disposal of a data processor are relevant for the purpose of determining which means are reasonably likely to be used for identification. This can be illustrated with an example from Zwenne, presented here in a slightly revised form.[875] If a random person finds some human hairs, those hairs should probably not be seen as personal data for the finder. But if the police has a hair sample and sends this to a DNA research institute to match them against a database with DNA samples, the sample should probably be regarded as personal data.

---

[873] Emphasis added. See also Bygrave 2002, p 318; Article 29 Working Party 2010, WP 171, p 9; Article 29 Working Party 2007, WP 136, p 12-21.
[874] See European Commission's Information Society and Media Directorate-General 2011, p. 18-21.
[875] Zwenne 2013, p. 26-27. Zwenne argues for a relative approach.

Sometimes, Data Protection Authorities say that personal data are identifiable for one party, and not identifiable for another party.[876] Hence, Data Protection Authorities sometimes take into account what means can be used by the firm holding the data. For instance, the English Information Commissioner's Office appears to favour the relative approach.[877] In sum, while recital 26 appears to dictate an absolute approach, the relative approach may be relevant when determining which methods are likely to be used for identification. The Working Party clearly advocates the absolute approach in a 2014 opinion on anonymisation techniques.[878]

Computer scientist Narayanan discusses various ways for ad networks to attach a name to data. For instance, many websites disclose identifying information about their visitors to ad networks – often inadvertently. Furthermore, some firms specialise in tying names to data held by ad networks. The goal of some web surveys – "Win a free iPod!" – is matching email addresses and names to data. If you provide your email address to a firm that also operates a cookie, that firm can tie the two together. If one firm has tied a name to a cookie-profile, it can provide the name to other firms that only had a nameless profile ("cookie matching"). Narayanan summarises: "[t]here is no such thing as anonymous online tracking."[879]

The discussion about behavioural targeting and the scope of data protection law resembles the discussion about IP addresses. The Working Party and many judges in Europe say that IP addresses should generally be considered to be personal data.[880] Others counter that IP addresses shouldn't be considered as personal data in all circumstances. First, some argue for a relative approach. For instance, Google says that IP addresses shouldn't be seen as personal data if the firm holding the IP address

---

[876] Impact Assessment for the proposal for a Data Protection Regulation (2012), Annex 2, p. 15-16.
[877] Information Commissioner's Office 2012, p. 21. The German situation is more complicated, but also boils down to a relative approach (see Korff 2010b, p. 4).
[878] Article 29 Working Party 2014, WP 216, p. 9.
[879] Narayanan 2011. See on cookie synching chapter 2, section 6.
[880] See about the status of IP addresses as personal data: Impact Assessment for the proposal for a Data Protection Regulation (2012), Annex 2, p. 14-16; Time.lex 2011. See for criticism on the Time.lex report: Zwenne 2013.

can't tie a name to it.[881] Second, sometimes IP addresses can't be used to identify a person.[882] For example, the country Qatar routed all internet traffic through a couple of IP addresses.[883] And some organisations access the internet though one IP address. In such cases, a mere IP address without any other information may not be enough to identify somebody.

In the 2012 *Scarlet* case, the European Court of Justice decided that the IP addresses in that case were personal data. Copyright organisation Sabam requested internet access provider Scarlet to install a filtering system to help enforce copyrights. Scarlet refused. Prompted by the Advocate General, the European Court of Justice decided that the IP addresses are personal data. "Those addresses are protected personal data because they allow those users to be precisely identified."[884] The Advocate General referred to opinions of the Article 29 Working Party and the European Data Protection Supervisor to support his conclusion that the IP addresses were personal data.[885]

Still, the discussion about IP addresses isn't over. The Court uses ambiguous language, but it may have suggested a relative approach.[886] For parties that aren't internet access providers, it's harder to tie an IP address to a name. They may still try to argue that IP addresses are not personal data in their hands.[887] In sum, European Data Protection Authorities generally consider IP addresses to be personal data, and judges tend to take a similar position.

---

[881] See e.g. Whitten 2008.
[882] See Zwenne 2013, p. 27-28.
[883] Zittrain 2008, p. 157.
[884] CJEU, C-70/10, Scarlet v Sabam, 24 November 2011, par. 51.
[885] Opinion AG (14 April 2011) for CJEU, C-70/10, Scarlet v Sabam, 24 November 2011, par. 75-80.
[886] In an earlier publication I assumed that the Court limited its remark to IP addresses in the hands of access provider Scarlet (Kulk & Zuiderveen Borgesius 2012). Now I believe the Court may have taken an absolute approach, as the Court talks about "users", and not about "subscribers." See the definition of "user" (article 2(a) of the e-Privacy Directive), and of "subscriber" (article 2(k) of the Framework Directive 2002/21). A full discussion of the Scarlet Sabam case falls outside this study's scope.
[887] In a 2013 opinion the Advocate General also sees IP addresses as personal data when they're in the hands of Google. This suggests an absolute approach (Opinion AG (25 June 2013), C-131/12, Google Spain, par. 48). The Court has neither confirmed nor disproved this view in the subsequent judgment. In October 2014, the German Bundesgerichtshof has asked preliminary questions to the CJEU regarding the question of whether dynamic IP addresses should be seen as personal data (Bundesgerichtshof 2014; see Husovec 2014).

The case law on IP addresses is relevant for two reasons. First, the discussion resembles the discussion about behavioural targeting profiles. The case law on IP addresses confirms that nameless data that refer to a device can be personal data. But there's an important difference between IP addresses and personal profiles that are used for behavioural targeting. Individual behavioural targeting profiles contain much more information than an IP address.[888] Second, firms that process data about individuals for behavioural targeting usually tie IP addresses to the data. For instance, an ad network typically needs the IP address of the receiving device to display the ad.

To conclude, if a firm processes data about an individual for behavioural targeting, and it's fairly easy for another party to tie a name to the data, the Data Protection Directive's preamble suggests that the data sould be regarded as personal data.

Situation (iv)

A firm processes data about an individual, and it would be difficult for *anybody* to tie a name to the data. As it's often fairly easy for a firm to tie a name to the data it processes for behavioural targeting, the number of situations in this category is likely to be small. This category was discussed in section 2: the Working Party says it's not relevant whether a firm can attach a name to the data. If the firm uses the data to single out a person, the firm processes personal data.

5.4 Data protection law should apply to behavioural targeting

Many scholars say a logical interpretation of data protection law implies that data about a nameless individual should be regarded as personal data.[889] This study agrees. Some say that, if necessary, the personal data definition should be adapted to

---

[888] See Van Hoboken 2012, p. 328.
[889] See for instance De Hert & Gutwirth 2008; Leenes 2008; Koëter 2009; Traung 2010; Traung 2012. But see Zwenne 2013, with references, for another view.

emphasise that it includes data used to single out a person. De Hert & Gutwirth have hinted at "a shift from personal data protection to data protection *tout court*."[890]

But, apart from an analysis of the law, why should data that are used to single out a person for behavioural targeting be regarded as personal data? First, the processing of information for behavioural targeting triggers many concerns that lie at the core of data protection law. The risks of large-scale data collection don't disappear merely because data about a person can't be tied to a name.[891] For instance, massive collection of information on user behaviour can cause a chilling effect; which remains true even if firms collect pseudonymous data. Firms compile detailed information about people, and can classify people, while the individual has little control over this process.[892] As Turow notes, "[i]f a company can follow your behaviour in the digital environment – an environment that potentially includes your mobile phone and television set – its claim that you are 'anonymous' is meaningless. (…) It matters little whether your name is John Smith, Yesh Mispar, or 3211466."[893] And a firm could discriminate against a person, for instance when the cookie says the person is "handicapped",[894] or is in the interest category "lesbian, gay, bisexual, transgender."[895]

True, certain risks are reduced when a firm doesn't attach a name to the data it holds about a person. Suppose a firm with pseudonymous profiles regarding people's browsing behaviour experiences a data breach: the firm accidentally publishes millions of browsing profiles on the web. People who see the data learn that the person behind ID *xyz* visited dirty-pictures.com, or another-embarrassing-website.com. But somebody who sees the pseudonymous browsing profile doesn't immediately learn the name of the person who visited those websites. Hence, the

---

[890] De Hert & Gutwirth 2008, p. 289.
[891] Article 29 Working Party 2013, WP 203, p. 46.
[892] See chapter 3, section 3.
[893] Turow 2011, p. 7 (see also p. 100).
[894] Rocket Fuel, Health Related Segments 2014. All the examples are taken from US companies, but it can't be ruled out that they also operate cookies on devices in the EU.
[895] Flurry (audiences). Flurry is firm offering analytics and advertising for mobile devices. Among the demographic data that advertisers can select, Flurry lists "race" (Flurry, factual).

privacy risks are reduced, because the leak concerns pseudonymous data. There's less risk of embarrassment or other unpleasant surprises for the person behind ID *xyz*. However, the AOL search data case illustrates that it may be possible to find the person behind a pseudonymous profile.[896]

Privacy risks are also reduced for another reason when a firm doesn't know the name of a person behind a cookie profile. For example, say a supermarket offers a loyalty card to customers, and knows the names of those customers. If a behavioural targeting firm wanted to tie a profile based on information gathered through a supermarket loyalty card to an online profile, it would be practical if a name were linked to both profiles. Without a name, it's harder to merge data from different databases.[897]

Nevertheless, many risks remain, even when firms don't tie a name to data. The behavioural targeting industry compiles large amounts of information about people, and if data protection law didn't apply, this industry could operate largely unregulated. We will see in later chapters that applying data protection law provides more protection to internet users than only applying the e-Privacy Directive's consent requirement for cookies and similar tracking technologies.[898]

Second, a name is merely one of the identifiers that can be tied to data about a person. In some situations, the name is the most practical identifier. But for many purposes, a name isn't the most effective identifier. If the purpose is sending messages to a phone, or tracking a phone's location, a phone number or one of the ID numbers of a phone is the easiest identifier. Furthermore, a unique number is often a better identifier than a name, because names may not be unique.[899]

For an ad network that wants to track a person's browsing behaviour, or wants to target ads to a person, a cookie is a better identifier than a name. Many firms aren't

---

[896] See section 3 of this chapter.
[897] See chapter 2, section 6.
[898] See chapter 6, 8 and 9.
[899] The Working Party notes that very common names by itself aren't always personal data, because they can't be used to identify people (Article 29 Working Party 2007, WP 136, p. 13).

interested in tying a name to data they process for behavioural targeting. When Mozilla, the firm behind the Firefox browser, considered blocking third party cookies by default, the Interactive Advertising Bureau (IAB) US reacted furiously.[900] The reaction suggests that the IAB sees the threat of not being able to use people's names for behavioural targeting as less serious than the threat that third party cookies won't work anymore.

Third, the goal of behavioural targeting is targeting the right person, with the right ad, at the right time. It would be odd to say that data used by a firm for individualised tracking and targeting aren't personal data. The whole point of behavioural targeting is singling people out, and targeting ads to *specific* individuals.

Seeing data that can single out a person as personal data corresponds with the rationale for the Data Protection Directive.[901] One of the Directive's goals is protecting privacy and other fundamental rights.[902] The European Court of Justice says that the Directive aims for a "high level" of protection,[903] and that fundamental rights guide the interpretation of the Directive.[904] Furthermore, "limitations in relation to the protection of personal data must apply only in so far as is strictly necessary."[905] According to the European Court of Human Rights, the right to private life is a broad term that should be applied dynamically and pragmatically.[906] This study suggests that data protection law, like the European Convention on Human Rights, should be seen as a living instrument. In the light of new developments such as behavioural targeting,

---

[900] Interactive Advertising Bureau Europe 2013. See chapter 2, section 2.
[901] See Korff 2010a, p. 47-48.
[902] Article 1(1) of the Data Protection Directive.
[903] ECJ, C-524/06, Huber, 16 December 2008, par. 50; CJEU, C-131/12, Google Spain, 13 May 2014, par. 66.
[904] ECJ, C-465/00, C-138/01 and C-139/01, Österreichischer Rundfunk, 20 May 2003, par. 68; CJEU, C-131/12, Google Spain, 13 May 2014, par. 68.
[905] See e.g. CJEU, C-293/12 and C-594/12, Digital Rights Ireland Ltd, 8 April 2014, par. 52; CJEU, Case C-473/12, 7 November 2013, Institut professionnel des agents immobiliers, par. 39 (with further references).
[906] See the case law mentioned in chapter 3, section 2. The Court says: "[t]hat broad interpretation [of he right to private life] corresponds with that of the Council of Europe's Convention of 28 January 1981 for the Protection of Individuals with regard to Automatic Processing of Personal Data (…)" ECtHR, Amann v. Switzerland, No. 27798/95, 16 February 2000, par. 65). See similarly ECtHR, Rotaru v. Romania, No. 28341/95, 4 May 2000, par. 43.

it wouldn't make sense to limit the scope of data protection law to data that can identify people by name.

Criticism on the singling out perspective

Some authors criticise the tendency of Data Protection Authorities to interpret the personal data definition broadly and point to several disadvantages.[907] The main points are summarised here. It's concluded that the arguments aren't persuasive.

First, it has been argued that firms have less incentives for investing in pseudonymisation technology if the law covers pseudonymised data.[908] While it may be true that firms have less incentive to pseudonymise data, the law requires appropriate security measures from data controllers, and pseudonymisation can improve security. For instance, pseudonymisation can help to keep data subjects' names hidden from employees that don't need to see the names.[909] Hence, pseudonymisation can improve data security. But replacing a name with another identifier isn't enough to keep data outside the scope of data protection law.[910]

Second, some suggest applying data protection law to behavioural targeting would be bad for business and innovation.[911] This argument isn't sufficient to keep behavioural targeting outside data protection law's scope. When information is within the scope of data protection law, that doesn't mean processing is prohibited. But it does imply that firms have to comply with the data protection principles. It's certainly true that some firms would make less profit when they have to comply with data protection law. But even if fundamental rights were ignored and only economic effects were taken into account, a more relevant question is whether society as a whole wins or loses. Chapter 7 shows that it's unclear whether more or less legal privacy protection is better from

---

[907] The most detailed and eloquent critique is offered by Zwenne (Zwenne 2013).
[908] Zwenne 2010, p. 336.
[909] The Data Protection Directive requires an appropriate level of security for personal data (article 17). See chapter 4, section 2 on the security principle.
[910] See European Agency for Fundamental Rights 2014, p. 45-46; Article 29 Working Party 2014, WP 216, p. 20.
[911] See e.g. Stringer 2013.

an economic perspective.[912] And while innovation – a term almost as vague as privacy – is important, it doesn't trump fundamental rights. If it were good for innovation if children below eight worked in factories, we still shouldn't allow it.[913] Moreover, if regulation pushes firms towards developing new and privacy preserving technologies, this is also innovation.

Third, some say that applying data protection law to data that identify nameless people could lead to peculiar situations. For instance, if a firm holding nameless profiles granted data subjects the right to access their data, the firm might have to ask the data subject to identify herself, which might involve asking for more personal data.[914] However, interpreting the data protection provisions in a reasonable manner can prevent absurd results.[915] As the Working Party puts it, "[i]t is a better option not to unduly restrict the interpretation of the definition of personal data but rather to note that there is considerable flexibility in the application of the rules to the data."[916]

Fourth, some say a broad interpretation of personal data implies that data protection law applies, even when there are no privacy threats. Some suggest that data protection law shouldn't be severed from the right to privacy.[917] That argument doesn't fit well with positive law, as the EU Charter of Fundamental Rights distinguishes the right to data protection and the right to privacy. Furthermore, many authors say it's an advantage that data protection law applies to all personal data, rather than just private personal data.[918]

---

[912] See chapter 7, section 2.
[913] Helen Nissenbaum made a remark among these lines at the Acatech Symposium Internet Privacy (26 March 2012, Berlin). Article 32 of the EU Charter of Fundamental rights: "The employment of children is prohibited." See for a general critique of the innovation argument Morozov 2013, and Richards 2014a, p. 28 - p. 36.
[914] Schwartz & Solove 2011, p. 1876. See also Zwenne 2013, p. 37. See on access rights and pseudonymous data: chapter 8, section 2.
[915] Like with any statue, there's also a risk that data protection law is applied in an unreasonable manner.
[916] Article 29 Working Party 2007, WP 136, p. 5.
[917] Cuijpers & Marcelis 2012.
[918] See e.g. De Hert & Gutwirth 2006, p. 94; Hildebrandt et al. 2008, p. 245. See also chapter 4, section 2, and chapter 9, section 2.

Fifth, some worry that almost everything could become personal data if data that are used to single out a person are seen as personal data. Enforcing data protection law would become too difficult. Data Protection Authorities would only be able to enforce the law against a few wrongdoers. This could lead to arbitrary decisions about enforcement, which would be bad for legal certainty. A related point is that the scope of the personal data definition becomes too uncertain. This would also be bad for legal certainty.[919]

There's merit to the point that the broad scope of data protection law makes enforcement difficult. But limiting the scope of data protection to exclude pseudonymous data wouldn't be the right reaction. Similarly, it's probably good that we have environmental law, even though it's impossible to catch everybody who breaches the law.[920] Furthermore, in legal practice the fringes of a definition can always provoke discussion. In sum, the criticism doesn't justify leaving behavioural targeting outside the scope of data protection law.

Merely ensuring that data protection law applies to behavioural targeting doesn't solve all privacy problems. But, with all its weaknesses, at least data protection law provides a framework to assess fairness when personal data are processed. And since data protection law requires firms to disclose information about their processing practices, it can help to make the processing transparent. When problems are found, this could lead to the conclusion that more regulation is needed.[921]

---

[919] This fifth point is made most convincingly by Zwenne 2013, in particular p. 33-35. Korff agrees that a broad interpretation of personal data can have drawbacks, but still argues for a broad interpretation (Korff 2010a, p. 47-48).
[920] See on privacy scholarship taking inspiration from environmental law Hirsch 2006.
[921] See also chapter 4, section 3.

5.5 Data protection reform and pseudonymous data

The 2012 European Commission proposal for a Data Protection Regulation led to much discussion about the law's material scope.[922] The proposal doesn't significantly alter the personal data definition. But the proposal includes "online identifiers" and "location data" in the list of examples of information that can be used to identify a data subject.[923] The preamble and the impact assessment that accompanied the proposal show that the European Commission intended data protection law to apply to behavioural targeting.[924]

The Commission's proposal chooses the absolute approach to identifiability. The definition says that the "means reasonably likely to be used by the controller *or by any other natural or legal person*" should be taken into account when determining identifiability.[925] The proposal defines personal data as "any information relating to a data subject."[926] The latter is defined as follows:

> "Data subject" means an identified natural person or a natural person who can be identified, directly or indirectly, by means reasonably likely to be used by the controller or by any other natural or legal person, in particular by reference to an *identification number*, location data, *online identifier* or to one or more factors specific to the physical, physiological, genetic,

[922] I took part in this debate, for instance at the Dutch and the European Parliament (Zuiderveen Borgesius 2012a).
[923] Article 4(2) of the European Commission proposal for a Data Protection Regulation (2012)
[924] See recital 20 and 46 and article 3(2)(b) of the European Commission proposal for a Data Protection Regulation (2012). See also Impact Assessment for the proposal for a Data Protection Regulation (2012), p. 31.
[925] The proposed definition incorporates parts of the old recital 26 ("the controller or by any other") in the definition of personal data.
[926] Article 4(2) of the European Commission proposal for a Data Protection Regulation (2012).

> mental, economic, cultural or social identity of that person (emphasis added).[927]

Recital 24 of the proposal discusses online tracking and elaborates on the use of "online identifiers." The recital begins with suggesting that data about a person that are attached to a unique identifier, such as a cookie, are usually personal data. A tracking cookie or another identifier can be used to profile individuals and to identify them.

> When using online services, individuals may be associated with online identifiers provided by their devices, applications, tools and protocols, such as Internet Protocol addresses or cookie identifiers. This may leave traces which, combined with unique identifiers and other information received by the servers, may be used to create profiles of the individuals and identify them. It follows that identification numbers, location data, online identifiers or other specific factors as such need not necessarily be considered as personal data in all circumstances.[928]

The recital's last sentence suggests that there may be circumstances where online identifiers shouldn't be considered as personal data.[929] It's true that in some circumstances unique identifiers might not relate to an individual, for instance when many people use the same computer. But the last sentence may create a gap in the data protection regime. Among others, the Working Party says the last sentence must

---

[927] Article 4(2) of the European Commission proposal for a Data Protection Regulation (2012) (capitalisation and punctuation adapted).
[928] Recital 24 of the European Commission proposal for a Data Protection Regulation (2012).
[929] During the last weeks before the European Commission proposal was presented, the last sentence was changed. An earlier version of the proposed Regulation concluded in the last sentence that the "Regulation should be applicable to processing involving such data" (European Commission proposal for a Data Protection Regulation (2012), leaked Interservice draft (2011), recital 23).

be deleted, to emphasise that data protection law fully applies to unique identifiers such as tracking cookies.[930]

Discussion on the scope of data protection law continued after the Commission's proposal. The proposal has led to an enormous amount of lobbying, including by firms from the US[931] During the discussions about the Data Protection Regulation proposals, a new legal concept was suggested: "pseudonymous data." The Interactive Advertising Bureau, and firms such as Yahoo and Amazon, both using behavioural targeting, lobbied for amendments that would introduce a lighter regime for "pseudonymous" data.[932] At least one non-governmental organisation was in favour of a lighter regime for pseudonymous data, because that would incentivise firms to pseudonymise data, which would help data security.[933]

Some European Parliament members proposed amendments to introduce a data protection-light regime for pseudonymous data. For instance, shadow rapporteur Alvaro proposed adding a rule that would legitimise the processing of pseudonymous data. "Processing of pseudonymized data shall be lawful."[934] Other Parliament members proposed leaving pseudonymous data largely outside the scope of data protection law.[935]

LIBE Compromise

In January 2013, the Rapporteur for the European Parliament, Albrecht, presented his draft report. The report codifies the Working Party's view on the definition of

---

[930] Article 29 Working Party 2012, WP 199, p. 5-6; Korff 2012, p. 32.
[931] See Albrecht 2013. Albrecht estimates that more than half of the firms that contacted him regarding the proposals are from the US (Traynor 2014). See generally on corporate lobbying in Brussels Horten 2011.
[932] See on the lobbying by the Interactive Advertising Bureau for a lighter regime for pseudonymous data Stringer 2013. Amazon proposed amendments, ready to submit (Amazon proposed amendments). See also Yahoo proposed amendments.
[933] Center for Democracy & Technology 2013a.
[934] Alvaro 2013, amendment 48, p. 31. The rule would imply that firms don't need another legal basis (such as consent) for the processing of pseudonymous data; see chapter 6.
[935] See amendment 327 by Jens Rohde & Bendt Bendtsen: "encrypted and some psydonymised [sic] data fall outside this regulation" (ITRE Amendments).

personal data, by adding the "single out" phrase to the personal data definition.[936] Hence, any data relating to a person that "can be identified *or singled out"* are personal data. The Albrecht report thus emphasises that data protection law applies to processing data about nameless individuals. The draft report by Rapporteur Albrecht also included a definition of "pseudonymous data", as a category of personal data.[937]

In March 2014, the European Parliament adopted a compromise text (the "LIBE Compromise"), which the Parliament's LIBE Committee prepared on the basis of the 3999 amendments by the members of parliament.[938] The LIBE Compromise defines personal data roughly the same as the 1995 Data Protection Directive.[939] But the LIBE Compromise adds location data and unique identifiers to the examples of possible identifiers. The preamble makes clear that the LIBE Compromise takes an absolute approach to identifiably. "To determine whether a person is identifiable, account should be taken of all the means reasonably likely to be used *either by the controller or by any other person* to identify or single out the individual directly or indirectly."[940]

Recital 24 of the LIBE Compromise suggests that in principle the regulation is applicable to processing unique identifiers such as cookies, IP addresses and RFID tags.[941] In other words, the Regulation seems to apply to data that can "single out" a

---

[936] He proposed the following definition: "data subject' means an identified natural person or a natural person who can be identified *or singled out*, directly or indirectly, *alone or in combination with associated data*, by means reasonably likely to be used by the controller or by any other natural or legal person, in particular by reference to *a unique identifier*, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural, social *or gender* identity *or sexual orientation* of that person" (the emphasised words are proposed) (amendment 84, article 4(1), Draft Albrecht report).

[937] Amendment 85, article 4(2)(a), Draft Albrecht report. The draft report suggests that under certain conditions, a system like Do Not Track could be used to signify consent to the processing of such data (amendement 105, article 7(2)(a)).

[938] LIBE Compromise, proposal for a Data Protection Regulation (2013). See for a concise overview of the discussions from January 2012 to January 2014: Burton & Pateraki 2013; Kuner et al. 2014.

[939] 'Personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, unique identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social or gender identity of that person." (Article 4(2) of the LIBE Compromise, proposal for a Data Protection Regulation (2013), capitalisation and punctuation adapted).

[940] Recital 23 of the LIBE Compromise, proposal for a Data Protection Regulation (2013) (emphasis added).

[941] Recital 24 of the LIBE Compromise, proposal for a Data Protection Regulation (2013).

person, including if no name is tied to the data.[942] However, the LIBE Compromise also introduces a new category of personal data: "pseudonymous data."

> "Pseudonymous data" means personal data that cannot be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organisational measures to ensure non-attribution.[943]

Such pseudonymous data are subject to a lighter regime in the LIBE Compromise. One of the main differences is that the LIBE Compromise allows processing pseudonymous data without consent in some circumstances.[944] But the introduction of the pseudonymous data category might lead to a level of protection that is too low.[945] At the time of writing of this study, the debate about the legal status of pseudonymous data is ongoing.

5.6 Special categories of data

Data protection law has a stricter regime for "special categories of data." These are "data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life."[946] According to the European Court of Justice, data concerning health must be

---

[942] See also recital 23 of the LIBE Compromise, proposal for a Data Protection Regulation (2013): "To determine whether a person is identifiable, account should be taken of all the means reasonably likely to be used either by the controller or by any other person to identify or single out the individual directly or indirectly."
[943] Article 4(2a) of the LIBE Compromise, proposal for a Data Protection Regulation (2013). The LIBE Compromise also includes a definition of encrypted data in article 4(2b). Recital 23 says the regulation doesn't apply to anonymous data: "information that does not relate to an identified or identifiable natural person."
[944] See chapter 6, section 2. The lighter regime for pseudonymous data has more consequences. See for instance recital 38 and 58a (on the balancing provision and profiling), health data (recital 122a and article 81(2)(a)), and article 10.
[945] European Commissioner Reding warns: "pseudonymous data must not become a Trojan horse at the heart of the Regulation, allowing the non-application of its provisions" (Reding 2014).
[946] Article 8(1) of the Data Protection Directive.

given a wide interpretation.[947] This suggests that "special categories of data" must be interpreted broadly.

Processing special categories of data is only allowed with the data subject's "explicit consent."[948] About half of the member states require such explicit consent to be in writing. Some member states have chosen not to allow people to override the prohibition with consent.[949] There are exceptions to the in-principle processing prohibition, for instance in the medical context and for churches. These provisions aren't, however, relevant for behavioural targeting.[950]

The stricter regime for special categories of data can be explained by the wish to prevent unfair discrimination.[951] "In general, information relating to the intimate private life of persons or information which might lead to unfair discrimination should not be recorded or, if recorded, should not be disseminated," said the Council of Europe in 1972.[952] And the Directive's preamble says that special categories of data "are capable by their nature of infringing fundamental freedoms or privacy."[953] The stricter regime for special categories of data also seems to be related to privacy as limited access, or as intimacy.[954] Certain types of data are considered particularly private or sensitive.[955] Data protection instruments such as the Data Protection

---

[947] ECJ, C-101/01, Lindqvist, 6 November 2003, par. 50.
[948] See article 8(2)(a) of the Data Protection Directive. See also chapter 9, section 5.
[949] For instance: Italy and Sweden require consent to be in writing (Impact Assessment for the proposal for a Data Protection Regulation (2012), Annex 2, p. 29). See article 8(2)(a) of the Data Protection Directive.
[950] Article 8(2)-8(7) of the Data Protection Directive. There's also an exception for sensitive data that are "manifestly made public by the data subject" (article 8(2)(e). It doesn't seem plausible that firms can invoke this ground for the gathering of data for behavioural targeting. An exception might be a firm that gathers information that people publish about themselves on the web.
[951] The United Nations guidelines use the header "principle of non-discrimination" for their provision on sensitive data, article 5 (UN General Assembly, Guidelines for the Regulation of Computerized Personal Data Files, 14 December 1990).
[952] Committee of Ministers, Resolution (73)22 on the protection of the privacy of individuals *vis-à-vis* electronic data banks in the private sector, 26 September 1973, article 1.
[953] Recital 33 of the Data Protection Directive.
[954] See Bygrave 2002, p. 132.
[955] See e.g. European Union Civil Service Tribunal, Civil Service Tribunal Decision F-46/095, V & EDPS v. European Parliament, 5 July 2011, par. 163; I. v. Finland, App. No. 25011/03, 17 Jul. 2008, par. 38. See along similar lines Z v. Finland (9/1996/627/811) 25 February 1997, par. 95. See on special categories of data also chapter 9, section 5.

Convention and the United Nations Data Protection Guidelines also have stricter rules for certain types of personal data.[956]

The European Commission proposal for a Data Protection Regulation retains the stricter regime for special categories of data. The categories remain roughly the same.[957] While the proposal always requires consent to be "explicit", the distinction between special categories of data and non-special personal data remains relevant. The Regulation only allows processing of special categories of data for direct marketing and behavioural targeting after obtaining the data subject's consent.[958]

Research suggests that many people indeed regard special categories of data, such as data regarding health, as sensitive. Many people also consider data regarding their finances sensitive.[959] The European Consumer organisation says financial data should be added to the category of sensitive data.[960] However, there are cultural differences between member states. For instance, in Finland data from the tax office about people's income are publicly available.[961]

Behavioural targeting and special categories of data

Do firms that engage in behavioural targeting process special categories of data? Some firms do, some don't, and many operate in a grey area. There are firms that clearly process special categories of data for behavioural targeting. Some firms target advertising based on categories such as "US Hispanics",[962] "arthritis", "cardiovascular

---

[956] Article 6 of the Data Protection Convention 1981; Article 5 of the UN General Assembly, Guidelines for the Regulation of Computerized Personal Data Files, 14 December 1990.
[957] Genetic data are added to the definition, and data about philosophical beliefs are deleted (article 9 of the European Commission proposal for a Data Protection Regulation (2012)). Genetic data are defined in article 4(10). See also article 33(2)(a), which suggests that processing operations involving certain kinds of data "present specific risks."
[958] See on the legal basis for processing (such as consent) chapter 6.
[959] See Leon et al. 2013. See also the Commission Regulation on Data Breaches (no. 611/2013), which gives examples of data that likely to adversely affect people's personal data or privacy in the case of a data breach. The list of examples includes "financial information (…) internet log files [and] web browsing histories (article 3(2) and recital 12).
[960] European Consumer Organisation BEUC 2013, p. 16.
[961] See ECJ, C-73/07, Satamedia, 16 December 2008.
[962] Batanga Network Inc.

general health", [963] "lesbian, gay, bisexual, and transgender," [964] or "disabled/handicapped consumers."[965] Such firms process special categories of data.

It's possible to use behavioural targeting without processing special categories of data. Suppose an ad network only works with websites about comic books. The firm puts cookies in one of three categories: people who like science fiction, people who like superheroes, and people who like other topics. Immediately after categorising people, the firm deletes the list of visited websites. In this example, the firm doesn't process special categories of data.

But many firms using behavioural targeting operate in a grey area – perhaps most of them. Say a firm puts people (or cookies) in the category "unions and labour movement", based on their surfing behaviour.[966] A person's interest in the labour movement could imply a political opinion. And certain website visits could suggest a person's sexual orientation or medical condition, even if there are no behavioural categories associated with the raw data.[967] In sum, behavioural targeting often entails the processing of data that could be considered "special categories of data."

e-Privacy Directive

In 1997, two years after the Data Protection Directive, the EU adopted the Directive on personal data processing in the telecommunications sector.[968] In 2002 it was replaced by the Directive "concerning the processing of personal data and the

---

[963] Yahoo! Privacy.
[964] Flurry (audiences). Flurry is firm offering analytics and advertising for mobile devices. Among the demographic data that advertisers can select, Flurry lists "race" (Flurry, factual).
[965] Rocket Fuel, Health Related Segments 2014. All the examples are taken from US companies, but it can't be ruled out that they also operate cookies on devices within the EU. See on political behavioural targeting also chapter 2, section 7, chapter 3, section 3, and chapter 9, section 5.
[966] Google Ad Interest Categories 2014.
[967] For instance, the Office of the Privacy Commissioner of Canada finds that "Google is delivering tailored ads in respect of a sensitive category, in this case, health" (Office of the Privacy Commissioner of Canada (Google) 2014, par. 26). Of course, that report doesn't concern EU data protection law, but the Canadian regime has similarities to the EU regime.
[968] Directive 97/66/EC (the ISDN Directive).

protection of privacy in the electronic communications sector." This e-Privacy Directive was meant to be more in line with new technologies.[969]

The e-Privacy Directive has a stricter regime that applies when certain types of firms process location data or traffic data. Such data may only be processed based on consent, or in some narrowly defined circumstances.[970] Traffic data, sometimes called metadata, are "any data processed for the purpose of the conveyance of a communication on an electronic communications network or for the billing thereof."[971] Examples of traffic data are the time of a communication, the email address of communicating partners, and the IP address used to access the internet.[972]

The Advocate General of the European Court of Justice says traffic data are "data which are in a sense more than personal."[973] Traffic data are "'special' personal data, the use of which may make it possible to create a both faithful and exhaustive map of a large portion of a person's conduct strictly forming part of his private life, or even a complete and accurate picture of his private identity."[974] With modern communication technology, the line between traffic data and communications content becomes increasingly blurry. For instance, the subject line of an email message could be seen as traffic data or as communications content, and traffic data can paint a detailed picture of a person's life.[975]

Location data are data "indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service."[976] Location data

---

[969] Recital 4 of the e-Privacy Directive. This study refers to the consolidated version (amended in 2009), unless otherwise noted. See on the e-Privacy Directive chapter 6, section 4, chapter 8, section 4.
[970] Article 6 and 9 of the e-Privacy Directive.
[971] Article 2(b) of the e-Privacy Directive.
[972] See recital 15 of the e-Privacy Directive, and the Data Retention Directive.
[973] Opinion AG (12 December 2013) for CJEU, C-293/12 and C-594/12, Digital Rights Ireland Ltd, 8 April 2014, par. 65.
[974] Opinion AG (12 December 2013) for CJEU, C-293/12 and C-594/12, Digital Rights Ireland Ltd, 8 April 2014, par. 74)
[975] See on the blurry line between traffic data, a EU law angle: Koops & Smit 2014; Breyer 2005; United Nations High Commissioner for Human Rights 2014, p. 6-7. See for a computer science angle Felten 2013; Mayer & Mutchler 2014.
[976] Article 9 of the e-Privacy Directive.

can be sensitive.[977] For example, a phone's location data can disclose visits to the hospital or a mosque, or the location of one's bed. The e-Privacy Directive's regime for traffic data and location data is similar to the regime for special categories of data in the Data Protection Directive. Unless a legal exception applies, consent is needed for the processing of traffic and location data.[978]

But the scope of these provisions in the e-Privacy Directive is narrow. The requirements regarding traffic and location data only apply to providers of publicly available electronic communications services, such as internet access providers or phone operators (telecommunication providers for short).[979] The e-Privacy Directive's background as a directive regulating telecommunications companies can help to explain the narrow scope of these provisions.[980] But many firms, such as ad networks and providers of smart phone apps, process more data of a sensitive nature than telecommunications providers. However, ad networks and apps providers aren't subject to the e-Privacy Directive's rules for traffic and location data. The Working Party suggests that when applying data protection law, location and traffic data should be treated as prima facie sensitive, although they're not within the definition of "special categories of data."[981]

In the behavioural targeting area, the most relevant provision of the e-Privacy Directive is article 5(3), which requires consent for the use of most tracking

---

[977] See Article 29 Working Party 2011, WP 185, p. 7.
[978] See article 5, 6 and 9 of the e-Privacy Directive.
[979] An "electronic communications service" is, in short, a service that consists wholly or mainly in the conveyance of signals on electronic communications networks (article 2(c) of the Framework Directive 2002/21/EC (amended in 2009)). It's thus a transmission service. See Zuiderveen Borgesius 2011a.
[980] See Arnbak 2013a, p. 9.
[981] Article 29 Working Party 2013, WP 203, p. 25; p. 66. See also the European Commission proposal for a Data Protection Regulation (2012): article 33(2)(a) suggests that certain processing operations involving location data "present specific risks." See also the Commission Regulation on Data Breaches (no. 611/2013), article 3(2) and recital 12. Financial information and web browsing histories are given as examples of data that are likely to affect privacy in case of a breach. See on the scope of the e-Privacy Directive also chapter 6, section 4, and chapter 9, section 5.

technologies. The scope of article 5(3) isn't limited to telecommunications providers. That provision is discussed in the next chapter.[982]

5.7 Conclusion

Two conclusions can be drawn from this chapter. First, an analysis of current law shows that data protection law generally applies to behavioural targeting. Second, from a normative perspective, data protection law should apply to behavioural targeting.

Personal data are "any information relating to an identified or identifiable natural person."[983] The Article 29 Working Party says that firms carrying out behavioural targeting usually process personal data; even if they don't tie a name to the data they hold about an individual. A name is not needed to identify a person. Firms process the data to single out a person. Therefore, the data processed for behavioural targeting are generally personal data. Moreover, it's often fairly easy for the firm using behavioural targeting, or for another party, to tie a name to the data.

Heated discussions about pseudonymous data ensued when the European Commission released its proposal for a new Data Protection Regulation. The debate focuses on two aspects. Should data protection law apply to pseudonymous data? And if pseudonymous data are within the scope of data protection law, should there be a lighter regime? At the time of writing of this study, the debate is ongoing.

This study argues that data protection law should apply to behavioural targeting, and argues against a lighter regime for pseudonymous data. First, many risks remain, regardless of whether firms tie a name to the information they hold about a person. For instance, surveillance can cause a chilling effect, even if firms collect pseudonymous data. And a cookie-profile that says a person is handicapped or from a

---

[982] See chapter 6, section 4.
[983] Article 2(a) of the Data Protection Directive.

bad neighbourhood could be used for unfair discrimination. Second, a name is merely one of the identifiers that can be tied to data about a person, and is not the most practical identifier for behavioural targeting. For an ad network that wants to track a person's browsing behaviour, or wants to target a person with online advertising, a cookie works better than a name. Third, the online marketing industry processes large amounts of information about people, which carries risks. If data protection law didn't apply, this industry could operate largely unregulated. For these reasons, data that are used to single out a person should be considered personal data.

The fact that data protection law applies doesn't imply that processing is prohibited. It means that the firm using behavioural targeting must process the data fairly, lawfully, and transparently. Of course, merely ensuring that data protection law applies doesn't solve all privacy problems. But at least, data protection law can be used to assess the fairness of processing. The next chapter discusses the role of informed consent in data protection law.

* * *