Available online at www.sciencedirect.com

ScienceDirect

www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation

Frederik J. Zuiderveen Borgesius *

Institute for Information Law (IViR), University of Amsterdam, Amsterdam, The Netherlands

Keywords:
Pseudonymous data
Data protection law
Personal data
Privacy
Online behavioural advertising
Behavioural targeting
Profiling
Cookie
IP address
Tracking

ABSTRACT

Information about millions of people is collected for behavioural targeting, a type of marketing that involves tracking people's online behaviour for targeted advertising. It is hotly debated whether data protection law applies to behavioural targeting. Many behavioural targeting companies say that, as long as they do not tie names to data they hold about individuals, they do not process any personal data, and that, therefore, data protection law does not apply to them. European Data Protection Authorities, however, take the view that a company processes personal data if it uses data to single out a person, even if it cannot tie a name to these data. This paper argues that data protection law should indeed apply to behavioural targeting. Companies can often tie a name to nameless data about individuals. Furthermore, behavioural targeting relies on collecting information about individuals, singling out individuals, and targeting ads to individuals. Many privacy risks remain, regardless of whether companies tie a name to the information they hold about a person. A name is merely one of the identifiers that can be tied to data about a person, and it is not even the most practical identifier for behavioural targeting. Seeing data used to single out a person as personal data fits the rationale for data protection law: protecting fairness and privacy.

1. Introduction

It is hotly debated whether data protection law applies to behavioural targeting. Behavioural targeting, or online profiling, is a type of personalised communication that involves monitoring people's online behaviour and using the collected information to show people individually targeted advertisements. Many behavioural targeting companies say that, as long as they do not tie names to data they hold about individuals, they do not process any personal data, and that, therefore, data protection law does not apply to them. This paper examines whether data protection applies to behavioural targeting, and whether, from a fundamental rights perspective, it should apply.

Behavioural targeting and data protection law are introduced in Section 2 and 3. In Section 4, it is shown that, from a doctrinal perspective, nameless data can be viewed as personal data when a company uses these data to *single out* a person, a view taken by European Data Protection Authorities. Apart from that, Section 5 explains that it is often fairly easy to tie a name to behavioural targeting data. The new Data Protection Regulation and its definitions of 'personal data' and 'pseudonymous data' are discussed in Section 6. Section 7

* Institute for Information Law (IViR), University of Amsterdam, Korte Spinhuissteeg 3, 1012 CX Amsterdam, The Netherlands. Tel.: +31627544960.

E-mail address: f.j.zuiderveenBorgesius@uva.nl.

http://dx.doi.org/10.1016/j.clsr.2015.12.013
0267-3649/© 2016 Frederik J. Zuiderveen Borgesius. Published by Elsevier Ltd. All rights reserved.

argues that data protection law should apply to behavioural targeting. Counter-arguments are considered in Section 8. The conclusion is provided in Section 9: data protection law generally applies – and should apply – to behavioural targeting.

2. Targeted online marketing

Information about millions of people is collected for behavioural targeting. For instance, Facebook collects information about at least 1.5 billion people.[1] Google says it 'reaches 90% of Internet users worldwide.'[2] Some lesser-known companies also process information about many people, such as the Rubicon Project ('600 million'),[3] and AddThis ('1.9 billion').[4]

Many types of companies are involved with behavioural targeting, and the resulting data flows are complicated. In a simplified example of behavioural targeting, an advertising network follows an internet user's behaviour, so it can display individually targeted ads to this user. Ad networks are companies that serve advertisements on thousands of websites. An ad network can track a person's visits to all websites on which it serves ads.

Ad networks often use cookies. These are small text files that website publishers can store on an Internet user's computer. If the cookie contains a unique identifier, website publishers can recognise the visitor's computer. Recognising a visitor's computer is useful, for instance if somebody has included items in a virtual shopping cart. Another example is that of language selected on a website, after which the website publisher can store a cookie on the visitor's computer to ensure that the website will be displayed in the selected language at every subsequent visit by the same individual.

There are several types of cookies. Session cookies are deleted when the user closes his or her browser. Persistent cookies are retained when the user closes the browser or turns off the computer. First party cookies are placed by website publishers. Third party cookies are placed through a website by other parties than the website publisher. Tracking cookies that are used to recognise people contain unique codes, such as 22be6e056ca010062||t = 1392841778|cs = 002213fd48e6bd6f7bf8d99065.[5] If a website publisher uses a cookie to remember a visitor's language settings, the publisher can use a cookie without a unique identifier, for instance FR for French, or EN for English.

When visiting a website, say Newspaper.com, it seems like all parts of the website are presented by one publisher. In reality, various elements on a website are often presented by different companies. The widget of www.newspaper.com showing the weather report might be served from www.weather.com. If ads are displayed on www.newspaper.com, these might be served from www.adnetwork.com, and a Facebook 'Like' button on a website is served by Facebook.[6] All these third parties can store and read their own cookies. During a single website visit, the visitor may receive dozens of third-party tracking cookies.[7]

If www.newspaper.com stores a cookie on a computer, in principle other websites, such as www.gossip.com, cannot read that cookie. Hence, in principle websites can only read their own cookies. However, ad networks have found a way around this system. Third parties such as Weather.com and Adnetwork.com can set and read their own cookies. Hence, Adnetwork.com can set and read its cookies through www.newspaper.com and www.gossip.com, if it serves ads on both websites. In this way, Adnetwork.com can recognise visitors on any website on which it serves advertising. Third-party cookies that are used to follow people around the web are referred to as tracking cookies. The Interactive Advertising Bureau, a trade association for online and mobile advertising, explains that 'cookies are used in behavioural advertising to identify users who share a particular interest so that they can be served more relevant adverts.'[8]

Apart from cookies, behavioural targeting companies use many other tracking technologies. Some technologies, such as flash cookies and other super cookies, are comparable to conventional cookies and involve storing a unique identifier on devices. While people can delete conventional cookies from their computers, super cookies are usually harder to delete. Some companies have used flash cookies to reinstall, or re-spawn, cookies that people deleted: 'zombie cookies'.[9]

Other tracking methods do not rely on storing an identifier on a device. For example, passive device fingerprinting involves recognising a device by analysing the information it transmits. A computer's browser can be recognised by looking at characteristics such as the browser type (e.g. Mozilla Firefox version 38.0.5), its settings, and installed fonts. A device fingerprint is 'a set of system attributes that, for each device, take a combination of values that is, with high likelihood, unique,

[1] Facebook says it had '1.55 billion monthly active users as of September 30, 2015' <http://newsroom.fb.com/company-info/> accessed 24 January 2016.

[2] Google Adwords, 'About the Google Display Network' (publication date unknown) <https://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=57174> accessed 24 January 2016.

[3] <http://rubiconproject.com/ad-buyers-solutions/> accessed 24 January 2016.

[4] 'AddThis offers unparalleled insight into the interests and behaviors of over 1.9 billion web visitors' <http://www.addthis.com/about> accessed 24 January 2016.

[5] This is the unique identifier in a cookie that Doubleclick, Google's ad network, placed on my computer.

[6] Güneş Acar, Brendan Van Alsenoy, Frank Piessens, Claudia Diaz and Bart Preneel, 'Facebook Tracking Through Social Plug-ins' (technical report prepared for the Belgian Privacy Commission) (V. 1.1, 24 June 2015) <https://securehomes.esat.kuleuven.be/~gacar/fb_tracking/fb_plugins.pdf> accessed 24 January 2016.

[7] Chris Jay Hoofnagle and Nathan Good, Web Privacy Census (1 June 2012) <http://ssrn.com/abstract=2460547> accessed 24 January 2016.

[8] Interactive Advertising Bureau, 'A Guide to online behavioural advertising' (Internet Marketing Handbook Series) (2009) <www.iabuk.net/sites/default/files/publication-download/OnlineBehaviouralAdvertisingHandbook_5455.pdf> accessed 24 January 2016, p. 4.

[9] Christian Olsen, 'Supercookies: What You Need to Know About the Web's Latest Tracking Devic' (Mashable) (2 September 2011) <http://mashable.com/2011/09/02/supercookies-internet-privacy/> accessed 24 January 2016. See also Mika D Ayenson, Nathaniel Good, Chris Jay Hoofnagle, Ashkan Soltani, Dietrich J. Wambach, 'Behavioral advertising: The offer you cannot refuse' (2012) *Harvard Law & Policy Review*, 6(2), 273–296.