Social media and cookies: challenges for online privacy

Jo Pierson and Rob Heyman

Jo Pierson is a Professor and Rob Heyman is a PhD Student, both at the Research Centre IBBT-SMIT, Department of Media and Communication Studies, Vrije Universiteit Brussel (VUB), Brussels, Belgium.

Abstract

Purpose – *The advent of Web 2.0 or so-called social media have enabled a new kind of communication, called mass self-communication. These tools and the new form of communication are believed to empower users in everyday life. The authors of this paper observe a paradox: if this positive potential is possible, the negative downside is also possible. There is often a denial of this downside and it is especially visible in social media at the level of privacy and dataveillance. The purpose of this paper is to illustrate this point through an analysis of cookies.*

Design/methodology/approach – *The paper illustrates how mass self-communication in social media enables a new form of vulnerability for privacy. This is best shown by redefining privacy as flows of Personal Identifiable Information (PII) that are regulated by informational norms of Nissenbaum's concept of contextual integrity. Instead of analysing these contexts on a general level, the paper operationalises them on the user level to illustrate the lack of user awareness regarding cookies. The results of the research were gathered through desk research and expert interviews.*

Findings – *The positive aspects of cookies, unobtrusiveness and ease of use, are also the main challenges for user privacy. This technology can be disempowering because users are often hardly aware of its existence. In that way cookies can obfuscate the perceived context of personal data exposure.*

Originality/value – *The research shows how user disempowerment in social media is often overlooked by overstressing their beneficial potential.*

Keywords *Cookies, PII, Privacy, Social media, Web 2.0*

Paper type *Research paper*

1. Mass self-communication and user (dis)empowerment

Tools and technologies for media and communication are undergoing major changes, based on economic transitions and digitisation. This goes together with an intensified state of convergence between the formerly strictly divided sectors of audiovisual media, telecommunication and computer industry. These new media have been described by Punie *et al.* (2009, p. 136) as:

> [. . .] a set of open, web-based and user-friendly applications that enable users to network, share data, collaborate and co-produce content.

Punie *et al.* (2009, p. 136) define these tools as ''social computing'' tools. We propose to use ''social media'' in order to highlight the changing communication processes typified by Castells' concept of ''mass self-communication''. Castells (2009) sees the latter as the novel quality of communication in contemporary society:

- Mass communication because social media can potentially reach a global internet audience.
- ''Self-communication'' because the message production is self-generated, the potential receiver(s) definition is self-directed and the message or content retrieval is self-selected.

However, the different forms of communication (mass media, interpersonal communication and mass self-communication) complement rather than substitute one another.

The notion of ''mass self-communication'' is a good signifier for the techno-dialectic changes taking place in communication and media production. He situates the current ICT and internet landscape as a conflict between the global multimedia business networks that attempt to commodify the internet and the unprecedented autonomy for communicative subjects to communicate at large, labelled as the creative audiences or users:

> Yet, this potential autonomy is shaped, controlled, and curtailed by the growing concentration and interlocking of corporate media and network operators around the world (Castells, 2009)[1].

As indicated by critical scholars like Van Dijck and Nieborg (2009) and Fuchs (2009), these changes in the internet landscape and the claims made on the societal impact are often overrated. Nevertheless we cannot overlook that these new media and internet are becoming an integrated part of everyday life in major parts of Western society, i.e. 47 per cent of American adults are on at least one social network site (Hampton *et al.*, 2011, p. 85). Haythornthwaite and Wellman (2002), Arsenault and Castells (2008) and Hampton *et al.* (2011) also stress how the greater communicative autonomy of the media consumers could help them to become media citizens, and thus restoring the balance of power *vis-à-vis* their would be controllers. This is however only possible if users are empowered, which means that they acquire the necessary know-how to operate social media applications.

1.1 User empowerment

The pros and cons of mass self-communication are linked to notions of respectively ''user empowerment'' and ''user disempowerment''. Empowerment in the general sense is defined as ''enabling people to control their own lives and to take advantage of opportunities'' (van der Maesen and Walker, 2002, p. 24) or in other words ''a process, a mechanism by which people, organisations, and communities gain mastery over their affairs.'' (Rappaport, 1987).

When applying this perspective of empowerment in the realm of new media and mass self-communication, we refer to Mansell (2002):

> [. . .] the implications of the new media are contradictory. Once connected, there are no grounds for simply assuming that citizens will be empowered to conduct their social lives in meaningful ways. There is, therefore, a growing need to examine whether the deployment of new media is consistent with ensuring that the majority of citizens acquire the necessary capabilities for interpreting and acting upon a social world that is intensively mediated by the new media.

Capabilities in this sense are the underpinning of the freedom of people to construct meaningful lives. We therefore define user empowerment in relation to social media as the capability for interpreting and acting on the social world that is intensively mediated by mass self-communication.

So, in order to be more empowered through social media, a user is already presupposed to have mastered these new empowering media technologies. As mentioned before the unprecedented autonomy of media consumers increases the chance of positive and negative consequences, and thus implies more responsibilities and capabilities to foster these new tools of (dis)empowerment. In this article we take a closer look at how disempowerment can take shape within the changing technological landscape of mass self-communication. To answer this question we focus on the issue of online privacy and dataveillance of consumers in relation to social media.

Therefore we take a critical view on online consumer privacy, from the agency perspective of user practices in social media. In this context mass self-communication by users is endorsed by social media companies because the commodification of their expressions is becoming a lucrative business. This phenomenon called ''dataveillance'' challenges user privacy. To illustrate user disempowerment in this realm, we introduce the perspective of contextual integrity. We investigate how the privacy of people is being reconfigured in a social media environment, by way of techniques for collecting personal information for

commercial purposes. In this paper we focus on the most common tool, being ''cookies''. The latter is based on desk research and expert interviews.

2. Challenges for privacy in dataveillance

2.1 Online privacy

The notion of privacy has a long tradition since it became prominent at the end of the nineteenth century, more in particular in the legal academic literature in the USA. The conceptual reconfiguration of this concept in society and law has been nicely framed by metaphors in fiction literature. Solove (2004) discussed the transition from a classic perception of privacy, which is ''the right to be left alone'' to the notion of ''dataveillance'' – as coined by Clarke (1991) as the junction between ''data'' and ''surveillance'' – which is more prevalent in online environments. Solove does not agree with the fact that media and legal academics still view privacy (only) as a right to be protected from totalitarian government and/or other entities prying into the ''secrets'' of ordinary people through surveillance. In this view people still have the power to conceal information. Solove points out that most consumers are even more powerless: they do not know what information has consequences because they do not know who is gathering it, why they gather it, to what cause they gather it and how they gather it. Since a large part of this process is driven automatically, Solove compares this to the secretive and elusive bureaucracy that trials Jozef K. in Kafka's ''The trial''. Jozef K. is unable to understand his trial, or even the reason for his trial and therefore he is more powerless than the main character in the novel *1984* by George Orwell.

2.2 Corporate dataveillance

When applying the notion of dataveillance to commercial online settings, some particular issues need to be stressed. Similar to the traditional media and their audiences, social media generate users that can be sold to advertisers (Smythe, 1977; Bermejo, 2009). In more traditional formats of online display advertising, like banners on web sites, it is quite obvious that one is approached as a consumer. But if somebody is sharing and tagging his pictures with friends on Facebook, this user will normally see himself as someone maintaining social relations with his friends and not as a consumer conveying (very) personal data and content to a US company in exchange for the ''free'' use of its social network services. In this way social relations are commodified intensively, by using this information for more personalised commercial communication and promotion of goods and services.

In this way personal identifiable information (PII) is the currency users pay to get access to social media applications like social network services (SNS). PII refers to information, which makes it possible to either directly or indirectly identify a person or what kind of data belong to that person. We need to stress the importance of indirectly identifiable information. This means that it is often still possible to identify users with anonymised PII, through the coupling with another piece of PII[2]. We involve all kinds of PII because our approach analyses the transaction of information regardless of its anonymity degree.

This commercial exchange does not have to be problematic. It can be a fair deal between the user and the digital service, as long as each party in the deal clearly understands the transactional terms. However users are often not fully aware about what kind of deal they have entered and how to possibly change the conditions (Graber *et al.*, 2002; McDonald and Cranor, 2008). To assess if a deal between a user and a social media service is fair in relation to online consumer privacy, we need a framework that explicitly relates privacy rules to the context. This is done by Nissenbaum' (2004) framework of ''contextual integrity''.

2.3 Contextual integrity

Contextual integrity was developed as a philosophical account of privacy and how personal information is transferred. Barth *et al.* (2006) do not propose this as a full definition of privacy, but as a normative model or framework:

> [. . .] for evaluating the flow of information between agents [individuals and other entities], with a particular emphasis on explaining why certain patterns of flow provoke public outcry in the name of privacy (and why some do not).

Contextual integrity was designed to answer whether a situation contained a privacy breach or not. To achieve this Nissenbaum defined the situation as a context with the following relevant entities: ''the one from whom the information flows, the one to whom the information flows, and the one – the information subject – about whom the information is.'' (Barth *et al.* 2006) These entities perform roles in our society such as patients and physicians or students and teachers. The relationship between these two roles defines what is said. This means a physician may ask about your health in his office and you may expect from him that he keeps this information to himself, unless he needs to share it with a colleague to help with a therapy. In this situation two sorts of information norms define the flow and content of the disclosed information. The relationship ''physician-patient'' defined what kind of information would be exchanged by whom. The norms that govern what is disclosed in a certain situation are norms of appropriateness, and they are context dependent.

Every situation contains a second set of norms which defines to what other contexts or persons this information may flow. These are the norms of distribution. This norm of information flow assesses the transfer of personal information from one party or context to another context. The question is which information from one context may be used in another context. Personal data that are revealed in one specific context will always carry a specific stamp from that context.

A good example is what happened with the social network site Buzz by Google. At the launch on 9 February 2010 Google Buzz automatically, without asking, published openly all personal networks of users based on the people they interact with via Gmail. However e-mail contact lists can hold very private information, like names of personal physicians, romantic relationships or the identities of anti-government activists. They wrongfully assumed that information in one context (of e-mail correspondence like Gmail) could be disclosed without any problem in another setting (of social network relationships like Buzz).

These norms of distribution and appropriateness were later translated into a logical system to assess privacy situations in a more abstract way (Barth *et al.*, 2006). In this logical system agents were able to distinct contexts and roles, but they were also able to see to what contexts the information flowed. This means that the agents in the system are fully aware of their context and therefore capable of controlling their flow of information, which results in full control over their privacy.

Bearing in mind the former examples and the logic system with its fully aware agents, we can now adapt this perspective to the changes in online privacy context. Online consumers are not fully aware of their context compared to the logic system's agents. Consumers still disclose PII, bearing a certain context and spectators in mind. Zuckerberg's infamous quote: ''The privacy is largely false, but for most students, the privacy is good enough'' (*New Yorker*, 2006), points out that there are two contexts, the context perceived by the SNS user and the context as it would have been perceived by the logical agents. We will call the latter ideal context the ''complete context'' and the former the ''perceived context''.

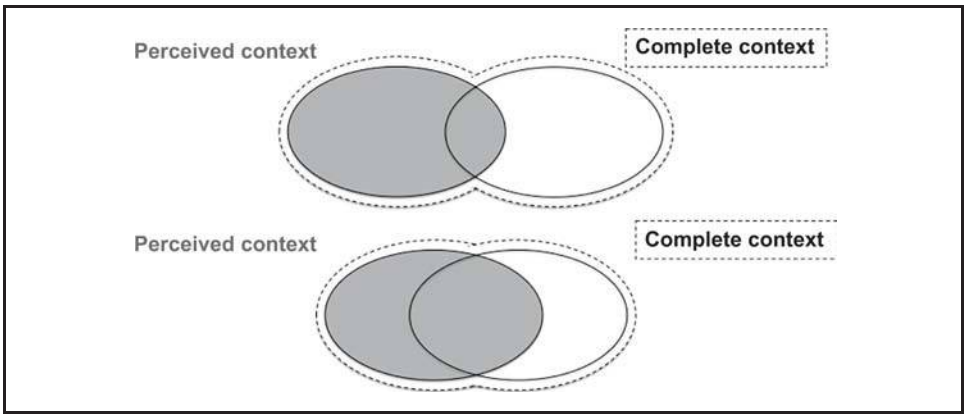
If we draw a map of both these contexts as shown in Figure 1, we can situate where certain technological processes of PII collection happen. We can also define empowerment as the degree of overlap between these two contexts, the bigger the overlap, the more empowered a user is to at least evaluate whether he or she should disclose PII in this context while taking privacy into account.

Papacharissi (2010) frames privacy as a new luxury, and a possible divide:

> As a luxury commodity, the right to privacy, afforded to those fortunate enough to be internet–literate becomes a social stratifier; it divides users into classes of haves and have-nots, thus creating a privacy divide. This privacy divide is further enlarged by the high income elasticity of demand that luxury goods possess (Papacharissi, 2010).

Figure 1 shows this divide in a clear way through the difference of overlap. Elaborating further on Papacharissi's privacy divide, we would like to point out the possibility of drawing

Figure 1 Perceived and complete context



the same sets of perceived and complete contexts for privacy solutions. In order to foster empowerment users do not only need to perceive challenges to their privacy, but that they also need to perceive solutions to this challenges.

3. Analysis of cookies as corporate dataveillance technique

For demonstrating the risk of disempowerment on the level of privacy awareness, we discuss the most familiar online PII collecting tool, the internet cookie. This is a little text file that is placed on the computer by visited web sites. The http cookie was introduced in 1994 in the early browser Netscape Navigator with the purpose of user convenience, namely remembering contents of web shopping carts (Schwartz, 2001). Another important property of cookies is the fact that it is sent automatically, which makes it very unobtrusive.

We will start by defining first party http cookies, third party http cookies and Flash cookies. Next we highlight the occurrence of the http and Flash cookies for commercial use. After this step we will elaborate how these cookies enable new forms of tracking and aggregating PII. Lastly we evaluate our findings against the proposed framework of the perceived and complete context. This evaluation shows that cookies are unobtrusive by nature and are implemented for more reasons than ease of use alone.

We have complemented our findings from the desk research with expert interviews, for additional insights on the technological possibilities and limits of cookies. The selection of experts was based on their technical knowledge of cookies[3].

3.1 First party http cookies

It was already mentioned that cookies were first developed to give web sites a memory. This memory is called a ''state'', and a state is a configuration last used by a user. To remember states, a cookie is able to store the interaction between the user and the web site. This type of information can be for example the user name, the ads clicked and the time spent on each web page (Whalen, 2002). It is important to note that this information is usually encoded. Therefore it is impossible to know what kind of information is being communicated through a cookie (Heyman, 2011). The information is encoded to keep the information safe from malicious parties. Other elements of information are easily added in cookies: every browser conveys information about the operating system, browser, the previous visited url and plug-ins in the header[4] (Eckersley, 2009). The amount of information stored in a http cookie is however limited to 4 kB.

According to our experts, this state information is of increased importance for modern internet applications such as social media and cloud computing, to enrich browsing experience. A typical example is StumbleUpon. The latter uses cookies to remember user preferences in its social tagging service:StumbleUpon uses Cookies to store your

preferences [. . .] You can reset your browser to refuse all cookies or to indicate when a cookie is being sent. *Be aware, however, that the Toolbar may not function properly* [. . .][5](own emphasis).The amount of information added to web sites by users increases and this information is easily stored as state information by cookies. Thus the social layer upon web sites relies on cookie technology although this is a much older innovation than social media.

3.2 Third party http cookies

Third party http cookies differ from first party cookies in two ways. First, they are not placed through the answer to a page request. They are placed through advertisements, images or scripts hosted on a first party web site by a third party server. These cookies do not require a user interaction to be loaded on the user's browser. Second, third party cookies are more persistent than first party cookies, because they are used across different web sites and internet sessions instead of one single visit. Some of these cookies have a default maximum age of more than 30 years.

Third parties may also track users through one by one pixels (Tappenden and Miller, 2009). These pixels are called ''beacons'' or ''gif/web/pixel bugs''. Pixel bugs are impossible to spot for users because they are blank images. It is important to note that third party cookies are not used for advertising only. Social media and other web applications that require much state information through different web sites, such as social tagging, need third party cookies as well to ensure an optimal working service. These social media plugins are however not only used to provide personalisation. For this Roosendaal (2010) refers to the Facebook like button:

> [. . .] the button is a piece of HTML code which includes the request to the Facebook server to provide the image when the web site is loaded. This implies that the button can be used to set third party cookies or to recognize them as well.

It is not only Facebook who gathers data in this way, but also the Google Buzz plugin and Twitter's Tweet button are used to track users (Efrati, 2011). This is problematic because users do not expect to be tracked via the plugins when they are not using them.

3.3 Flash cookies or local shared objects (LSO)

Flash cookies were developed by Macromedia Flash, which became Adobe Flash after Abobe acquired its rival Macromedia in 2005. The first Flash cookie enabled Flash player was Flash Player 6, released in March 2002. This type of cookie was also made to remember states, but there are some important differences. First of all, a Flash cookie was not removable until September 2006, when Flash made the option available through its web site (Adobe Flash Player, n.d.). Second, the cookies were not removable through a browser until January 2011 with the implementation of NPAPI ClearSiteData[6] (Julian *et al.*, 2011). Third, the amount of available information space has grown to 100 kB instead of 4 kB compared to the http cookie. Finally, Flash cookies do not have an expiry date.

Flash cookies are installed via any Flash application on a web site if the user installed the Flash player plug-in. All cookies are accepted by default. Cookie preferences can be changed in the ''Adobe Flash Player Settings Manager'' shown in Figure 2[7].

However, an option to delete all sorts of cookies (http and Flash) can also have negative effects for users. Deleting these PII collecting tools can lead to loss of other saved data, diminishing the convenience of use. For example one can lose all saved game data, which is stored in LSOs, as mentioned by a gamer:

> The problem is, all my saved Flash data cookies [games, etc.] are now completely deleted! [. . .] The only solution seems to be to not delete your browser's cookies as well. Great security feature right? (steelstr45, 2011).

This development also upsets Flash application developers because their applications loose their convenience every time a user cleans his or her browser history.

Figure 2 Image of Adobe Flash Player Settings Manager

Source: Adobe Flash Player (2009)

3.4 Cookie occurrence

Previous research has demonstrated the relevance of cookies, given their increasing usage on the internet. The studies point at three central tendencies. First, the amount of web sites using cookies has risen. Second, the amount of cookies per web site increased, and, third, the more popular a web site is the more cookies are used (see Table I).

Miyazaki (2008) was one of the few to do research that compares two moments in time using the same sample and method. This method has a small bias because many popular newcomers have emerged between 2000 and 2007, like Wikipedia (2001), Myspace (2003) and Facebook (2004). Cookies in general rose 14 per cent and third party cookies rose 17,7 per cent. The occurrence of cookies on a web site increased as well. Miyazaki (2008) found that for web sites that had at least one cookie placed on their home page, the average number of cookies on the home page increased from 2.45 in 2000 (range from 1 to 12) to 8.71 in 2007 (range from 1 to 59), which means a significant increase ($t = 11.8$, $p < 0.01$). The same study also shows that the average number of third party home page cookies grew from 1.57 (range from 1 to 6) to 3.84 (range from 1 to 28), which is again a significant increase ($t = 4.66$, $p < 0.01$).

Table I Overview of studies on cookie occurrence

Source	*Period*	*Sample*	*n*	*http cookies (%)*	*3rd party cookies (%)*	*Flash cookies (%)*
FTC	2000 February- March	335 random e-commerce web sites	335	57		
FTC	2000 February- March	100 most visited web sites	91	78		
Miyazaki	2000	Media Matrix Top 500	406[a]	81.3	32.5	
Miyazaki	2007	Media Matrix Top 500 (from 2000)	406[b]	95.3	50.2	
Soltani *et al.*	2009	Quantcast Top 100	100	98		54
Tappenden and Miller	2009	Alexa Top 100,000	98,006	67.4	54.3	

Notes: The values in the table are percentages of the total sample that deployed either http cookies, third party cookies or Flash cookies. The FTC studies were cited by Miyazaki (2008, p. 21); [a]the result only refers to those 406 web sites from the Media Matrix Top 500 that still existed in 2007 in the follow-up study; [b]the remaining 406 web sites from the Media Matrix Top 500 in 2000 were used

Figures by Tappenden and Miller (2009) for http cookie deployment are significantly lower than those by Miyazaki (2008) and Soltani *et al.* (2009). This is due to the following pattern: the more popular a web site is, the more likely the chance that the site uses cookies to gather information. Since the scale of Tappenden and Miller's research is 1,000 times bigger, it is logical that cookie deployment is lower because they include more less-popular web sites than Soltani *et al.*'s top 100 sample. The same pattern seems to exist in the FTC's random sample vs the FTC's 100 busiest web site sample (Miyazaki, 2008).

4. Dataveillance enabled by cookies

In this section we explain dataveillance techniques enabled by cookies. These techniques solidify or enlarge the complete context and often hamper the perceived context. The first technique is behavioural advertising through third party tracking cookies. Secondly we discuss the use of zombie cookies as a strategy to prevent cookie deletion by users.

4.1 Behavioural advertising

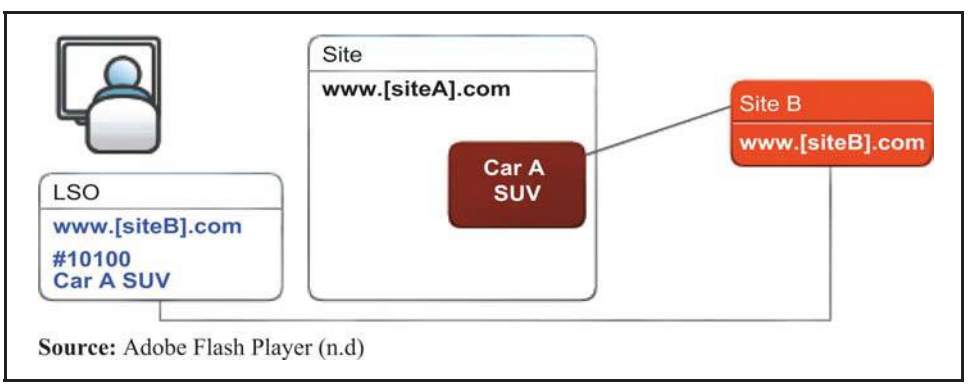
Third party cookies are most often used to profile users and then – depending on the third party – to serve advertisements in line with their behaviour or to aggregate the user data for another party i.e. data miner. This process is explained with Flash cookies, although it is exactly the same for persistent http cookies.

A user surfs to site A and a third party cookie is active on this site in the red box: ''Car A SUV''. This window possibly contains a Flash element or a pixel bug to send a cookie instruction to the browser. In this case the user clicks on this advertisement[8] because it is a straightforward way to explain behavioural advertising. The user now has a cookie of site B, which is accessible for all advertisements of site B. It is shown as an LSO on the left side of Figure 3. In this cookie two values are present, his ID ''10100'' and the fact that he clicked ad ''Car A SUV'', indicating he wanted to know more about the SUV.

The user in this example (Figure 4) leaves site A for site C, which also has an advertisement of Site B portraying another kind of SUV advertisement. The user clicks the advertisement again to learn more about it. His browser receives a new instruction; it needs to add a new value ''Car B SUV'' to the LSO. The user is profiled as a potential buyer of SUVs because he has clicked two advertisements about SUVs.

In Figure 5 the same user surfs to site D, which is also in the same ad network of the third party web site B. This time he will not receive a randomly generated ad. This ad is chosen for users carrying the SUV-interested cookie only. The owner of the tracking cookie can sell this specially chosen ad space to SUV advertisers who wish to buy ads that are guaranteed to be viewed by interested users.

Figure 3 Cookie implementation



Source: Adobe Flash Player (n.d)

Figure 4 Profiling of online behaviour information

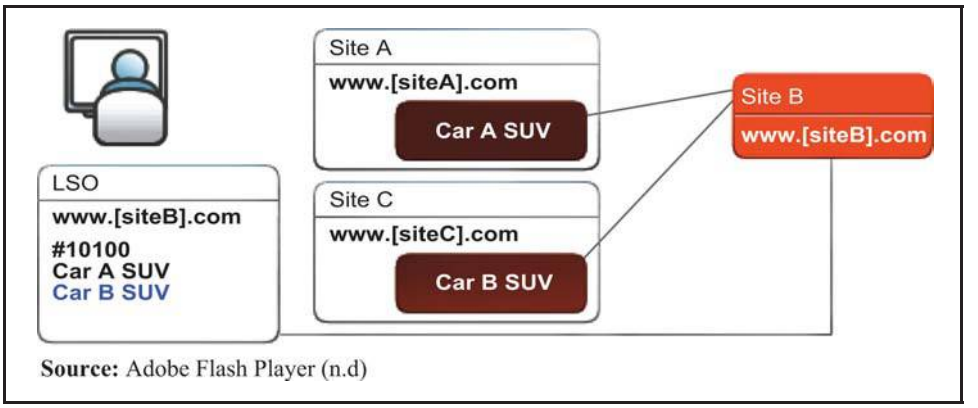
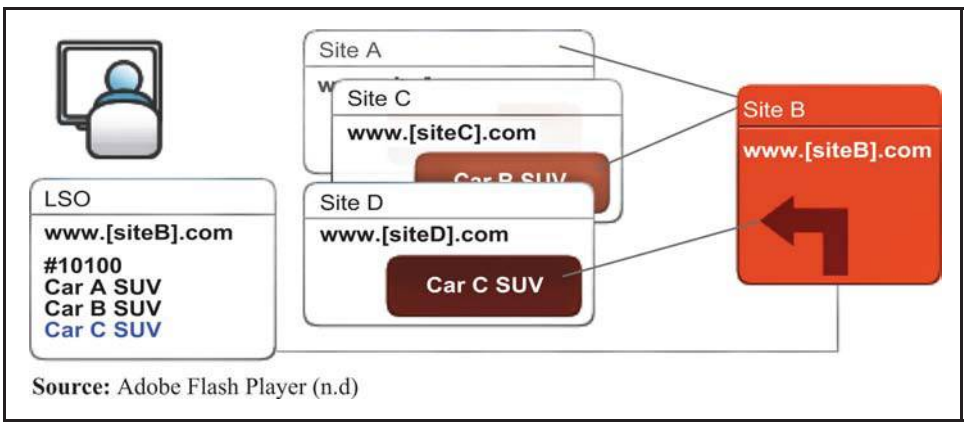


Figure 5 Behavioural targeting



4.2 Cookie doubles

This phenomenon was first noticed and researched by Soltani *et al.* (2009). They found that ''zombie cookies'' were implemented by firms such as United Virtualities after they learned that 30 per cent of internet users were deleting http cookies. The zombie cookie or:

> Persistent Identification Element (PIE) is tagged to the user's browser, providing each with a unique ID just like traditional cookie coding. However, PIEs cannot be deleted by any commercially available adware, spyware or malware removal program[9]. They will even function at the default security setting for Internet Explorer (Soltani *et al.*, 2009).

In order to achieve this kind of persistence the PIE is not one, but two cookies. The first one is the http cookie and the second one is the Flash cookie that revives the http cookie in case of deletion (Soltani *et al.*, 2009). Soltani *et al.* examined the top 100 Quantcast web sites by checking what cookies were added after every single visit to one of these sites. Cookies were then categorized and deleted to ensure correct measurement. This study was done manually to simulate a user who pays a normal visit. They found an overlap between http and Flash cookies:

> Of the top 100 web sites, 31 had at least one overlap between a HTTP and Flash cookie. For instance, a web site might have an HTTP cookie labeled ''uid''[10] with a long value such as 4a7082eb-775d6-d440f-dbf25. There were 41 such matches on these 31 sites (Soltani *et al.*, 2009).

These results do not indicate the use of zombie cookies, but the fact that a Flash cookie served as a backup for the deleted http cookie. This means that 31 out of the 100 examined

sites used cookies that were hard to remove for users unaware of Flash cookies. However they did found zombie cookies on About.com (third party cookie by SpecificClick), Hulu.com (third party cookie by Quantcast) and across domains between AOL (www.aol.com) and MapQuest (www.mapquest.com). A hacker, Samy Kamkar, showed that there are even more technological means to make a cookie persistent[11].

5. Conclusion

For illustrating and better understanding the (dis)empowering characteristics of social media in relation to privacy, we took the case of cookies and the kind of dataveillance practices these corporate profiling techniques enable. The main advantage of cookies – its unobtrusive way to store states – enables a lot of positive and empowering uses when they are used to improve user browsing experience by adding a social or personalised layer. Cookies are built in such a way that information sharing becomes less tedious by removing the need to (re)create and direct data. In this way the threshold to mass self-communicate has been lowered for users. In fact data is now spontaneously generated and automatically communicated with one single mouse click or less.

However our research on cookies has also demonstrated that these forms of corporate dataveillance simultaneously incorporate a risk of user disempowerment. As the value of personal data (PII) grows in the realm of mass self-communication, they are increasingly used as a currency instead of being treated as personal property of the users. Hence, depending on the business model, this trend raises the pressure for internet companies and advertisers to collect a maximum amount of personal data with the least possible threshold. The challenge is then to organise a fair exchange between users and suppliers of digital services. However, we find that the deal is not always balanced. Users are often unaware about the (online) context in which their personal data are being collected, processed and distributed, which leaves them little to no room for control if they wanted to. We find how cookies often aggravate the situation. The contextual integrity perspective has shown that cookies are an implicit technology on both the norms of distribution, i.e. to whom the information is sent, and the norms of appropriateness, i.e. what is communicated. The implicit character of cookies is a disempowering feature to all users who do not have a notion of this phenomenon, because it is no factor in their evaluation of a context. This absence of cookies in the perceived context disables all other actions to empower users in their privacy control. Hence cookies as tools for collecting personal data often obfuscate the context for users, which leads to less overlap between the perceived context and the complete context.

The challenges of user disempowerment and online privacy can be addressed on different levels: on user level, on technology level, and on policy level:

1. First of all, on user level, future research has to take a critical look at the differences on the micro-level of everyday consumer practices between various consumers and consumer groups, in order to assess this ''privacy divide'' in an everyday surveillance environment. This not only means investigating what consumers know about exchanging personal data (awareness), but also what they are able to do (capabilities), what their preferences are (attitudes) and what they effectively do (practices).
2. The outcome on user level needs to be matched with the second level of technological affordances and industry developments with regard to new techniques for tracking and exposing online consumer behaviour, as illustrated by the case of cookies. This can for example be enabled by a socio-technological approach of ''privacy by design''(Cavoukian, 2009, p. 350; Kool *et al.*, 2011, p. 85). The interaction between technological design and user insights engenders the mutual shaping of new technological means for preventing disempowerment and furthering empowerment.
3. On a third level, these user and technological perspectives can also inform policy and (self)regulation. In order to increase the perceived context policy needs to address transparency and awareness, by which users can know and understand the exchange taking place: personal data for ''free'' services. The awareness also refers to knowing about possible countermeasures that a consumer can take (e.g. cultivating privacy

protecting habits, using privacy enhancing technologies, etc.). Finally consumers also require the necessary capabilities to interpret and act upon the social world that is intensively mediated by mass self-communication, in order to convert knowledge into everyday practices. Policy can take initiatives to strengthen the digital literacy capabilities regarding privacy, via school and at home. And if all measures fail there still needs to be a sufficient degree of enforcement. All this is for consideration in the revision of the EU e-Privacy Directive (EU, 2002), which particularises and complements the general EU Data Protection Directive in the electronic communications sector (EU, 1995, p. 31).

In that way citizens and consumers would have the possibility to apply the notion of mass self-communication also on the disclosure of their own personal data. This means that users of social media can have control on the production of their personal data, on who potentially receives these data, and on what is exposed explicitly in their digital footprint or implicit via cookies and other collecting tools. Only by taking these actions on the level of user, technology and policy on local, national and European level, we will be able to keep privacy as a normal good, so as a good that everyone may afford or even as a public good (Papacharissi, 2010).

Notes

1. As indicated by Fuchs (2009). There is however no clear definition of the notion of ''autonomy''.
2. A postal code for example is anonymous because many people share the same postal code, but it becomes more identifiable information when this is coupled with the date of birth. Date of birth is equally anonymous if it is isolated, but the aggregation of PII makes this kind of information pseudonymous or even personally identifiable.
3. The first expert is Professor Marc Langheinrich who teaches information security at the Faculty of Informatics in the Università della Svizzera Italiana (USI) in Italy. The second expert is Prof. Frank Piessens from the Computer Science department at the Katholieke Universiteit Leuven (KUL) in Belgium. Security is his main research expertise. Mr Jeroen van de Gun was our third expert. He is a volunteer from the Dutch Bits of Freedom organisation, an advocacy initiative that raises online privacy awareness. Besides questions on technological issues, we also discussed possible other cookie-like technologies, as well as applications or techniques for detecting and blocking cookies.
4. The header is sent to a server to ask for instructions to load a web page.
5. www.stumbleupon.com/privacy/
6. NPAPI:ClearSiteData allows browsers to clear plugin data, which also includes Flash cookies. This project was a cooperation between contributors of Mozilla, Chromium, Greenbytes, Adobe and Apple.
7. This settings panel is accessible on Adobe.com or by right clicking any Flash content on a given web site, selecting ''settings''.
8. Other types of interaction or even no interaction at all may trigger the beacon, pixel bug or Flash element to send a cookie.
9. Adware is defined as a piece of added software that has a different function from the main software component where it was installed with. Spyware is a special form of adware to gather PII and Malware is adware used to malicious ends. These removal tools can be downloaded as plug-ins or stand-alone applications.
10. uid = user identification.
11. A hacker, Samy Kamkar, made the ''evercookie'' to illustrate what kind of tracking is possible with the current available technology. He makes use of 13 different cookie-like technologies that will reinstall deleted cookies. See Kamkar (2010).

References

Adobe Flash Player (n.d.), ''What are third-party local shared objects?'', Adobe, available at: www.adobe.com/products/flashplayer/articles/thirdpartylso/ (accessed 4 February 2011).

Adobe Flash Player (2009), ''Settings Manager – Global Security Settings panel'', Adobe, available at: www.macromedia.com/support/documentation/en/flashplayer/help/settings_manager04.html (accessed 9 February 2011).

Arsenault, A.H. and Castells, M. (2008), ''The structure and dynamics of global multimedia business networks'', *International Journal of Communication*, Vol. 2, pp. 707-48.

Barth, A., Datta, A., Mitchell, J.C. and Nissenbuam, H. (2006), ''Privacy and contextual integrity: framework and applications'', *2006 IEEE Symposium on Security and Privacy*, p. 15.

Bermejo, F. (2009), ''Audience manufacture in historical perspective: from broadcasting to Google'', *New Media & Society*, Vol. 11 No. 1 and 2-154.

Castells, M. (2009), *Communication Power*, Oxford University Press, Oxford.

Cavoukian, A. (2009), *Privacy by Design – Take the Challenge*, Information and Privacy Commissioner, Toronto, Ontario.

Clarke, R. (1991), ''Information technology and dataveillance'', in Dunlop, C. and Kling, R. (Eds), *Controversies in Computing*, Academic Press, New York, NY.

Eckersley, P. (2009) Vol. 22, ''How online tracking companies know most of what you do online (and what social networks are doing to help them)'', Electronic Frontier Foundation, available at: www.eff.org/deeplinks/2009/09/online-trackers-and-social-networks (accessed 22 December 2010).

Efrati, A. (2011), '''Like'' button follows web users'', *'Like'' button follows web users, The Wall Street Journal*, 8 May, available at: http://online.wsj.com/article/SB10001424052748704281504576329441432995616.html?mod = WSJ_hp_MIDDLENexttoWhatsNewsThird (accessed 22 June 2011).

EU (1995), *Data Protection Directive 95/46/EC*, European Parliament, Brussels.

EU (2002), *Directive on Privacy and Electronic Communications 2002/58/EC*, European Parliament, Brussels.

Fuchs, C. (2009), ''Social software and Web 2.0: their sociological foundations and implications'', in Murugesan, S. (Ed.), *Handbook of Research on Web 2.0, 3.0, and X.0: Technologies, Business, and Social Applications*, Vol. 2, IGI-Global, Hershey, PA, pp. 764-89.

Graber, M.A., D'Alessandro, D.M. and Johnson-West, J. (2002), ''Reading level of privacy policies on internet health web sites'', *Journal of Family Practice*, Vol. 51 No. 7, pp. 642-5.

Hampton, K.N., Goulet, L.S., Rainie, L. and Purcell, K. (2011), *Social Networking Sites and Our Lives*, Pew Research Center's Internet & American Life Project, Washington, DC.

Haythornthwaite, C.A. and Wellman, B. (2002), ''The internet in everyday life'', in Wellman, B. and Haythornthwaite, C.A. (Eds), *The Internet in Everyday Life*, Blackwell, Oxford, pp. 3-41.

Heyman, R. (2011), Interview with Frank Piessens, 2 March.

Julian, R. *et al.* (2011), ''NPAPI: ClearSiteData'', MozillaWiki, available at: https://wiki.mozilla.org/NPAPI:ClearSiteData (accessed 16 June 2011).

Kamkar, S. (2010), ''Evercookie – virtually irrevocable persistent cookies'', 20 August, available at: samy.pl/evercookie/

Kool, L., van der Plas, A., van Eijk, N.A.N.M. and van der Sloot, B. (2011), *A Bite too Big: Dillemma's bij de implementatie van de cookiewet in Nederland*, TNO-rapport, TNO, Delft.

McDonald, A.M. and Cranor, L.F. (2008), ''The cost of reading privacy policies'', *ACM Transactions on Computer-Human Interaction*, Vol. 4 No. 3, pp. 1-22.

Mansell, R. (2002), ''From digital divides to digital entitlements in knowledge societies'', *Current Sociology*, Vol. 50 No. 3, pp. 407-26.

Miyazaki, A.D. (2008), ''Online privacy and the disclosure of cookie use: effects on consumer trust and anticipated patronage'', *Journal of Public Policy & Marketing*, Vol. 27 No. 1, pp. 19-33.

New Yorker (2006), Interview with Mark Zuckerberg, CEO of *Facebook*.

Nissenbaum, H. (2004), ''Privacy as contextual integrity'', *Washington Law Review*, Vol. 79, pp. 101-39.

Papacharissi, Z. (2010), ''Privacy as a luxury commodity'', *First Monday*, Vol. 15 No. 8.

Punie, Y., Lusoli, W., Centeno, C., Misuraca, G. and Broster, D. (2009), *The Impact of Social Computing on the EU Information Society and Economy*, IPTS European Commission – Joint Research Centre, Seville.

Rappaport, J. (1987), ''Terms of empowerment/exemplars of prevention'', *American Journal of Community Psychology*, Vol. 15 No. 2, pp. 121-48.

Roosendaal, A. (2010), ''Facebook tracks and traces everyone: like this!'', Tilburg Law School Research Paper No. 03/2011, available at: http://ssrn.com/abstract = 1717563 (accessed 30 November 2010).

Schwartz, J. (2001), ''Giving web a memory cost its users privacy'', *The New York Times*, available at: www.nytimes.com/2001/09/04/technology/04COOK.html (accessed 2 February 2011).

Smythe, D.W. (1977), ''Communications: blindspot of western Marxism'', *Canadian Journal of Political and Social Theory*, Vol. 1 No. 3, pp. 1-27.

Solove, D.J. (2004), *The Digital Person: Technology and Privacy in the Information Age*, New York University Press, New York, NY.

Soltani, A., Canty, S., Mayo, Q., Thomas, L. and Hoofnagle, C.J. (2009), ''Flash cookies and privacy'', SSRN preprint, available at: http://papers.ssrn.com/sol3/papers.cfm (accessed 22 December 2010).

steelstr45 (2011), ''10.3 and LSO's (flash cookies) being deleted'', Adobe Forums, 22 May, available at: http://forums.adobe.com/thread/855496?tstart = 0 (accessed 16 June 2011).

Tappenden, A.F. and Miller, J. (2009), ''Cookies: a deployment study and the testing implications'', *ACM Transactions on the Web*, Vol. 3 No. 3, pp. 1-49.

van der Maesen, L.J.G. and Walker, A.C. (2002), *Social Quality: The Theoretical State of Affair*, European Foundation on Social Quality, Amsterdam.

Van Dijck, J. and Nieborg, D. (2009), ''Wikinomics and its discontents: a critical analysis of Web 2.0 business manifestos'', *New Media & Society*, Vol. 11 No. 5, pp. 855-74.

Whalen, D. (2002), ''The unofficial cookie FAQ'', Cookiecentral.com, 27 February, availabe at: www.cookiecentral.com/faq/#about

Corresponding author

Rob Heyman can be contacted at: Rob.Heyman@vub.ac.be

This article has been cited by:

1. Barbara Carminati, Elena Ferrari, Marco Viviani. 2013. Security and Trust in Online Social Networks. *Synthesis Lectures on Information Security, Privacy, and Trust* **4**, 1-120. [CrossRef]
2. Eric P. Delozier. 2013. Anonymity and authenticity in the cloud: issues and applications. *OCLC Systems & Services: International digital library perspectives* **29**:2, 65-77. [Abstract] [Full Text] [PDF]