

Novática, founded in 1975, is the oldest periodical publication amongst those specialized in Information and Communications Technology (ICT) existing today in Spain. It is published by **ATI** (*Asociación de Técnicos de Informática*) which also publishes **REICIS** (*Revista Española de Innovación, Calidad e Ingeniería del Software*).

<<http://www.ati.es/novatica/>>
<<http://www.ati.es/reicis/>>

ATI is a founding member of **CEPIS** (Council of European Professional Informatics Societies), an organization with a global membership of about 200,000 European informatics professionals, and the Spain's representative in **IFIP** (International Federation for Information Processing), a world-wide umbrella organization for national societies working in the field of information processing. It has a collaboration agreement with **ACM** (Association for Computing Machinery) as well as with **AdaSpain**, **Ai2**, **ASTIC**, **RITSI** and **Hispalux** among other organisations in the ICT field.

Editorial Board

Ignacio Aguiló Sousa, Guillem Alsina González, María José Escalona Cuaresma, Rafael Fernández Calvo (Chairman), Jaime Fernández Martínez, Luis Fernández Sanz, Didac Lopez Viñas, Celestino Martín Alonso, José Oñofre Montes Andrés, Francisco Noguera Puig, Ignacio Pérez Martínez, Andrés Pérez Payeras, Víktor Pons i Colomer, Juan Carlos Vigo López

Chief Editor

Llorenç Pagés Casas <pages@ati.es>

Layout

Jorge Llácer Gil de Rames

Translations

Grupo de Lengua e Informática de ATI <<http://www.ati.es/pt/lengua-informatica/>>

Administration

Tomás Brunete, María José Fernández, Enric Camarero

Section Editors

Artificial Intelligence
Vicente Botti Navarro, Vicente Julián Inglada (DSIC-UPV), <vbotti.vinglada@dsic.upv.es>

Computational Linguistics

Xavier Gómez Guinovart (Univ. de Vigo), <xgg@uvigo.es>

Manuel Palomar (Univ. de Alicante), <mpalomar@dsi.ua.es>

Computer Architecture

Enrique F. Torres Moreno (Universidad de Zaragoza), <enrique.torres@unizar.es>

José Flich Cardó (Universidad Politécnica de Valencia), <jflich@d9sca.upv.es>

Computer Graphics

Miguel Chover Sellés (Universitat Jaume I de Castellón), <chover@lsi.uji.es>

Roberto Vivó Hernando (Eurographics, sección española), <rvivo@dsic.upv.es>

Computer Languages

Oscar Belmonte Fernández (Univ. Jaime I de Castellón), <belferm@lsi.uji.es>

Inmaculada Coma Tatay (Univ. de Valencia), <inmaculada.coma@uv.es>

e-Government

Francisco López Crespo (MAE), <flc@ati.es>

Sebastià Justicia Pérez (Diputació de Barcelona), <sjusticia@ati.es>

Free Software

Jesús M. González Barahona (GSYC-URJC), <jgb@gsyc.es>

Israel Herráiz Tabernero (Universidad Politécnica de Madrid), <isra@herrai2.org>

Human-Computer Interaction

Pedro M. Latorre Andrés (Universidad de Zaragoza, AIPO), <platorre@unizar.es>

Francisco L. Gutiérrez Vela (Universidad de Granada, AIPO), <fgutierrez@ugr.es>

ICT and Tourism

Antóns Aguiar Maldonado, Antonio Guevara Plaza (Univ. de Málaga), <aguiar.guevara@lcc.uma.es>

Informatics and Philosophy

José Ángel Olivás Varela (Escuela Superior de Informática, UCLM), <joseangel.olivas@uclm.es>

Roberto Feltrero Oreja (UNED), <rfeltrero@gmail.com>

Informatics Profession

Rafael Fernández Cano (ATI), <rfcanvo@ati.es>

Miquel Sàrries Griño (ATI), <miquel@sarries.net>

Information Access and Retrieval

José María Gómez Hidalgo (Optenei), <jmgomez@optenei.com>

Manuel J. María López (Universidad de Huelva), <manuel.mana@dieia.uhu.es>

Information Systems Auditing

Marina Touriño Troitillo, <marinatourino@marinatourino.com>

Manuel Palao García-Suelto (ATI), <manuel@palao.com>

Knowledge Management

José Baiget Solé (Cap Gemini Ernst & Young), <josbaiget@ati.es>

Language and Informatics

M. del Carmen Ugarte García (ATI), <cugarte@ati.es>

Law and Technology

Isabel Hernando Collazos (Fac. Derecho de Donostia UPV), <isabel.hernando@ehu.es>

Elena Davara Fernández de Marcos (Davara & Davara), <edavara@davara.com>

Networking and Telematic Services

José Luis Marzo Lázaro (Univ. de Girona), <joseluis.marzo@udg.es>

Juan Carlos López López (UCLM), <juancarlo.lopez@uclm.es>

Object Technology

Jesús García Molina (DIS-UM), <jmolina@um.es>

Gustavo Rossi (LPIA-UNLP Argentina), <gustavo@sol.info.unlp.edu.ar>

Personal Digital Environment

Andrés Marín López (Univ. Carlos III), <amarin@it.uc3m.es>

Diego Gachet Páez (Universidad Europea de Madrid), <gachet@uem.es>

Real Time Systems

Alejandro Alonso Muñoz, Juan Antonio de la Puente Alfaro (DIT-UPM), <almonso.puente@di.upm.es>

Robotics

José Cortés Arenas (Sopra Group), <joscortare@gmail.com>

Juan González Gómez (Universidad Carlos III), <juan@iearobotics.com>

Security

Javier Arellano Bertolín (Univ. de Deusto), <jarellito@deusto.es>

Javier López Muñoz (ETSI Informática-UMA), <jlm@lcc.uma.es>

Software Engineering

Javier Dolado Cosin (DLSI-UPV), <dolado@lcc.upv.es>

Daniel Rodríguez García (Universidad de Alcalá), <daniel.rodriguez@uah.es>

Students' World

Federico G. Mon Trotti (RITSI), <gnu.fede@gmail.com>

Mikel Salazar Peña (Asoc. de Jóvenes Profesionales, Junta de ATI Madrid), <mikelbo_uni@yahoo.es>

Technologies and Business

Didac López Viñas (Universitat de Girona), <didac.lopez@ati.es>

Francisco Javier Cantais Sánchez (Indra Sistemas), <fjcantais@gmail.com>

Technologies for Education

Juan Manuel Dodero Beardo (UC3M), <dodero@inf.uc3m.es>

César Pablo Córcoles Briongo (UOC), <ccorcoles@uoc.edu>

Technological Trends

Alonso Álvarez García (TID), <aad@tid.es>

Gabriel Martí Fuentes (Interbits), <gabim@atnet.es>

University Computer Science Teaching

Cristóbal Pareja Flores (DSIP-UCM), <cpareja@sip.ucm.es>

J. Ángel Velázquez Iturbide (DLSI I, URJC), <angel.velazquez@urjc.es>

Web Standards

Encarna Quesada Ruiz (Virati), <encarna.quesada@virati.com>

José Carlos del Arco Prieto (TCP Sistemas e Ingeniería), <jcarco@gmail.com>

Copyright

© ATI 2013

The opinions expressed by the authors are their exclusive responsibility

Editorial Office, Advertising and Madrid Office

Plaza de España 6, 2ª planta, 28008 Madrid

Tfn. 91 4029391; fax. 91 2039395 <novatica@ati.es>

Layout and Comandant Valenciana Office

Av. del Reino de Valencia 23, 46005 Valencia; Tfn. 963740173 <novatica_prod@ati.es>

Accounting, Subscriptions and Catalonia Office

Via Laietana 46, ppal. 1ª, 08003 Barcelona

Tfn. 93 41 25235; fax 93 41 27713 <secregen@ati.es>; <novatica.subscriptions@atinet.es>

Aragón Office

Lagasca 9, 3-B, 50006 Zaragoza Tfn./fax 976235181 <secreara@ati.es>

Andalucía Office

<secreand@ati.es>

Galicia Office

<secregal@ati.es>

Advertising

Plaza de España 6, 2ª planta, 28008 Madrid.

Tfn. 91 4029391; fax. 91 2039395 <novatica@ati.es>

Legal deposit: B 15.154-1975 -- ISSN: 0211-2124; CODEN NOVAEC

Cover Page: Dancing House - Concha Arias Pérez / © ATI

Layout Design: Fernando Agraeta / © ATI 2003

editorial

Novática: Reaching beyond International Borders

> 02

Didac López Viñas, President of ATI

From the Chief Editor's Pen

Privacy: Our Contribution to a High-Level Debate in the Digital Age

> 02

Llorenç Pagés Casas, Chief Editor of Novática

monograph

Privacy and New Technologies

Guest Editors: Gemma Galdon Clavell and Gus Hosein

Presentation. Privacy, Technology and Policy: Social Networks, Data Mining and Surveillance

> 04

Gemma Galdon Clavell, Gus Hosein

Privacy and Surveillance Primer

> 11

Aaron Martin

European Data Protection and the Haunting Presence of Privacy

> 17

Gloria González Fuster, Rocco Bellanova

Secrecy Trumps Location: A Short Paper on Establishing the Gravity of Privacy Interferences Posed by Detection Technologies

> 23

Mathias Vermeulen

Surveillance Technology and Territorial Controls: Governance and the 'Lite Touch' of Privacy

> 26

Darren Palmer, Ian Warren

Google: Navigating Security, Rights to Information and Privacy

> 32

Cristina Blasi Casagran, Eduard Blasi Casagran

Human Traces on the Internet: Privacy and Online Tracking in Popular Websites in Brazil

> 37

Fernanda Glória Bruno, Liliane da Costa Nascimento, Rodrigo José Firmino, Marta M. Kanashiro, Rafael Evangelista

Social Networks and the Protection of Personal Information. When Privacy Is Not Perceived As a Right

> 44

Massimo Ragneda

Privacy and Body Scanners at EU Airports

> 49

Joan Figueras Tugas

Fernanda Glória Bruno¹,
Liliane da Costa Nascimento¹,
Rodrigo José Firmino², Marta
M. Kanashiro³, Rafael Evangelista³

¹Federal University of Rio de Janeiro (UFRJ),
Brazil; ²Pontifícia Universidade Católica do
Paraná (PUCPR), Brazil; ³State University
of Campinas (Unicamp), Brazil

<bruno.fernanda@gmail.com>,
<liliane.line@gmail.com>,
<rodrigo.firmino@pucpr.br>,
<mmk@unicamp.br>,
<rae@unicamp.br>

1. Introduction

The intensive monitoring and data gathering of people's online activities is the core element of a new business model for the Internet, in which we pay for information and services offered as if they were for free with personal information [1]. These data, collected throughout the Internet as we navigate from website to website, are successively stored, processed, and then (as it is claimed by the tracking industry) used to predict our future behavior as consumers and deliver us tailored advertisements. These processes have remained largely oblivious to Internet users, who usually do not know when or by whom their information is being collected, nor to which purposes it will be used.

This everyday monitoring of peoples' ways of life has given rise to a series of debates about privacy. It is known that several forces are driving changes in the meaning, practices, contours and experience of privacy on different levels: conceptual, commercial, political, social, subjective etc. These processes are not homogeneous and in some cases, they have been happening in spite of actors' consciousness about its reach and complexities. In this context, abstract definitions of privacy do not allow us to grasp what is at stake [2][3]. We consider it essential to observe and identify, from a sociotechnical perspective [4], the actions and practices through which this renegotiation has been performed.

This paper aims to contribute to this approach, by focusing on a specific field of practices and monitoring technologies related to online activities in Brazil. It presents the results of a preliminary survey whose main objective is to reveal and discuss how online tracking is taking place in seven websites ranked among the most visited by Brazilians.

Therefore, we start by briefly presenting the general framework of online tracking, with its activities of data gathering and processing. Hereupon, we present the survey methods and results. Then, we situate public debate and

Human Traces on the Internet: Privacy and Online Tracking in Popular Websites in Brazil

Abstract: *The Internet has made our actions and lives increasingly traceable. Data about our habits, preferences, tastes and interests can be easily collected, stored and processed on the web, which has an inestimable value to several private parties interested in predicting our online and offline behaviors. This paper aims to shed some light on how this data collection takes place in popular websites in Brazil. In order to do so, we investigate tracking mechanisms used in five websites and two social network websites. We identify the HTTP cookies, Flash cookies and web beacons used in each of them. The trackers are analyzed in quantitative and qualitative terms, based on the practices of the companies responsible for operating them. Based on that, we discuss the contributions and limitations on the notion of privacy related to the use of trackers in the Brazilian context.*

Keywords: *Brazil, Cookies, Online Marketing, Online Tracking, Privacy, Web Beacons.*

Authors

Fernanda Glória Bruno is an Associate Professor in the graduate program in Communication and Culture at the Federal University of Rio de Janeiro (UFRJ), Brasil. She has a PhD in Communication from the Federal University of Rio de Janeiro (Doctoral Research Fellow at Paris Descartes University, Sorbonne, 2000). She is also a National Counsel of Technological and Scientific Development/CNPq researcher and founding member of the Latin America Network of Surveillance, Technology and Society Studies. In 2010-2011 she was a visiting researcher at Sciences Po (CERL and Médialab), Paris (France).

Liliane da Costa Nascimento is a PhD Candidate in Communication and Culture at the Federal University of Rio de Janeiro (UFRJ), Brasil. She holds a master's degree from the same university. Her researches interests include surveillance technologies, cyberculture and social control. She is a member of the Latin American Network of Surveillance, Technology and Society Studies.

Rodrigo José Firmino is a lecturer in the Postgraduate Program of Urban Management at the Pontifícia Universidade Católica do Paraná, in Curitiba, Brazil. Rodrigo is also a CNPQ Research Fellow. He worked as a Postdoctoral Fellow at University of São Paulo, Brazil, researching the co-development of urban and technological strategies for cities in developing countries. He obtained his Ph.D. on Urban Planning from the School of Architecture, Planning and Landscape at the University of Newcastle upon Tyne, U.K. He also holds an MPhil in Architecture and Urbanism from the University of São Paulo, and trained as an architect and planner at State University of São Paulo. Rodrigo is one of the founding members of the Latin American Network of Surveillance, Technology and Society Studies.

Marta M. Kanashiro is a professor in the Postgraduate Programme in Scientific and Cultural Communication at the University of Campinas (Unicamp), Brazil. Marta is also a researcher in the Laboratory of Advanced Studies on Journalism and in the research group Knowledge, Technology and Market (CTeMe), both in the same university. She obtained his Ph.D on Sociology in University of São Paulo (USP) and his main interests focus on Science Communication and Sociology of Technology from a critical perspective on new developments in science or new technologies. She has been studying surveillance technologies in Brazil, since 2001. Her research includes an investigation about the use of CCTVs in the city center of São Paulo (2002-2005) and the use of biometrics in the Brazilian Id Card (2007-2011). She is a Founding member of the Latin American Network of Surveillance, Technology and Society Studies.

Rafael Evangelista is a professor in the Postgraduate Program of Science and Cultural Communication and a researcher in the Laboratory of Advanced Studies on Journalism at the State University of Campinas (Unicamp), Brasil. He holds a Ph.D in Anthropology and an MPhil in Linguistics both from Unicamp. His main research interests and publications are on the cultural and social aspects of Internet and network society, free and open source software communities, commons-based peer production, and science communication. He is a Member of the Latin American Network of Surveillance, Technology and Society Studies.

regulatory framework on privacy and data protection in Brazil and finally, we raise privacy questions related to the subject of control over personal information.

“ Through this process, which is opaque to the great majority of users, companies can place small files at visitors' computers and track their web surfing activity over time and across sites ”

2. The Growing Value of Personal Information

While in the offline world data gathering is frequently related to a single domain or activity, on the Internet any action can be traced. Over time, it has enabled huge databases to emerge, containing data about almost any aspect of people's lives. The question then becomes how to cope with this large amount of information and extract value from it, which has been answered with the emergence of sophisticated computational and statistical technologies.

The consumer profiling industry is largely based on methods such as profiling and Knowledge Discover in Databases (KDD), also called Data Mining (DM). In short, they enable scouring databases for hidden patterns without the need to draw from hypothesis formulated by a human being [5].

In general, databases can be analyzed in two main ways, each one having different privacy implications. Descriptive practices consist of scanning the data available to retrieve information about the database as a whole. Predictive tasks, in their turn, enable the making of guesses about a future condition. Once analysts have segmented their databases, they claim that they become able to anticipate the future conduct of the individuals previously classified, or predict the behavior of others about whom they have almost no information.

As it is well known, the Internet works based on a constant exchange of information – users are always sending requests to the websites they access and downloading data from its host servers. Through this process, which is opaque to the great majority of users, companies can place small files at visitors' computers and track their web surfing activity over time and across sites. These pieces of software, called cookies, collect the information that will enable behavioral targeting and influence the advertisements one sees when visiting a website. Other trackers, such as web beacons, can also be used to this purpose¹. What our clicks can reveal is attached to the content available on the website we visit².

Thus, the knowledge about individuals is restricted and no single entity collects all personal information on the web. It drives efforts to collect data across websites, using cookies and also other tracking methods³. As Peter Eckersley argues, "the core function of

the cookie is to link what you do on Web site A to what you do on Web site B" [11].

In the online market case, it is done by ad networks, which manage the placement of advertisements in a group of websites. Advertisements are not stored in the publisher's server, but in an ad server, which delivers the advertisements to the websites visitors⁴ are seeing. It enables companies to track our online habits and preferences through different and not related websites, which has important privacy implications⁵.

3. Data Collection and Tracking in Brazilian Websites

There are several reasons to look at online tracking in Brazilian sites. Brazil has the largest online population in Latin America. Worldwide, it is the eighth country in number of Internet users. According to a survey carried out in 43 countries [12], there are 41.5 million active Brazilian users⁶, who spend, in average, 24.3 hours online per month, 2 hours more than the global average. In addition, online population in Brazil has grown by 19% from March 2010 to March 2011, a number 8% bigger than the increase of worldwide online population in the same period.

Two statistics brought out by this research are of particular interest to the purposes of this article. The first one shows that the habit of online purchasing is becoming more popular among Brazilians, which moves tracking industry forward. In December 2010, 69.6% (seven out of ten Brazilian users) visited retail websites, a rate that is the highest in the region. A comparison with the same month in 2009 shows a 9% growth. The second statistic we highlight is about social networking usage. Brazil is the fifth largest social networking population in the world, and social networking reach in the country is 85.3%, 14.8% higher than the worldwide average. According to another research, the usage of social network websites is the third biggest online activity for users in Brazil [13].

Following these data, we have examined two social network websites (Orkut and Facebook⁷) and five websites listed among the fifteen most visited in the country⁸ (Terra⁹, UOL¹⁰, Globo.com¹¹, Yahoo!Brasil¹² and YouTube¹³). We have looked for HTTP cookies, Flash cookies and web beacons used in each of them and analyzed the trackers found in quantitative and qualitative terms, based on the practices of the companies responsible for their operation.

4. Methods

The data for this survey were collected in February 2010¹⁴. The methodology employed has been partially drawn up from an analysis conducted by The Wall Street Journal [14] and varies according to the kind of trackers analyzed.

The search for HTTP cookies was done basically in three steps: (1) by deleting all web browser's cookies; (2) by visiting, on average, 30 different pages in each domain; (3) and finally, checking the HTTP cookies stored in the web browser. Flash cookies analysis followed a similar method, but since these trackers are not stored in the web browser, we used Adobe Flash Player: Settings Manager¹⁵, an Adobe Panel controlled via webpage that shows a list of domains storing flash content in a computer. At last, to conduct the web beacon's examination, we used Gosthery¹⁶, a browser tool that notifies the user about the presence of web beacons and gives information about companies operating them. During the process, we never logged in and always took care not to access external links.

In social network websites, we examined five social applications in each website, chosen according to their popularity¹⁷. It is worth mentioning that, in this case, we conducted our analysis logged in as registered users. However, since motivations to track can decrease once you are identified to the system, we focused on these external trackers operated by social platform developers or companies working for them.

After their identification, we analyzed each cookie domain, in order to know which company had set them. Sometimes, the domain's name gave us this information, but when it was not explicit, we had to search for it using Robtex¹⁸, a tool that provides domain name consulting. Then, we visited each company's website to learn about the service it provides. We also searched for their privacy policies and checked if they offered an opt-out mechanism for their trackers. Based on this information, we classified the trackers found.

5. Results

On the five websites analyzed, we have found a total of 334 HTTP cookies (see **Figure 1**) (174 set by third parties and 53 set by different domains), 3 Flash cookies and 25 web beacons (considering only the exclusive ones). The examination of trackers distribution across sites has revealed that the number of third party HTTP cookies found on Terra,

“ After their identification, we analyzed each cookie domain, in order to know which company had set them ”

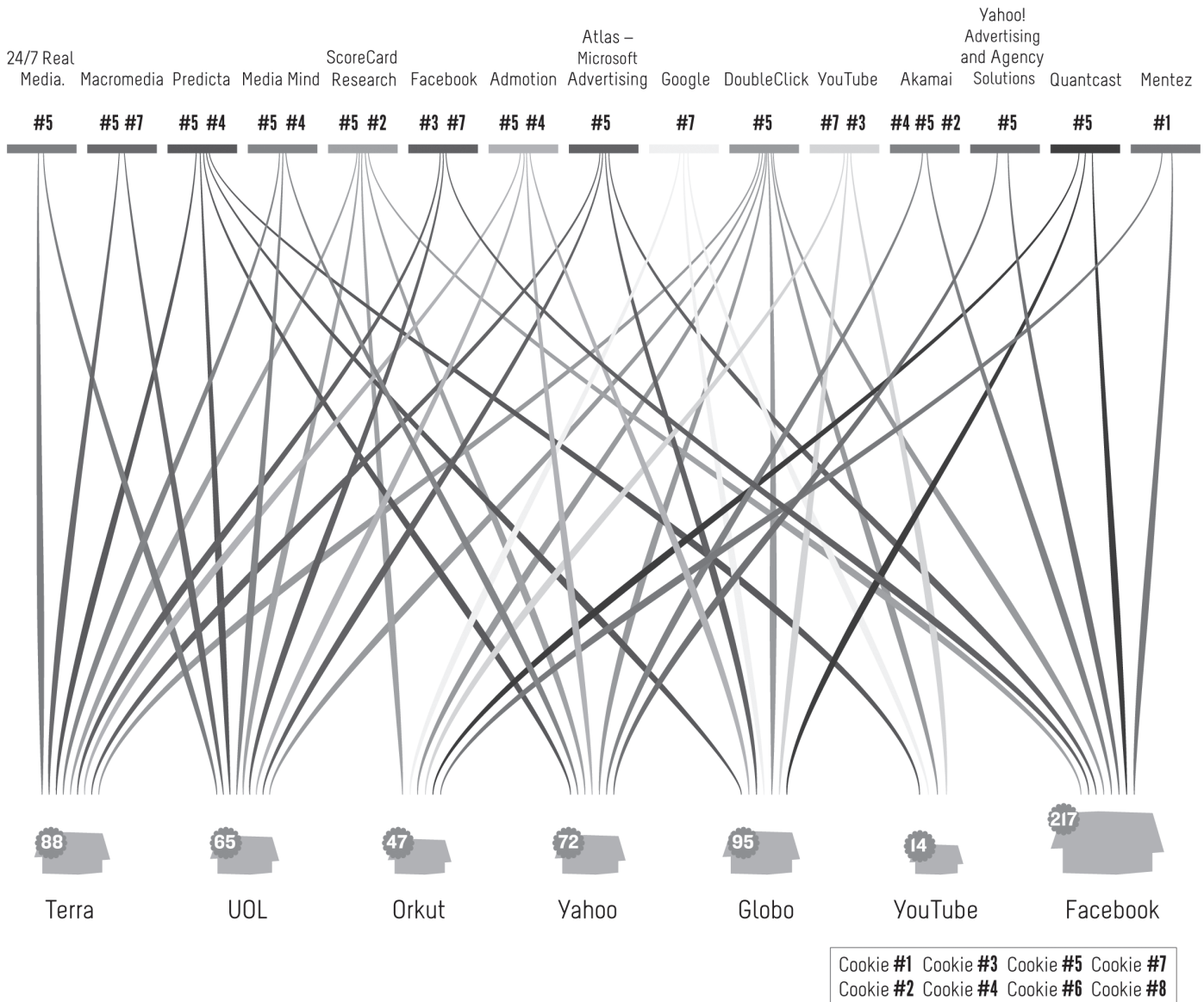


Figure 1. Characteristics of the Most Common HTTP Cookies, the Websites Where They Were Found and the Companies that Set Them¹⁹.

UOL and Yahoo is 4.2 times bigger than the number of cookies found on Globo.com and YouTube. When it comes to web beacons, UOL and YouTube have used less trackers of this type (1 and 3, respectively) than Yahoo, Terra and Globo.com (which have set 6, 7 and 8 web beacons, respectively).

We also examined these five websites' privacy policies to understand how privacy is considered and informed to the public in these documents. Except for UOL, all policies mentioned that cookies from third parties are allowed. In these cases – all policies observed – data collection and use is governed by the partners' policies. This claim resounds as an attempt to reallocate responsibility, putting

on users' shoulders the obligation to analyze all "chain of policies" and take a stand in each situation. Given the expressive number of external cookies revealed by our survey, and how obscure this process may be, we can consider this practice as a source of privacy concerns. On the one hand, the user is charged with the responsibility and work to manage its own privacy. On the other hand, there is no guarantee that he/she can count on the required transparency to negotiate privacy in an autonomous way.

Almost all policies have also emphasized that cookies do not collect personal information and that companies can share aggregated data with commercial partners. In fact, this

perspective can be put in check if we consider that the notion of personal data is being challenged by the fact that digital databases can be easily re-identified through pieces of information other than names [15]. Moreover, the lack of identification, in a traditional sense, of a specific individual, does not prevent consumer's engagement. Even anonymous, databases can be used to classify and influence personal conduct, as well as to sort the opportunities offered to the individuals [16].

Social network websites analysis has revealed an expressive asymmetry in the number of cookies found. On Orkut, we have identified 43 HTTP cookies, while on Facebook, there

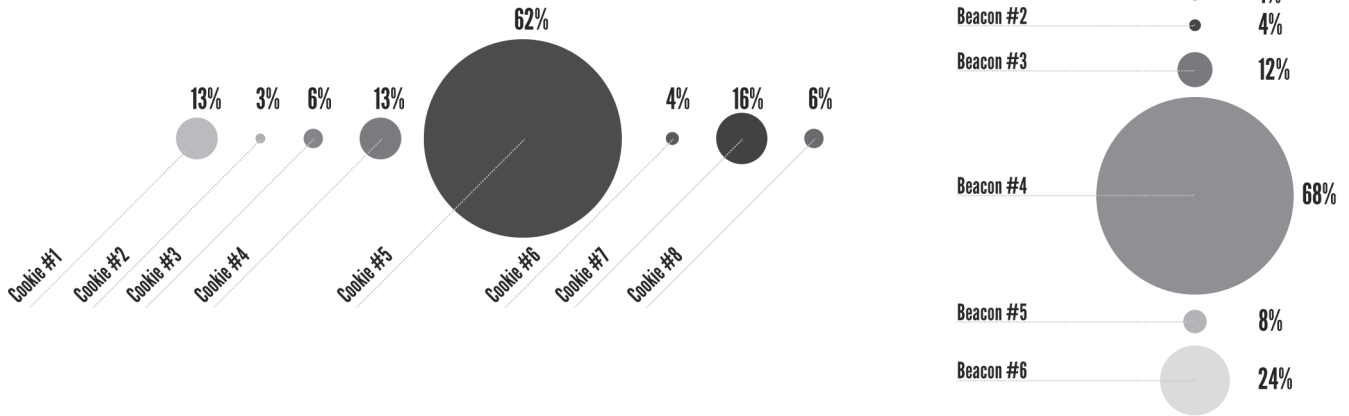


Figure 2. Cookies (on the left) and Beacons (on the right) Occurrence per Type.²¹

were 217. The number of web beacons, in turn, has ranged from 15 on Facebook to 18 on Orkut (considering only the exclusive ones)²⁰. The expressive number of companies operating trackers called our attention. In only seven websites, we identified 69 companies operating cookies and 23 companies operating web beacons. Only four of these companies are located in Brazil: Predicta, Navegg, Zura! and Boo-Box. We also observed the predominance of companies from the online marketing field – 62% of those setting cookies and 68% of those setting beacons provided services such as advertisement serving and optimization (see Figure 2). This evidence indicates the importance of discussing and identifying

how most parts of personal data storage, monitoring and classification are tied to the dynamics of online marketing.

We have ranked companies responsible for the trackers according to the number of websites in which they appeared (see Figure 3). As previously highlighted, the amount of data a company can collect increases with the number of websites in which they deliver advertisements. Doubleclick, Google’s branch on online marketing, has set trackers in all websites analyzed and leads the ranking for HTTP cookies, followed by Predicta, a national company which have set cookies in five out of seven websites. Google Analytics, in turn,

leads the ranking of web beacons, appearing in all websites analyzed, followed by Google’s Ad Sense, which have set web beacons in three out of seven websites.

Our investigation showed that 19% of companies that have set cookies did not offer a privacy policy and 46% did not provide an opt-out option. Among those who operate web beacons, the number of companies that did not offer a privacy policy drops to 4%. We also have found that, among companies that have set beacons, 32% are Network Advertising Initiative (NAI)²² members, while 56% are certified by TRUSTe²³ or Safe Harbor²⁴ signatories. Among companies operating cook-

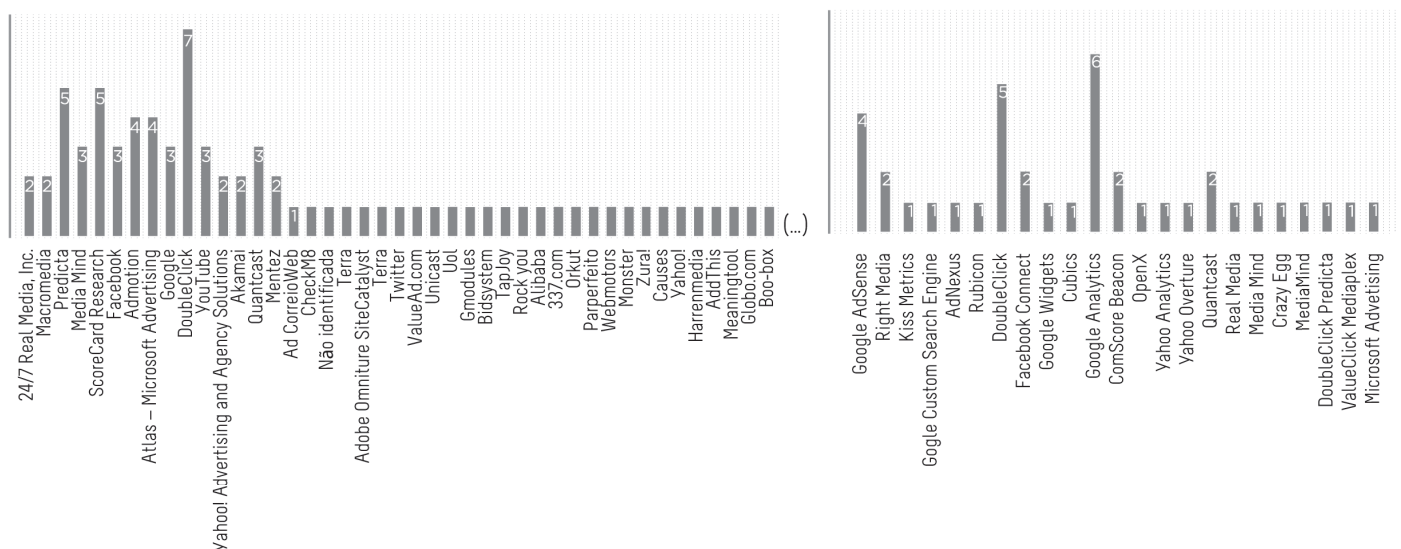


Figure 3. Number of Websites in which We Have Found Cookies (top) and Beacons (bottom) Set by each Company.

“ Social network websites have helped to raise awareness about the reach and visibility of data its users usually post on their profiles, mainly due to Orkut mass adoption in the country ”

ies, these numbers drop to 19% and 23%, respectively; 67% of companies operating cookies and 32% of those operating beacons are not NAI, neither TRUSTe nor Safe Harbor signatories.

It is also worth highlighting that, except for a few cases, companies have not been explicit about the usage and opt-out options for web beacons and Flash cookies. In general, their privacy policies employed only the term cookie. Another fact with potential privacy implications, is that, sometimes, the given opt-out mechanism consisted in setting a cookie to indicate the user's option of not being tracked. As this cookie comes from the same domain that trackers cookies came from, it makes the process uncertain concerning the ways companies manage personal data.

6. Public Debate and Regulatory Framework

Despite all tracking evidences discussed above, public debate about online privacy in Brazil is still incipient. Social network websites have helped to raise awareness about the reach and visibility of data its users usually post on their profiles, mainly due to Orkut mass adoption in the country. However, institutional audiences and issues like behavioral targeting, data collection and process have been out of agenda, except for a few cases which gained some relevance and media coverage in the country.

In 2010, Velox and Speedy, well known Brazilian Internet service providers, started to test a spy tool able to register users navigation data, including pages visited and time spent in each one. This software has been developed by Phorm, a British company whose monitoring technologies have already raised controversy in several countries²⁵. Oi, Velox controller, and Phorm have also announced an association with popular Brazilian Internet portals, including Terra, UOL and Estadão, which could target advertisements to their visitors based on information collected by the spy software.

Together, Oi and Telefonica provide 55% of broadband Internet access in the country. But as they act in different geographical areas, the Administrative Council of Economic Defense (CADE) approved Oi partnership with Phorm without restriction in October 2011. Telefonica partnership with Phorm was also approved two months later²⁶. The Department of Consumer Protection and Defense, part of the Brazilian Ministry of Justice, has filed an administrative lawsuit against Oi in June 2010, to examine evidences of privacy intrusion. The company

was invited to give some explanations about the software, but it did not respond. The lawsuit is still under analysis.

Currently, the most modern provision about data protection in Brazilian law is article 43 of the Consumer Defense Code, according to which people should be informed about their inclusion in databases and have open access to registered information about them [28]. The Complementary Law of Bank Secrecy (CL 105/201) and Habeas Data Law (9507/97) also contain legal provisions about data protection. Besides, the Brazilian Federal Constitution proclaims that "*privacy, private life, honor and image of people are inviolable*".

However, Brazil still does not have a dedicated personal data protection law. While in developed countries these laws started to emerge in the 1970s, only now legislators are discussing the matter in Brazil. The country is late even when compared to other Latin American nations: Chile consolidated a law about the issue in 1999, followed by Argentina, whose personal data protection law is the only one in the region in compliance with European rules. In 2010, Mexico also approved its Federal Law of Personal Data Protection.

Aiming to surpass this delay, the Ministry of Justice put a draft bill proposition before a public hearing in November 2010²⁷. Inspired by the European Data Protection Law (95/46/EC), it aims to update citizens' rights about the current context of technological acceleration and progressive surveillance. It does not forbid data processing, admitting that, in some cases, it has a social value and can be useful to individuals. But it set the rules and conditions under which institutions can do it, and provides that people should not be submitted to decisions which heavily affect them, based only on an automated data processing. The proposition also gives people the right to ask institutions about techniques employed and patterns informing data processing. It also provides that they should consent and be notified of data collection in the very moment it takes place.

Other important draft bill proposition, also put before public hearing by the Ministry of Justice is the Internet Law Framework, which aims to define the technical structure and values that should guide Internet development in the country. It respects data protection directives and provides that people should be informed about how their data will be used.

According to that, data processing, distribution and third parties access should depend on the users consent. This proposition has received more than two thousand contributions during the public hearing and it was presented to the National Congress in August 2011.

7. Privacy: Strength and Limitations

Taking the discussion above into account, is privacy a useful way to frame inquiries about data collection and usage as observed in Brazilian websites? Privacy is proclaimed to be a fundamental value for freedom and democracy [2][17]. On the other hand, some philosophers, sociologists and even legal scholars claim it is an inconsistent concept to face contemporary nuances of personal information flow [3]. The most common criticisms are directed towards the concept broadness and individualistic dimension²⁸.

One of the most prominent definitions of privacy is that of control over personal information. In Alan Westin's ([19], p. 7) words, "*privacy is the claim of individuals, groups, or institutions to determine for themselves when, how and to what extent the information about them is communicated to others*". This concept of privacy, which has its roots in liberalism, is widely quoted as evidence of privacy as seclusion and individualism. As Steeves ([20], p. 11) says, Westin's analysis contains social aspects, but they fade as the focus of the author's argumentation shifts to the flow of information. Thus, privacy becomes antisocial and it is finally located in the individual's unilateral control against disclosure of his/her information.

As we have seen, the idea of individual control over personal information is recurrent in the online tracking discourse. Privacy policies and corporations claim that individuals can block cookies or opt-out while social actors and even legislators reinforce the importance of user notification and consent. But in practice, what we see is a context of pervasive and noiseless data collection, which cannot be faced only with individual informed choice, but one which depends on collective action²⁹. Especially concerning Internet usage, conditions to autonomy cannot be centered on the subject, since they involve a network of technical, human, administrative, political, juridical and several other actors.

As we see it, the nature of this technology and the way it affects individuals should be a matter of public concern. It is eminently social, since it is the collective dimension of

“ Privacy policies and corporations claim that individuals can block cookies or opt-out while social actors and even legislators reinforce the importance of user notification and consent ”

databases which permits, through statistical processing, the decision making process which is going to influence individual conduct. Furthermore, online tracking can be related to the emergence of potential privacy problems with important social implications: threats to freedom of choice and discrimination; data usage to purposes unknown by the user; discrimination practices, with the denial of products and services to a specific group or individual and the leverage of price according to different profiles; boxing, which is defined as the limitation of the consumers' vision and choices by his/her digital history ([24], p. 673)³⁰.

In this sense, some questions can be raised: Is enhancing notice and respecting users consent all we need? Do we need to have the power of choice or specialists and regulators to answer these questions for us? Yes or no, opting in or out, does it define enough space for negotiating a value as important as privacy? We believe that the freedom of choice, control and notification play an important role in rescuing privacy, but we also need an adequate regulatory framework, able to reflect society's stand on the question.

8. Conclusion

HTTP cookies and web beacons are popular tracking mechanisms in Brazilian websites and the majority of them are set by companies in the online marketing field. Third party cookies are also widely used in the websites analyzed and tracking processes are obscure. Opt-out options are not always available and information about their reach is sometimes unclear and limited. This context is built in a legal and technological shell that is too complicated for common people to understand. Privacy policies are generic and apparently designed to place doubt and responsibility on the shoulders of the public, opening enormous possibilities for companies to collect, manage and use personal data. In addition, an adequate regulatory framework is still under development in the country.

Opt-out options and user choice are hard to exert given the lack of transparency of such a context. Moreover, they cannot be taken as an easy way to get rid of the obligation of giving an adequate political response to the privacy problems that arise by behavioral targeting practices. If current practices shrink the space for negotiation, it thus requires us to rescue the social value of privacy. Hence, it needs to be regulated and discussed in the field of collective action.

Acknowledgments

This investigation is part of a research project titled "The social effects of regulation and data processing in Latin America" (Forthcoming) that approaches three main fields: Internet, identification schemes and video surveillance. It has been carried out in two countries (Brazil and Mexico) and was funded by the International Development Research Centre – IDRC <<http://www.idrc.ca/EN/Pages/default.aspx>>.

References

[1] O. Tene, J. Polonetsky. *To Track or 'Do Not Track': Advancing Transparency and Individual Control in Online Behavioral Advertising*. August 31, 2011. <<http://ssrn.com/abstract=1920505>>.

[2] D.J. Solove. *Understanding Privacy*. Cambridge: Harvard University Press, 2008. ISBN: 0674027728.

[3] F. Stalder. Privacy is not the Antidote to Surveillance. *Surveillance and Society*, 1 (1): 120-124, 2002.

[4] B. Latour. *Reassembling the Social*. Oxford: Oxford University Press, 2005.

[5] T.Z. Zarsky. Mine your own business! *Yale Journal of Law and Technology*, 5 (1), 2002.

[6] A. Soltani et al. *Flash cookies and privacy*. August 10, 2009. <<http://ssrn.com/abstract=1446862>>. Accessed on May 12, 2011.

[7] K. McKinley. *Cleaning Up After Cookies*. Technical report, iSEC PARTNERS, 2008, <https://www.isecpartners.com/files/iSEC_Cleaning_Up_After_Cookies.pdf>. Accessed on May 1, 2012.

[8] C. Jackson, A. Bortz, D. Boneh, J. Mitchell. Protecting Browser State from Web Privacy Attacks. *Proceedings of the 15th International Conference on World Wide Web*, May 23-26, 2006, Edinburgh, Scotland.

[9] J. Angwin. Latest in Web Tracking: Stealthy 'Supercookies'. *The Wall Street Journal*, August 19, 2011. <<http://online.wsj.com/article/SB10001424053111903480904576508382675931492.html>>. Accessed on April 29, 2012.

[10] B. Krishnamurthy, C. Wills. On the Leakage of Personally Identifiable Information Via Online Social Networks. *WOSN'09*, August 17, 2009, Barcelona, Spain.

[11] Tanzina Vega. Code That Tracks Users' Browsing Prompts Lawsuits. *The New York Times*, September 20, 2010. <http://www.nytimes.com/2010/09/21/technology/21cookie.html?_r=1&pagewanted=all>. Accessed on May 1, 2012.

[12] A. Banks. *State of the Internet in Brazil*. February 2011. <http://www.comscore.com/Press_Events/Presentations_Whitepapers/2011/State_of_the_Internet_in_Brazil>. Accessed on January 24, 2012.

[13] CGI.br. *Pesquisa TIC Domicílios 2010*. <<http://www.cetic.br/usuarios/tic/2010/index.htm>>. Accessed on February 11, 2012.

[14] J. Valentino-Devries. What they know about you. *The Wall Street Journal*, July 31, 2010. <<http://online.wsj.com/article/SB10001424052748703999304575399041849931612.html>>.

Accessed on January 20, 2012.

[15] P.M. Schwartz, D.J. Solove. The PII Problem. *NYU Law Review*, 86 (6): pp. 1814-94, 2011.

[16] D. Lyon. *Surveillance as Social Sorting*. London and New York: Routledge, 2003. ISBN: 0415278724

[17] C. Bennett. In defence of privacy. *Surveillance & Society* 8(4): pp. 485-496, 2011.

[18] A. Miller. *Assault on Privacy*. Ann Arbor: University of Michigan Press, 1971.

[19] A. Westin. *Privacy and Freedom*. Nueva York: Atheneum, 1970. ISBN: 0370013255.

[20] V. Steeves. Reclaiming the Social Value of Privacy. *Lessons from the Identity Trail*, ed. I Kerr. Oxford: Oxford University Press, 2008.

[21] K. Purcell, J. Brenner, L. Rainie. Targeted advertising: 59% of internet users have noticed it, but most don't like it. *Pew Internet & American Life Project*, March 9, 2012.

[22] G. McMillan. Less Than 1% of Firefox Users Use 'Do Not Track' Function. *Time Techland*, April 25, 2011. <<http://techland.time.com/2011/04/25/less-than-1-of-firefox-users-use-do-not-track-function/>>. Accessed on April 29, 2012.

[23] A. Acquisti, L. John, G. Loewenstein. What is privacy worth? *Twenty First Workshop on Information Systems and Economics (WISE)*, Phoenix, Arizona, 2011.

[24] M. Abrams. Boxing and concepts of harm. *Privacy and Data Security Law Journal*, September 2009, pp. 673-676.

[25] L. Scism, M. Marent. Test Data Profiles to Identify Risky Clients. *The Wall Street Journal*, November 18, 2010. <http://online.wsj.com/article/SB10001424052748704648604575620750998072986.html?mod=wsj_share_twitter>. Accessed on May 1, 2012..

[26] K. Hill. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. *Forbes*, February 12, 2012. <<http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/>>. Accessed on April 30, 2012.

[27] S. Stevenson. What your Klout score really means. *Epicenter*, April 24, 2012. <http://www.wired.com/epicenter/2012/04/ff_klout/all/1>. Accessed on April 30, 2012.

[28] Idec, Instituto Brasileiro de Defesa do Consumidor. Proteção de dados: Sorria, você está sendo monitorado. *Revista do Idec*, n. 150, December 2010.

Notes

¹ Cookies were created to cache applications state on users' computers and enhance their navigation performance, benefiting the user. On the one hand, they are useful as they make it possible for features such as saving passwords, retaining preferences (as volume or language) and files to be cached. But they can also be used to collect user's personal and Internet connection information (as an IP address and operational system details) and searches that users might have done in search engines. Since they are stored in the machines, giving it an identification number, they also enable

the tracking of pages visited thereon. HTTP cookies can be easily blocked through web browser's settings, but this can disturb the normal functioning of the website applications. Flash cookies are more resilient. They are not controlled by the browser. Thus, whatever you do in your browser (such as choosing private navigation options, erasing cache and so on) will not affect their functioning. In addition, Flash cookies are not stored in the same location as HTTP cookies, what makes it harder for the user to identify them [6]. Finally, web beacons are even harder to block because they are not files stored in web browsers. One example of this mechanism is the use of 1px transparent images, placed on a sequence of web pages, whose successive requests can be used to track user navigation.

² It can tell the kind of articles read, if you are in a news website, or can be collated with the product description bought, for example, enabling different profiles to emerge.

³ It includes cookies as tracking methods that use browser information, client browser state or content cached in a web browser [7][8]. An example is "history stealing" tracking, in which a website checks if a user has visited other specific [9]. Other important source of information includes social network sites, that can be used to identify a user through its profile page on a certain social network [10].

⁴ It enables agencies to manage the advertisements, their distribution and performance.

⁵ When someone visits a website, the ad server delivers a cookie attached to the banner one sees. This cookie is stored on the user's computer, and when the person visits another website showing ads delivered by the same server, his/her browser sends the cookie back to the server. Thus, the person is identified and based on information that has been previously collected about his/her, the system can "decide" which ad to show.

⁶ From the age of 15 and older, accessing Internet from home and work computers.

⁷ Orkut <<http://www.orkut.com>> has led social network market in the country until December 2011, when Facebook <<http://www.facebook.com>> registered 36.1 million visitors and finally took over after a year of unprecedented growth.

⁸ According to the ranking made by Alexa <<http://www.alexa.com/topsites/countries/BR>>. This ranking is ordered considering a combination of average daily visitors and pageviews over the past month.

⁹ <<http://www.terra.com.br/porta/>>.

¹⁰ <<http://www.uol.com.br/>>.

¹¹ <<http://www.globo.com/>>.

¹² <<http://br.yahoo.com/>>.

¹³ <<http://www.youtube.com/>>.

¹⁴ We have used Mozilla Firefox 4.0 and Windows XP Service Pack 2.

¹⁵ <http://www.macromedia.com/support/documentation/en/flashplayer/help/settings_manager07.html>.

¹⁶ <<http://www.ghostery.com/>>.

¹⁷ For this analysis, we have chosen the most popular apps of five different categories. This measurement of popularity considers the application usage by social network community as a whole, not only used by Brazilians. The Orkut applications chosen were: Segredos do Mar, Buddy Poke, Colheita Feliz, Baby Adopter and Musica. On Facebook, we analyzed the following applications: Causes, Texas Hold'em, Phrases, Badoo and Quiz Planet.

¹⁸ <<http://www.robtext.com>>.

¹⁹ Cookie #1: Placed by a social network website platform developer or a non advertising partner of these developers; Cookie #2: Placed by a research company which produces general reports of internet use; Cookie #3: Social media websites cookie placed in third party websites; Cookie #4: Placed by a company which does not quote its name in the URL associated with the cookie; Cookie #5: Placed by a company which offers market targeting solutions or audience and interests measurement solutions for social apps developers and advertisement publishers; Cookie #6: Placed by a company which offers traffic and access measurement solutions, semantic content analysis solutions or mapping tools to understand user behavior, optimizing websites and apps; Cookie #7: Placed by the website owner or a non advertising partner of the website; Cookie #8: Cookie placed by a third party website which offers its service embedded in a specific section of the analyzed website.

²⁰ For social network websites, Flash cookies have not been examined.

²¹ See note 19 for a description of the types of cookies. Beacon #1: Placed by a research company which produces general reports of internet use; Beacon #2: Social media websites beacon placed in third party websites; Beacon #3: Placed by a company which does not quote its name in the URL associated with the beacon; Beacon #4: Placed by a company which offers market targeting solutions or audience and interests measurement solutions for social apps developers and advertisement publishers; Beacon #5: Placed by a company which offers traffic and access measurement solutions, semantic content analysis solutions or mapping tools to understand user behavior, optimizing websites and apps; Beacon #6: Placed by the website owner or a non-advertising partner of the website.

²² Network Advertising Initiative (NAI) <<http://www.networkadvertising.org>> is a coalition of online marketing companies involved with regulation and consumers education about online advertising. It also offers a centralized opt-out mechanism for some member companies.

²³ TRUSTe <<http://www.truste.com>> is a United States company which certificate websites according to its own privacy policies.

²⁴ Safe Harbor Privacy Principles is a process that the United States corporations use to indicate comply with European Union Data Protection Directive - EU Directive 95/46/EC.

²⁵ In 2009, the service, tested in the United Kingdom by the British Telecom, was considered illegal by the European Commission.

²⁶ The lawsuits 08012.010585/2010-29 and 08012.003107/2010-62 are available at <<http://www.cade.gov.br/Default.aspx>>.

²⁷ The public debate is available at <<http://culturadigital.br/dadospeoais>>. The period for public hearing ended in March 2010. By now, the bill have not been sent to the National Congress.

²⁸ Miller [18, p. 25] says "*privacy is difficult to define because it is exasperatingly vague and evanescent*". Stalder [3, p. 3] claims that the "*bubble theory of privacy – based on concepts of individualism and separation – (...) applies a 19th century conceptual framework to a 21st century problem*". Lyon ([16] considers the concept is insufficient to face discrimination conditions prompted by social sorting: "*surveillance is not merely a matter of personal privacy but of social justice*".

²⁹ This approach is partly based in recent studies which show that user's practices and discourse towards privacy are ambiguous. On the one hand, users claim to be uncomfortable with behavioral advertising [21], but on the other hand, they usually do not take any attitude, no matter how effortless, to prevent their data from being collected [22]. Added to this, they seem to be likely to give away personal data even for small rewards [23]. It claims for a discussion that considers the value of online tracking not only to individuals, but also to broader society.

³⁰ Even though an adequate comprehension of these implications (or privacy problems) would require a deeper investigation – what would surpass the objectives of this preliminary study – harmful consequences of profiling practices are shown in the literature and newspaper articles. For instance, British insurer Aviva has been using market data to estimates people's risk for illnesses related to their lifestyles, raising concerns about denial of applicants and the leverage of price according to the consumer profile [25]. Target, a US-based retail chain, has been using purchase information to predict pregnancy, what ended up revealing to a father the pregnancy of his teenage daughter [26]. Sam Fiorella has been overlooked for a job despite her 15 years of experience because of her Klout.com score, a service that measures user's online influence without they even know about its existence, based on public information of social media accounts [27].