Use of Taxonomy of Privacy to Identify Activities Found in Social Networks' Terms of Use

Fernando de Assis Rodrigues* and Ricardo César Gonçalves Sant'Ana**

* São Paulo State University—UNESP, Univ. Estadual Paulista, Department of Information Science, Av. Hygino Muzzi Filho, 737, Marília—São Paulo, Brazil, 17525-900, <fernando@elleth.org>
** São Paulo State University—UNESP, Univ. Estadual Paulista, Department of Information Science, Av. Hygino Muzzi Filho, 737, Marília—São Paulo, Brazil, 17525-900, <ricardosantana@marilia.unesp.br>

Fernando de Assis Rodrigues is a lecturer for the Department of Information Science at São Paulo State University (UNESP) in Marília, Brazil. He is a member of the New Information Technologies (GPNTI/UNESP) and the Information, Data and Technology (IDT/USP) research groups and head-editor of the *Electronic Journal of Digital Competencies for Family Farming*.

Ricardo César Gonçalves Sant'Ana is a professor at the School of Science and Engineering and Coordinator of the Research Line Information and Technology in Information Science Program at São Paulo State University (UNESP), Brazil. He is a member of the New Information Technologies (UNESP) and the Information, Data and Technology (USP) research groups and coordinator of the Digital Competencies for Family Farming (CoDAF) outreach program. He also is an ad-hoc reviewer for several journals and funding agencies.

Rodrigues, Fernando de Assis and Ricardo Cesar Goncalves Sant'Ana. 2016. "Use of Taxonomy of Privacy to Identify Activities Found in Social Networks' Terms of Use." *Knowledge Organization* 43: 285-295. 22 references.

Abstract: The objective of this paper is to describe, on behalf of social network, elements which allow for identification of possible activity that can present potentially harmful effects to users' privacy, executed by either internal or external agents. To achieve this, the Taxonomy of Privacy was used to establish a way to categorize these possible actions found in terms of use, focusing on current guidelines set about issues related with gathering and storing personal data to increase users' perception about privacy issues. The universe of research was delimited to the study from the three prominent social networks at the time. The results are divided into two parts: the first part tries to identify evidence of activity with potential to be harmful to privacy through a linkage of characteristics from excerpts of the terms of use with concepts found of taxonomy; and the second shows comparisons between classifications made possible by taxonomy and their level of occurrence in terms of use studied. It was concluded that applying an appropriate taxonomy can help with the study of terms of use, enabling a perception of potential harmful activities under those terms. Also, it allows new proposals of applications of this methodology in other contexts.

Received 1 April 2016; Accepted 2 April 2016

Keywords: data, use, information, activities, terms of use, personal, users, privacy

1.0 Introduction

The penetration of information and communication technologies (ICT) in human actions and activities—centralized on a digital networked society, in which the main asset is the information (Castells 2010b)—enabled the establishment of a system (Castells 2010a) that would support the formation of networks with thousands of connected users for the purpose of exchanging information about numerous topics and also representing a new place of social and cultural organization. The Internet provides an infrastructure that allows a flow of a growing amount of types of data sets and documents. From this infrastructure, associated with the Hypertext Markup Language (HTML) created in 1989, emerged platforms (Adamic and Adar 2003; CERN 2015), which provide support to information networks and inter-relationship between people, called social networks or online social networks. Social networks have been present since the beginning of the Internet (Mislove et al. 2007, 30), and

with the maturity of the technology involved in this scenario, these social networks offer specific services for the interrelationship of members (Adamic and Adar 2003; Mislove et al. 2007), providing an exchange of information in multimedia formats such as images, videos, audios, hyperlinks and texts.

In December 2014 (Facebook Inc. 2015b; Jonathan Blake 2014; WORLDOMETERS 2015), Facebook had an average of 1.19 billion monthly active users, accounted for approximately 17% of the world population. Other social networks reached a monthly average of over 250 million active users, such as Instagram and Twitter. Therefore, these social networks reached significant numbers when compared to the human world population (about 7 billion). Mislove (2007, 29–30) considers these social networks as an integral part of most accessed and used services available over the Internet. These social networks, which are designed, developed and maintained by private companies, bring up existing concerns in other contexts such as the use of sensitive user data by companies, data exposure to governments and even to other users, and cyberstalking and digital spaces liable to provide hospitality for intolerant actions. In all these scenarios, there are data (Fogel and Nehmad 2009; Krasnova et al. 2009; Young and Quan-Haase 2009; Chen and Zhao 2012) that expose the underlying issues related to privacy. Problems related to user privacy in social networks are not caused only by the use of ICT (which can act as a catalyst agent in the automated collection of large amounts of data about users), but also by user activity, external agents and by data controllers that have sufficient expertise to gather and process user data together with other data sources, establishing new data (Solove 2006; Fogel and Nehmad 2009) showing potential harm to user privacy.

In this paper, two dimensions of privacy to be analyzed are highlighted (Sant'Ana 2013): the results of interactions between social networks and users during the data gathering phase and the definition of what is done with user data after storage in a database. In both dimensions, use of ICT (Vasalou et al. 2011) could affect privacy. However, to generalize activities which are harmful to user privacy, executed by social network information holders, is a complex issue primarily because users are informed and must agree to the terms of use when registering a new account to access the service. The terms of use of these networks have two roles in this process: 1) peacemaker, as an element of perception of security to users by establishing legal limits and guarantees on what is done with personal data; and 2) transparency, as an element between the user and the service about what will be done with users' data, diluted in high network complexity and in volume and variety of actions and activities likely to be performed (Castells 2010b; Castells 2010a). In this way, digital information environments do not offer a minimum set of perception elements to monitor potential harmful actions and their effects or impacts on user privacy. Based on these privacy issues, Solove (2006) proposes a taxonomy focused on the difficulty of segregating and categorizing existing types of activities potentially harmful to privacy, called "Taxonomy of Privacy." It is divided into groups and each one has a definition of a specific activity, describes the *modus operandi*, characteristics of infringement and possible damages.

1.1 Objective and procedures

The objective of this paper is to describe, on behalf of social network terms, elements which allow for the identification of possible activity that can present potentially harmful effects to users' privacy, executed by either internal or external agents. To achieve this, the "Taxonomy of Privacy" proposed by Solove (2006) was used to establish a way to categorize these possible actions found in terms of use, focusing on the current guidelines that these terms set about issues related with gathering and storing personal data to increase users' perceptions about privacy issues. The universe of research was limited to the study of document collections that compose terms of use from the following social networks: Facebook, Instagram and Twitter. This choice took into consideration the number of active users of these networks.

The scope of this paper does not include analyzing whether these terms of use guarantee or not the privacy of personal data for a particular activity, nor to identify the source (external or internal to social network services) of a potential agent able to execute these activities. The scope is restricted to identifying and categorizing, through the application of an appropriate taxonomy, potentially harmful activities to the privacy of its users. The methodology is outlined as a systematic documentary analysis from the reading of the document collections that compose each of the terms of use, divided into three stages. The first stage consists of detailing all groups with activity harmful to privacy, categorized as proposed by Solove's taxonomy. The second step identifies characteristics of the terms of use texts about members' personal data through the selection of specific excerpts from the available documents. The third stage presents the results of correlation between "Taxonomy of Privacy" subgroups and characteristics identified in terms of use, divided into two parts. The first part identifies potential activities that could be harmful to users' privacy through linking characteristics found in terms of use with taxonomy of privacy concepts, and the second part presents comparisons between classifications possible by the application of taxonomy of privacy and the level of occurrence in the terms of use studied.

2.0 Taxonomy of Privacy structure

The taxonomy of privacy is composed of four groups: "information collection," "information processing," "information dissemination" and "invasion," which in this study were identified as groups I, II, III and IV, respectively. The four groups are divided into subgroups (Table 1). Each subgroup is a type of harmful activity for privacy and Table 1 presents a summary describing all subgroups found in the taxonomy structure (a total of sixteen).

2.1 Information collection (group 1)

The first group is called "information collection" and involves privacy violations of activity at the time of data gathering of an individual or a collective group of individuals. This group is divided into two subgroups: "surveillance" and "interrogation." The surveillance subgroup concentrates on activity with the purpose of monitoring an individual or an entity in private or public space. For example, a service available on the Internet can process data gathered at different times about a user, based on justification that the data can be used to improve the user's experience on the platform. Another example can be a service available on the Internet which can process data gathered at different times about a user, based on justification of the data it can be used to improve a user's experience on the platform and perform surveillance actions such as content-targeting based on data collected about user paths (including geographical coordinates, humidity, atmospheric pres-

Group	Subgroup	Activities
Information Collection *(Group I)*	Surveillance	Activities with the purpose of monitoring an individual or an entity in their private or public space.
	Interrogation	Activities with data gathering processes, based on interrogation or interview.
Information Processing *(Group II)*	Aggregation	Activities related to bind user data with other data sources in order to reveal hidden facts when they were analyzed separately.
	Identification	Activities that are results of a user data binding process allowing agents to identify or re-identify user data to their respective individuals or entities.
	Insecurity	Activities that do not show reliability to those involved in issues of personal data access.
	Secondary Use	Activities that involve a use of data gathered for a specific purpose and subsequently used for other purposes.
	Exclusion	Activities that show transparency during the individual personal data storage process, the sharing of their data to third parties or the lack or an inability to participate in decisions involving gathering, storage, use and sharing of their own personal data.
Information Dissemination *(Group III)*	Breach of Confidentiality	Activities that occur as a breach of trust between parties to maintain the confidentiality of information about individuals.
	Disclosure	Activities that disseminate information about an individual that cause changes in the way that other people judge his/her character.
	Exposure	Activities linked to emotional and physical exposure and attributes of individual intimacy to a third party, such as nudity, bodily functions and private information.
	Increased Accessibility	Activities aimed at amplifying the access to personal data beyond expected or previously agreed between parties.
	Blackmail	Activities of control, domination, intimidation or threats to individuals or to groups by third parties.
	Appropriation	Activities that use personal data of an individual for benefit of a third party or to validate a service or a product without the full consent or understanding from the individual.
	Distortion	Activities that disseminate information that may be false, out of context or have the possibility of misinterpretations about an individual.
Invasion *(Group IV)*	Intrusion	Activities with the purpose of raiding private information or individual issues.
	Decisional Interference	Activities where there is state involvement in private matters that, somehow, try to perform or to change decisions on behalf of the individual.

Table 1. Groups, subgroups and activities, adapted from Solove (2006).

sure, altitude and azimuth), device and network data, recorded voice commands, user tastes and experiences about visited places, time spent in a public or private place, information about network connection, and metadata of images, audios and videos. The interrogation subgroup concentrates all activities with data gathering processes based on interrogatory or interview. Some examples of this subgroup are the services that appear in the registration process, forms with required fields requesting information that might be sensitive for a certain audience—and if the user does not have interest in sharing such information, they will not have access to the service.

2.2 Information processing (group 2)

The second group is called "information processing." This group involves activities harmful to privacy from the storage process (persistence), and the handling and use of data about individuals. This group is divided into five subgroups: "aggregation," "identification," "insecurity," "secondary use" and "exclusion." The aggregation subgroup is linked to activities related to the process of combining data from multiple sources about individuals, in order to reveal facts hidden when they were analyzed separately. For example, a user of a social network can provide data on relationship status when filling out profile information and also by searching for pages about places to go on a honeymoon. An external agent, which has access to the gathering of these data, can infer whether this individual is prone or not to buy future products and services for honeymooners through specialized algorithms for data aggregation. By combining data from these sources, an external agent performs an activity found from the aggregation subgroup. The identification subgroup consists of activities that come from the results of the user data binding process allowing agents to identify or re-identify user data to their respective individuals or entities.

Services on the Internet that provide access to data about their users may be subject to collection of these data sets by external agents, and if they have the necessary expertise and skills to recombine these collected data sets with data from other sources, it may increase the information repertory about a particular user and (re)-identifying that user in various domains (even linking these data with a personal ID such as social security or credit card number). This type of activity is part of the identification subgroup. The insecurity subgroup consists of activities that do not show reliability to those involved with issues of personal data access. For example, when a social network is the target of unauthorized external data gathering through techniques that exploit vulnerabilities, the result is a leak of personal data that has no possibility of returning to its original state (when there was no leak). The data access policy has been compromised and there is no guarantee that there are no other copies of that data within third parties.

In the secondary use subgroup, activities are contained that involve the use of data gathered for a specific purpose and subsequently used for other purposes. When personal data are gathered about individuals with a purpose, for example, to create a dynamic photograph album based on location, and shared to third parties for customization of advertisements. The exclusion subgroup consists of the activities that show some opacity to individuals in processes of personal data storage and data sharing to third parties, creating a lack or an inability to fully participate in decisions involving gathering, storage, use and sharing of their own personal data. For example, in a social network of which an individual is unaware, the user has no access to or does not participate in decisions about how personal data use is an activity of the exclusion subgroup.

2.3 Information dissemination (group 3)

The third group is called "information dissemination" and involves publishing activities, exposure and dissemination of information about individuals. It is divided into seven subgroups: "breach of confidentiality," "disclosure, exposure," "increase accessibility," "blackmail," "appropriation" and "distortion." The breach of confidentiality subgroup consists of activities that occur when there is a breach of trust between parties to maintain the confidentiality of information about individuals. A service that establishes non-sharing of personal data to third parties in their terms of use, and subsequently, the data are available to a preselected external public or are directly or indirectly publicly accessible, develops an activity bound to the breakage of secrecy subgroup. The disclosure subgroup consists of activities that disseminate information about an individual and causes changes in the way that other people judge that individual's character. For example, when it is not transparent to users which information repertory will be available to their peers and to connections of their peers, this type of activity overshadows the real reached audience from the delimited audience, which may result in a judgment about the individual's character on issues of a private nature by disseminating personal data.

The exposure subgroup consists of activities linked to exposure of emotional and physical attributes of individual intimacy to a third party, such as nudity, bodily functions and private information. Multimedia content sharing websites can store and preserve data containing personal photos and videos without the consent or full understanding of those involved, revealing intimacies to third parties. Even when that content could be removed, the intimacy of those involved has already been revealed, part of an ac-

tivity from the exposure subgroup. The increased accessibility subgroup consists of activities aimed at amplifying the access of personal data beyond expectations or previously agreed upon between parties. For example, when a website shares user data with other services, whether owned by them or by a third party, it extends access to these data beyond previous consent, even when this process is explained in the terms of use. Therefore, since user data will be linked to the terms of use of these services, and these services may have this in their terms of use, there will be different limitations about how personal data will be (re)shared with their own partners.

The blackmail subgroup categorizes activities of control, domination, intimidation or threats to individuals or to groups by third parties. An example of activity in this subgroup is the occurrence of blackmail, intimidation and threats of groups or individuals through the use of extortion to raise funds from disclosure of personal data (such as intimate photographs). In the appropriation subgroup are concentrated activities that use an individual's personal data for the benefit of a third party or to validate a service or a product without the full consent or understanding of the individual. For example, a service or a social network that uses personal data or photographs of its members as a way to confirm a kind of social approval of a brand or a product performs an activity bound to the appropriation subgroup. The distortion subgroup consists of activities that disseminate information that might be false, out of context or able to possibly misinterpret an individual. Websites that disseminate third party personal data as an online public catalog—extracted or linked from social network users data (using techniques like data crawling or data scraping)—offer public decontextualized information that may open a misleading interpretation about an individual, an activity bound to the distortion subgroup.

2.4 Invasion (group 4)

The fourth group is called "invasion" and is comprised of invasion activities against the privacy of individuals. The group is divided into two subgroups: "intrusion" and "decisional interference." The intrusion subgroup consists of activities with the purpose of raiding private information or data regarding individual issues. Examples are the mandatory installation of tools or a required use of a service with a purpose of recording user actions in a particular digital environment without consent or the party's full understanding. The decisional interference subgroup consists of state involvement activities in private matters that, somehow, try to perform or change decisions on behalf of the individual. When a state interferes with private nature actions, such as running investigations involving data-sharing of examination of a human body part, it might interfere with things that should be free of interference from others.

3.0 Characteristics of terms of use

The social networks Facebook, Instagram and Twitter have in the footnotes of their respective websites a specific area for access to the terms of use. They are offered through a hyperlink with a label titled "terms," "terms of use" or "terms of service" or its equivalent in other languages. Their characteristics are described individually in the following sections.

3.1 Facebook

The "Terms of Use" consists of a main document called the "Statement of Rights and Responsibilities" and eleven additional documents about specific services or issues with which the social network has direct or indirect involvement: data policy with information about the policy on personal data; payment terms with additional terms of payments made through the social network; platform page with information about data exchange with external applications; Facebook "Platform Policies" with the guidelines for developing external applications; advertising guidelines, guidelines about issues involved with advertising through social network; self-serve ad terms with guidelines about advertising for applications and services connected with Facebook; promotions guidelines with guidelines for competitions and awards; Facebook brand resources with guidelines for how partners may use Facebook intellectual property; how to report claims of intellectual property infringement with copyright guidelines; pages terms with usage guidelines for managing pages, and community standards, containing information about good conduct, fairness and standards on user actions and on his/her coexistence among its members. All of these documents (Facebook Inc. 2015c) are available even when users are not logged on to the social network.

According to documents available, when users post their pictures and videos on the platform, it guarantees the rights shared to Facebook for use and transfer of their personal data to partners in any place and at any time, stopping only when a user chooses to delete an account. In "data policy" (Facebook Inc. 2015a), it is explained that the platform can share user data with other partner companies, such as external sales teams, advertising companies, and regional offices, among others. The document collection does not have detailed information about the technical procedures executed to guarantee effective exclusion of personal data at the request of a user. However, Facebook Inc. (2015c) explains that personal data that are related to licensed media content with copyright or already shared

with third parties may be continually available even when a data exclusion is requested by a user.

In the fourth section entitled "Registration and Account Security," on the statement of rights and responsibilities, establishment of certain behaviors allowed to users, highlighting that social network users must agree to keep all of their personal information "accurate and up-to-date" (Facebook Inc. 2015c, 1). Otherwise, the company reserves the right to block user account access if it specifically considers that personal data or activity is not in accordance with predetermined guidelines. The seventeenth section called "Definitions," contains the definition of the meaning of certain terms, such as the term "platform," which is defined as a set of all of the services and interfaces for external application programming offered by Facebook. These platforms enable the sharing of personal data by Facebook with "others, including application developers and website operators, to retrieve data from Facebook or provide data to us."

In the "Data Policy" (Facebook Inc. 2015a) it is explained that metadata identified in multimedia content uploaded by users are stored, such as geospatial location, the date of creation of content, device and Internet Service Provider, IP address, the language, protocols and phone number. The data policy also specifies that users' information can be used to customize advertisements and measurement systems for the display of relevant ads or not and to measure effectiveness and reach of advertising.

3.2 Instagram

The "Terms of Use" (Instagram Inc. 2013c) is divided into five documents: terms of use, with general information; privacy policy with information related to privacy policy adopted by the social network; API Terms of Use with information on the rights and responsibilities of external applications that use data of its users collected through application programming interface (API); community guidelines, containing information about approved coexistence attitudes among users; and, intellectual property with information about copyright and licensing of products and services. All of these documents are available even when users are not logged on to the social network.

According to the documents available, when users post their pictures and videos on the platform, the right shared to Instagram for the use and transfer of their personal data to partners in any place and at any time is guaranteed, stopping only when an account is deleted. The "Terms of Use" (main document) (Instagram Inc. 2013c) explains that users have access to applications developed by third parties on their account with consent prior to sharing personal data, such as account name, user name and profile photograph. The document collection does not have detailed information about the technical procedures executed to guarantee an effective exclusion of personal data by user request. However, (Twitter Inc. 2014a) personal data that were already shared with partners may continue to be available even when a data exclusion is requested by a user.

The "Terms of Use," items 4, 5 and 6 of the general conditions section (Instagram Inc. 2013c) say that Instagram can block user account access, arbitrarily delete or modify contents or change a user name and monitor activities of its members without notice at any time. In the "API Terms of Use" (Instagram Inc. 2013a), it is explained that developers of external applications should inform users how Instagram's personal data are collected, stored, processed and disseminated. The privacy policy (Instagram Inc. 2013b) informs users of services that all metadata identified in multimedia content uploaded by users are stored, such as geospatial location, the date of creation of content, device and Internet Service Provider, IP address, the language, protocols and phone number.

3.3 Twitter

The "Terms of Use" consists of a main document called "Twitter Terms of Service" and five additional documents about specific services or issues with which the social network has direct or indirect involvement: "The Twitter Rules" with Twitter community principles and tool use guidelines; "Twitter Privacy Policy" with information about collecting, use and share of data; "Developer & Policy Agreement" with rights and duties of partners that intend to develop external applications and use Twitter users' data; "Twitter Commerce Terms" with terms of payments and promotions made through the social network; and, "Inactive Account Policy" with information about exclusion and decay processes of user accounts. All of these documents (Twitter Inc. 2014a) are available even when users are not logged on to the social network. The main document established in the privacy section says that all topics concerning the use of personal information by the service or by a third party is guided by the "Twitter Privacy Policy" document (Twitter Inc. 2014a; Twitter Inc. 2014b).

According to documents available, when users post their pictures and videos on platforms they guarantee the right to be shared on Twitter for use and transfer of their personal data to partners, in any place and at any time, stopping only when a user chooses to delete the account. The service (Twitter Inc. 2014b) also reserves the right to share users' personal data, such as geospatial location, the date of creation of content, texts, images, hyperlinks and languages codes with partner companies and advertising agencies, through its API. By default, all posts are available publicly and content, associated with personal data about

users, may be used for customized advertisements (Twitter Inc. 2014a; Twitter Inc. 2014b).

4.0 Results

The first part of the results presents the characteristics of possible activities harmful to privacy starting from elements identified in terms of use of each social network, linked to subgroup concepts defined in Solove's taxonomy of privacy. The second part presents comparisons between classifications possible through an application of the taxonomy of privacy and their level of occurrence of subgroups in terms of use.

4.1 Linking between taxonomy of privacy and characteristics found in terms of use

In this section, each paragraph is composed of the characteristics identified in the terms of use and their link with potential activities found in each subgroup of the privacy taxonomy—presented at the end of each sentence. Each characteristic can be linked with one or more subgroup, and subgroups can be linked to one or more characteristics.

4.1.1 Facebook

On Facebook, users are not able to find a place or a main document that lists all third party companies that have access to personal data, for which is possible to link this kind of activity with the subgroup increase of accessibility. It is not possible to delineate, in a transparent way, which kind of information is created from personal data by third party companies or how they handle personal data sets with other data sources, toward possible activities bound to the subgroups secondary use and aggregation. Although the "Terms of Use" (main document) established in the "Sharing Your Content and Information" section that all personal data are users' property, there is some opacity about how third party companies manage data rights, acting as a possible catalyst for a breach of confidentiality of activity environment, as set out in the breach of confidentiality subgroup.

According to the "Data Policy," personal data could be shared with advertising companies. These advertising companies are able to bind personal data (such as user tastes or interests by pages, products or services) to third party products and services offered on social networks, without adequate transparency to users about how this process is carried out (for example, to determine what the target public reached by a specific advertising or duration time of this campaign was), concomitant with activities related to the appropriation subgroup. Despite the "Terms of Use" defining which activities are allowed to users, there is some opacity in arbitration procedures of alleged violations (like a suspension of a user account), directly linked to inherent questions to the subgroups exclusion, surveillance and insecurity.

The personal data deletion process is not transparent to users, especially considering all the content already shared with third party companies and data linked to copyright content that will not be deleted. This scenario creates some uncertainty for users about how to interact with available content on the network since there is no guarantee whether the published user content will be deleted completely by a user request thus linking this activity to the insecurity subgroup.

It is possible to execute data-gathering through the social network API, allowing occurrences of automated data-gathering actions by external agents without a prior consent of users, since Facebook is allowed to set up new agreements with other partner companies over time, triggering activities bound to the break of confidentiality subgroup. Facebook can share the photo, name and other variables not identified about their users with partner companies and there is no guarantee for users that these companies will not link user data with their own databases, allowing that there may be instances of activities related to the groups aggregation and appropriation.

It is not transparent to users whether the social network is or is not sharing the metadata of multimedia user content with partner companies. If these metadata have been shared, it is possible to link to their other sources of personal data, ensuring the third party additional sensitive information, such as geospatial and temporal location of content—an activity that is part of the aggregation subgroup. Advertising company partners of Facebook can process user data with other data sources in order to customize advertising campaigns for users. It is also allowed to increase access of this kind of information to third parties through the charging of an advertising fee. These activities are linked to the subgroups secondary use and aggregation. *A priori*, custom external advertisement is needed to access personal data to increase accuracy. For that, third party companies need to get unique identifiable data access to distinguish each user, allowing identification of a single user in a group. This activity is linked to the identification subgroup.

4.1.2 Instagram

On Instagram, users are not able to find a place or a main document that lists all third party companies that have access to personal data, making it possible to link this kind of activity with the subgroup increase of accessibility. It is not possible to delineate in a transparent way

which information is created from personal data by third party companies or how they handle personal data sets together with other data sources, toward possible activities bound to the subgroups secondary use and aggregation. There is low transparency in the personal data deletion process, because even with an appropriate place to request removal of all personal data from the social network by a user, part of the personal data is not effectively located in databases. The process removes the visibility of personal data to other users, but part of these data are still stored in company databases, linking this to the subgroup insecurity. Despite the "Terms of Use" defining which activities are allowed to users, there is some opacity in the arbitration procedures of alleged violations (for example, in an arbitrary exclusion of content generated by a user), directly linked to inherent questions for subgroups exclusion, surveillance and insecurity.

It is possible to execute data-gathering through the social network API, allowing occurrences of automated data-gathering actions by external agents without prior consent of users, since it is allowed for Instagram to set up new agreements with other partner companies over time, triggering activities bound to the break of confidentiality subgroup. It is not transparent to users whether the social network is or is not sharing the metadata of multimedia users' content with partner companies. If metadata have been shared, it is possible to link them to other sources of personal data, ensuring third parties additional sensitive information, such as geospatial and temporal location of content—an activity that is part of aggregation subgroup.

4.1.3 Twitter

On Twitter, personal data such as name, user name, profile picture and location are, by default, part of a public data catalog. There is no local or no appropriate list for users to retrieve when, what, how, where and by whom their sets of personal data are collected and what other data sources are aggregating to them, part of possible activities of the subgroups aggregation and appropriation. Twitter users are not able to find a place or a main document that lists all third party companies that have access to personal data, which is possible to link this kind of activity with the subgroup increase of accessibility.

It is not possible to delineate, in a transparent way, what information is created from personal data by third party companies, or how they handle personal data sets with other data sources, toward possible activity bound to the subgroups secondary use and aggregation. The Twitter settings are adjusted by default on a model where all content posted by users becomes public to any interested party without the need for identification to access and data-gathering, making it possible for third parties to perform activities that might harm privacy through an opacity of information disclosure process in which information is apparently private to users but could be public to everyone or even retransmitted incompletely. These activities are linked to the subgroups disclosure, secondary use and exposure.

4.2 Results from taxonomy of privacy application

Figure 1 shows a chart with a total amount of potentially harmful activity that was linked to each "Terms of Use," grouped by social networks. It is divided into three columns, each representing the number of subgroups identified for each social network. If the sub-group has been linked to a "Terms of Use" at least once it is added to 1, otherwise it is added to 0, which is not accumulative for recurrence.

The taxonomy of privacy has a total of 16 sub-groups, and social networks (columns) had a total of links with subgroups varying between 6 and 9 with an approximate average of 7.34 (horizontal line in evidence). All "Terms of Use" presented at least one link to subgroups aggregation, secondary use, increased accessibility, surveillance, insecurity, exclusion, breach of confidentiality, appropriation, identification, disclosure and exposure, totaling 22 links with subgroups.

Table 2 systematizes the taxonomy subgroups per occurrence ranges, grouped by total number of social networks whose "Terms of Use" are linked with a subgroup at least once. It is divided into three columns: the first column with subgroup names is followed by a column with the total number of social networks with whose "Terms of Use" are related with the subgroup—ranging from 0 (when subgroup not been linked to any "Terms of Use") to 3 (when all "Terms of Use" show at least one potential activity linked to the sub-group).

The third column (Range) establishes a division by occurrences of range of links found between taxonomy subgroup activities and "Terms of Use," such that: Range 1 concentrates on the subgroups with a higher incidence; Range 2 subgroups with an intermediary incidence; Range 3 subgroups with a low incidence; and, Range 4 with no incidence of the sample.

Table 3 exhibits a redistribution of possible activities of subgroups harmful to privacy sorted in their respective taxonomy groups with subgroup results grouped by occurrence ranges of links found between taxonomy subgroup activities and "Terms of Use," where rows represent "Taxonomy of Privacy" groups and columns divide occurrence ranges (from Table 2).

It is possible to establish that "Terms of Use" of social networks have: 1) higher incidence spots, that support pos-

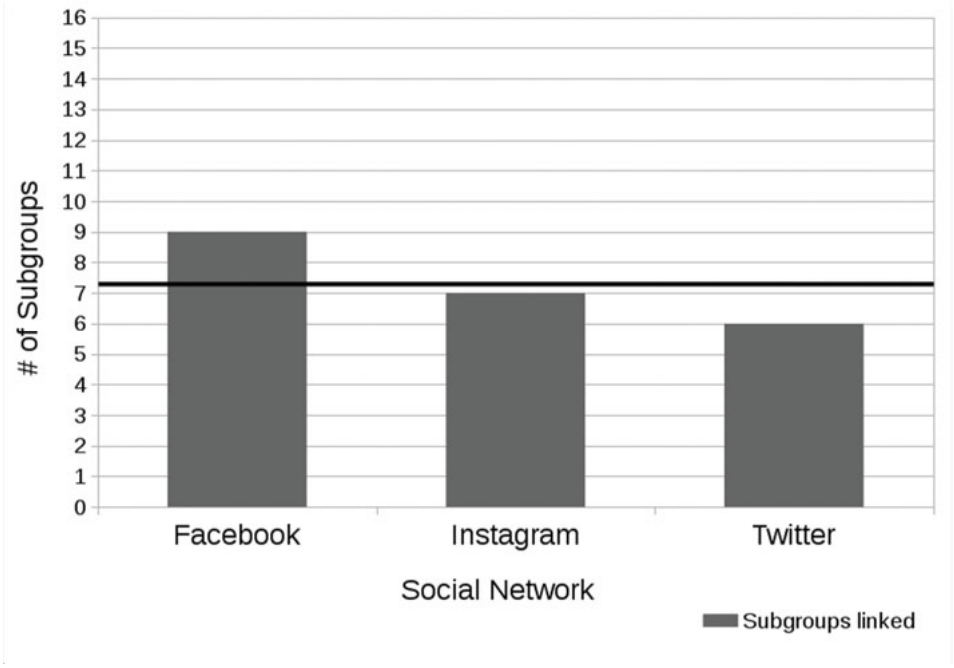

Figure 1. Chart of subgroups linked to Terms of Use, grouped by social networks.

Subgroup	**Total Social Networks linked to Subgroup**	**Range**
Aggregation	3	Range I
Secondary Use	3	
Increased Accessibility	3	
Surveillance	2	Range II
Insecurity	2	
Exclusion	2	
Breach of Confidentiality	2	
Appropriation	2	
Identification	1	Range III
Disclosure	1	
Exposure	1	
Interrogation	0	Range IV
Blackmail	0	
Distortion	0	
Intrusion	0	
Decisional Interference	0	
Total	22	-

Table 2. Subgroups, grouped by ranges and ordered by number of social networks linked.

sibilities of activity harmful to privacy linked to groups information processing and information dissemination, mostly bound to issues related with subgroups aggregation, secondary use and increased accessibility; 2) medium-range incidence spots that support possibilities of activity harmful to privacy linked to all groups, mostly bound to issues related with subgroups surveillance, insecurity, exclusion, breach of confidentiality and appropriation; and, 3) lower incidence spots that support possibilities of activity harmful to privacy linked to the groups information processing and information dissemination, mostly bound to issues related with subgroups identification, disclosure and

Group/Range	Range I	Range II	Range III	Range IV
Information Collection *(Group I)*		Surveillance		Interrogation
Information Processing *(Group II)*	Aggregation Secondary Use	Insecurity Exclusion	Identification	
Information Dissemination *(Group III)*	Increased Accessibility	Breach of Confidentiality Appropriation	Disclosure Exposure	Blackmail Distortion
Invasion *(Group IV)*				Intrusion Decisional Interference

Table 3. Subgroups of potential harmful privacy activities, divided by the hierarchy established by taxonomy of privacy, divided by range of occurrences in "Terms of Use" of social networks.

exposure. Interrogation (from information collection), blackmail and distortion (from information dissemination), intrusion and decisional interference (from invasion) compose subgroups without incidence and therefore were not linked to any use of the term.

5.0 Conclusions

The rounded average of 7 subgroups identified in each "Terms of Use" (Figure 1) from a total of 16 subgroups opens space for a reflection toward a necessity to go further on privacy issues explicit about collections of documents that compose the terms of use of commercial services like social networks, especially through studies—guided by these thematic axes of harmful activities—helping to explicate information, before hidden, such as which are the most recurrent activities or those more likely of achievement and what are the possible gaps in the "Terms of Use" texts that might ensure a legality of potential harmful activity to user privacy.

It was concluded that applying an appropriate taxonomy can help with the study of terms of use, enabling a perception of potential harmful activities under those terms. The application of Solove's "Taxonomy of Privacy" also allowed for a new classification of these subgroups when linked to terms of use characteristics from a sort by occurrence ranges (Table 2)—where this classification shows the most recurrent possible activities in terms of use (Range I) to the possibilities that have not been checked (Range IV). This tracking can help, for example, in the development of upgrade strategies to terms of use of these services through the presentation of potential gateways of harmful activities more adherent to social networks. Also, it allows new proposals of applications of this methodology in other digital contexts or in different scenarios and opens the possibility of applying these results to the definition of monitoring strategies for users' privacy in digital environments.

References

Adamic, Lada A. and Eytan Adar. 2003. "Friends and Neighbors on the Web." *Social Networks* 25:211–30. doi: 10.1016/S0378-8733(03)00009-1

Castells, Manuel. 2010a. *The Power of Identity,* 2nd ed. Vol. 2, The Information Age: Economy, Society, and Culture. Malden, MA: Wiley-Blackwell.

Castells, Manuel. 2010b. *The Rise of the Network Society*, 2nd ed. Vol. 1, The Information Age: Economy, Society, and Culture. Malden, MA: Wiley-Blackwell.

Chen, Deyan and Hong Zhao. 2012. "Data Security and Privacy Protection Issues in Cloud Computing." In *Software Engineering & Digital Media Technology: Proceedings of the 2012 International Conference on Computer Science and Electronic Engineering (ICCSEE 2012), March 23–25, Hangzhou, China.* [United States]: IEEE Computer Society, 647-51. doi:10.1109/ICCSEE.2012.193

Conseil Européen pour la Recherche Nucléaire (CERN). 2015. "The Birth of the Web." http://home.web.cern.ch/topics/birth-web

Facebook Inc. 2015a. "Data Policy." Facebook Inc. https://www.facebook.com/about/privacy/

Facebook Inc. 2015b. "Statistics." http://newsroom.fb.'com/company-info/

Facebook Inc. 2015c. "Terms of Service." Facebook Inc. https://www.facebook.com/legal/terms

Fogel, Joshua and Elham Nehmad. 2009. "Internet Social Network Communities: Risk Taking, Trust, and Privacy Concerns." *Computers in Human Behavior* 25:153–60.

Instagram Inc. 2013a. "API Terms of Use." Instagram Inc. https://instagram.com/about/legal/terms/api/

Instagram Inc. 2013b. "Privacy Policy." Instagram Inc. https://help.instagram.com/155833707900388

Instagram Inc. 2013c. "Terms of Use." Instagram Inc. https://help.instagram.com/478745558852511

Jonathan Blake. 2014. "Instagram Now Bigger than Twitter," December 10, British Broadcasting Corporation (BBC) edition, sec. Newsbeat. http://www.bbc.co.uk/

newsbeat/article/30410973/instagram-now-bigger-than-twitter

Krasnova, Hanna, Oliver Günther, Sarah Spiekermann and Ksenia Koroleva. 2009. "Privacy Concerns and Identity in Online Social Networks." *Identity in the Information Society* 2:39–63.

Mislove, Alan, Massimiliano Marcon, Krishna P. Gummadi, Peter Druschel and Bobby Bhattacharjee. 2007. "Measurement and Analysis of Online Social Networks." In *Proceeding of IMC '07 Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, October 23-26, 2007, San Diego, CA*, ed. Constantine Dovrolis and Matthew Roughan. NY: ACM Press, 29-42. doi: 10.1145/1298306.1298311

Sant'Ana, Ricardo César Gonçalves. 2013. "Ciclo de vida dos dados e o papel da ciência da informação." In *Anais do XIV Encontro Nacional de Pesquisa em Ciência da Informação 29 de outubro de 2013 – 01 de novembro de 2013, Universidade Federal de Santa Catarina*. Florianópolis, Brasil: ANCIB. http://enancib2013.ufsc.br/index.php/enancib2013/XIVenancib/paper/viewFile/284/319

Solove, Daniel J. 2006. "A Taxonomy of Privacy." *University of Pennsylvania Law Review* 154:477–564. doi:10.2307/40041279

Twitter Inc. 2014a. "Terms of Service." Twitter Inc. https://twitter.com/tos/previous/version_8?lang=en

Twitter Inc.. 2014b. "Twitter Privacy Policy." Twitter Inc. https://twitter.com/privacy/previous/version_9?lang=en

Vasalou, Asimina, Alastair J. Gill, Fadhila Mazanderani, Chrysanthi Papoutsi and Adam Joinson. 2011. "Privacy Dictionary: A New Resource for the Automated Content Analysis of Privacy." *Journal of the American Society for Information Science and Technology* 62:2095-2105. doi:10.1002/asi.21610

WORLDOMETERS. 2015. "World Population." Worldometers.info. http://www.worldometers.info/

Young, Alyson L. and Anabel Quan-Haase. 2009. "Information Revelation and Internet Privacy Concerns on Social Network Sites: A Case Study of Facebook." In *Proceedings of the Fourth International Conference on Communities and Technologies*, *June 25 - 27 2009, University Park, PA,* ed. John M. Carroll. NY: ACM, 265–74. doi:10.1145/1556460.1556499