Privacy and Social Network Applications

Antoni Roig

Researcher of the IDT (Law and Technology Institute) of the Autonomous University of Barcelona

Abstract. Privacy technological threatens are no limited to data protection. Social Network Applications (SNA) and ubiquitous computing or Ambient Intelligence face other privacy risks. The business model of SNA and the improvement of data mining allow social computation. SNA regulation should favor privacy-by-design and Privacy Enhancing Technologies (PET). Default friendly-privacy policies should also be adopted. The data portability of the applications shifts SNA into a new field of ubiquitous computing. Therefore, the solutions of the Ambient Intelligence should be also analyzed in the context of SNA.

Keywords: Social Network Applications, Privacy, Privacy Enhancing Technologies, Privacy-by-Design, Ubiquitous Computing, Ambient Intelligence.

1. Major Privacy Concerns

1.1. Control

Personally Identifiable Information (PII) is willingly provided by users of Social Networks Applications (SNA). So one of the major privacy concerns is the lack of self-control over these data. In fact, the business model of the SNA consists of exploiting the value of users’ PII. A recent study gives us a more precise idea of the amount of PII introduced in the SNA (Fogel, Nehmad, 2009). More than three-quarter of students have created a social networking profile in Facebook, and about one-half in MySpace. The average years for the profile displayed was 1.9 years. With regard to a daily visit to one’s profile, the average was 2.4 times. Other profiles were viewed on average 4 times. Concerning daily hours spent viewing profiles, the average was 1 h. The average number of “friends” on profiles was 239. Almost three-quarter allowed anyone to view their profile without restricting views to those specifically accepted. Almost 10% included their phone number and home address on their profile.

1. This study is within the frame of the funded research project “Freedom of speech In the Context Of Web 2.0 and Social Networks: Redefinition, Guarantees and Limits”, Cotino as main researcher. Ministry of Science and Innovation (DER2009-14519-C05-01).
The “News feed” and “Beacon” features in Facebook are interesting examples of control concerns. Facebook released the News Feed feature on September 5, 2006. The feature culls new PII that users post on their personal profile pages and delivers it to the website’s initial page (Hoadley, C. M. et al., 2009): for instance, “Alice’s status changed from ‘single’ to ‘in a relationship’.” Facebook indicated that it would make new information easier than ever to find. In response to the widespread concerns, Facebook immediately took down the News Feed applications and worked nonstop for two days on providing a wider variety of privacy preferences. Then Facebook re-released the News Feed applications with new privacy control features. On September 8, 2006, Facebook’s CEO, Mr. Zuckerberg, apologized for this privacy outcry and said: “this was a big mistake on our part, and I’m sorry for it. . . But apologizing isn’t enough. I wanted to make sure we did something about it, and quickly. So we have been coding nonstop for two days to get you better privacy controls.”

With News Feed, no new information was revealed; users could only see changes of their friends’ pages. So, why were users so uncomfortable with it? A plausible explanation is that the new interface offered lesser levels of perceived control over PII (Xu, H., 2009). One possible conclusion is that privacy concerns can be lessened by offering more control functions: first of all, control of PII disclosure; but also control access to disclosed information. With News Feed, obtaining information about other users was easier, which leads to a lower perception of control.

1.2. Transparency

Online privacy policies are difficult to understand. Most privacy policies require an ability to decode legalistic, confusing, or jargon-laden phrases. Privacy researchers and industry groups have thus devised several standardized privacy policy formats to help people compare policies (McDonald, A.M. et al., 2009).

Another helpful tool that shows what a user is sharing with whom is Facebook’s profile preview tool. Go to Settings’ Privacy Settings, then Profile, and type a friend’s name in the box on the top. You will see your profile as that friend would view it, and then you can adjust your privacy settings accordingly (Larkin, 2009).

1.3. Unauthorized use

Facebook states that it will do everything possible to protect the information posted on the site but “cannot and do not guarantee that User Content you post on the Site will not be viewed by unauthorized persons” (Facebook, 2009). Indeed, we will see later that other users or even non-users can accede and use PII most of the time from SNA users.

---

2. Legal Framework

2.1. Data protection regulation

Data protection regulations are considered by some authors as a reference basis for the development of methodologies tailored to design privacy-aware systems (Guarda, Zannone, 2009). The first step is to summarize the privacy principles:

(1) Fair and Lawful Processing: the collection and processing of personal data shall neither unreasonably intrude upon the data subjects’ privacy nor unreasonably interfere with their autonomy and integrity, and shall be compliant with the overall legal framework.

(2) Consent: personal data shall be collected and processed only if the data subject has given his explicit consent to their processing.

(3) Purpose Specification: personal data shall be collected for specified, lawful and legitimate purposes and shall not be processed in any way incompatible with the purposes for which data have been collected.

(4) Minimality: the collection and processing of personal data shall be limited to the minimum necessary for achieving the specific purpose. This includes that personal data shall be retained only for the time necessary to achieve the specific purpose.

(5) Minimal Disclosure: the disclosure of personal data to third parties shall be restricted and only occur upon certain conditions.

(6) Information Quality: personal data shall be accurate, relevant, and complete with respect to the purposes for which they are collected and processed.

(7) Data Subject Control: the data subject shall be able to check and influence the processing of his personal data.

(8) Sensitivity: the processing of personal data, data which are particularly sensitive for the data subject, shall be subject to more stringent protection measures than other personal data.

(9) Information Security: personal data shall be processed guaranteeing a level of security appropriate to the risks presented by the processing and the nature of the data (Guarda, Zannone, 2009).

2.2. Recommendations

Up to now, one report of 2008 is perhaps the most relevant legal framework on WBSN and privacy (“Rome Memorandum”). One key legal preliminary consideration is that there are only very few rules governing the publication of personal data at the initiative of private individuals. This is due to the fact that it had not been a major issue in the real world, and it has been only rel-
evant on the Internet with WBSN. Another sociological key element is that a new generation of “digital natives” has appeared, and they seem more comfortable with publishing (sometimes intimate) details of their lives on the Internet. The Rome Memorandum recommendations to regulators are:

– Introduce the option of a right to pseudonymous use.
– Ensure that service providers are honest and clear about what information is required for the basic service. Specific problems exist with consent of minors.
– Obligation to data breach notification for social network services.
– Possibly attributing more responsibility to WBSN providers for personal data content on WBSN.
– Improve integration of privacy issues and tools into the educational system.

Another interesting document is the European Network and Information Security Agency Position Paper 1. Some of the recommendations are:

– WBSN should, where possible, use contextual information to educate people in ‘real-time’.
– Awareness-raising campaigns should also be directed at software developers to encourage security conscious development practices and corporate policy.
– The regulatory framework governing WBSN should be reviewed and, where necessary, revised:
  • What is the legal position on deletion of user generated content by service providers if it is classed as WBSN spam?
  • What is the legal position on image-tagging by third parties?
  • Who is responsible for security flaws resulting from user-generated markup or scripting?
  • How should privacy policies of embedded third party widgets be communicated to users?
  • What exactly constitutes personal data in a WBSN environment?
  • What is the legal position on profile-squatting?
  • Should the posting of certain classes of data by minors (location data) be made illegal?
– Users should be given accurate information on what is done with their data before and after account closure. WBSN should be used in a controlled and open way (i.e. not banned or discouraged) with co-ordinated campaigns to educate students, teachers and parents.
Recently, we can mention the Working Paper n°163 of the article 29 Group dealing with online social communities, June 12, 2009\(^3\). This study considers that the European Directive of data protection covers also the SNA scenario:

“The new aspects of this recommendation are perhaps the reference to security tools and “privacy-friendly” default settings. For the first time, a recommendation mentions Privacy Enhancing Technologies as a solution, even if limited to the problem of the privacy of young users.”

2.3. Beyond data protection

But privacy cannot be limited to the data protection regulation. The German Constitutional Court published a decision in February 2008. This decision constitutes a new “basic right to the confidentiality and integrity of information-technological systems” as part of the general personality and privacy rights in the German constitution\(^4\). The ruling explains the relevance of using information-technological systems for the expression of personality. These systems are defined as technological tools that “alone or in their technical interconnectedness can contain PII of the affected person in a scope and multiplicity such that access to the system makes it possible to get insight into relevant parts of the conduct of life of a person or even gather a meaningful picture of the personality”. Weiss considers that this description could be easily applied to social network profile data (Weiss, 2009).

3. Privacy Preserving Tools and Procedures

3.1. SNA need new PETs

At least, the ISO has achieved consensus on four components of privacy as follows (Wright \textit{et al.}, 2009):

- **Anonymity** ensures that a subject may use a resource or service without disclosing user identity.
- **Pseudonymity** ensures that a user may use a resource or service without disclosing identity, but can still be accountable for that use.
- **Unlinkability** ensures that a user may make multiple uses of resources or services without others being able to link these uses together.

---


Unobservability ensures that a user may use a resource or service without others, especially third parties, being able to observe that the resource or service is being used\(^5\).

General Privacy Enhancing technologies, like anonymity, pseudonymity, and access management are useful for surfing the Internet. Nevertheless, SNA do not seem a priority because the increasing of the PII is at the core of business model. This is a major problem for PeerSoN, a decentralised SNA based on Peer-to-Peer (Buchegger et al.). More, privacy in SNA can not be considered as data protection, and thus be limited to access control and processing of data. Anonymity is not completely abandoned (Kacimi et al., 2009) and access control can be managed with semantic web framework (Carminati et al., 2009). We can also adapt interesting cryptographic solutions to Facebook (Lucas, Borisov, 2008). The main aim of cryptographic solutions is to what extent a user can ensure his privacy while benefiting from existing online services. NOYB is a novel approach that provides privacy while preserving some of the functionality provided by online services (Guha et al., 2008). The identity management can also be improved by reputation systems adapted to SNA, which are platforms of several online communities linked such as Facebook (Pingel and Steinbrecher, 2008).

However, SNA need new PETs focused on transparency, automatic compliance assurance functions and proactive communications techniques on risks (Weiss, S., 2009).

### 3.2. Transparency

The complete transparency and control of the usage of the user’s PII is only possible with privacy-by-design practices for designers and developers. PETs cannot be simply an added tool, but have to be incorporated at the first stages of the design or the application development. For instance, PriS is an engineering method which incorporates privacy requirements early in the system development process (Kalloniatis et al., 2008). Privacy requirements are considered as organisational goals that need to be satisfied. PriS provides a description of the effect of privacy requirements on business processes, and it allows the identification of the best privacy-preserving system architecture.

### 3.3. Default privacy settings

A recent study shows that while there is some use of privacy settings by SNA users now, there is still a significant portion of SNA users who have not

---

Privacy and Social Network Applications

changed their permissive settings and allow unknown users to view private bits of information. SNA must clearly indicate the bare minimum of private information needed for a particular set of interactions (Krishnamurthy et al., 2008). The default privacy settings should be the bare minimum. Furthermore, a privacy tool to identify the metrics bare minimum would be useful. It would also allow users to compare various SNA with a clear comparative element of analyses (Krishnamurthy et al., 2008).

Another aspect of the problem is helping a new user to identify suitable default privacy settings. An interesting tool allows a choice of default privacy settings that have been learnt from a set of users' privacy settings. This solution has the advantage that users can easily understand and modify these settings to specify their desired initial policy (Ravichandran et al., 2008).

The adoption of a sort of proactive communication would be also useful. Indeed, there should be options for the user to easily report privacy invasions.

4. New challenges: Data Portability

Data mining and screen scrapping applications automatically infer real-world connections, and discover communities and individuals. Indeed, identifying consumer preferences is a key challenge in customizing electronic commerce sites to individual users. People linked in SNA often share preferences, allowing inference of interest in products based on knowledge of a consumer’s network friends and their interests (Hogg, T., 2009). SNA operators should also be aware of the importance of not only protecting user profile data, but the structure of the social graph. In September 2007, Facebook started making public search listings available to those not logged in to the site to encourage visitors to join the SNA. A recent study concludes that they should not assist data aggregators by giving away public listings because it allows data retrieval. (Bonneau et al., 2009). Another way of getting PII, particularly link privacy attack, uses the lookahead tool that indicates the friends known by a user. A recent study suggests SNA not to limit the lookahead to 1 or 2, whenever possible (Korolova et al., 2009). An extreme use of SNA, the so-called Antisocial Networks, are distributed systems based on social networking Websites that can be exploited by attackers and aimed at carrying out network attacks (Athanasopoulos, 2008). These examples are external discovery or disuse of PII.

The problem comes also from the inside. PII data from SNA have more and more applications to run with or “mash up” applications. In fact, a recent study found that about 90% of the top 150 Facebook applications get access to PII which should not be allowed to manage (Felt, Evans, 2007). Some possible privacy improvements in Facebook: one Facebook user can see the applications he has currently authorized going to Settings-Application Settings (Larkin, 2009). If a friend installs an application, the program will be able by default to see anything a user has shared with that friend. To restrict the data available, a user can go to Settings-Privacy Settings and click the Applications link. Then,
he can click the *Settings* tab on the top, and deselect any checked boxes on that page for info he doesn’t want to be shared. These settings only affect the applications his friends have installed. Afterwards he has to choose *Authorized* from the ‘Show’ drop-down menu (Larkin, 2009).

Google announced the adoption of standards such as Friend-Of-A-Friend (FOAF) and XHTML Friends Network (XFN) in their OpenSocial application to give access to the developers to the network graph, i.e. the map of connections between friends. This is data portability based on known standards, but in many cases there is no standard guidelines in the process.

That is the reason why Weiss proposes a privacy threat model for SNA portability (Weiss, 2009):

- Information privacy needs to be controlled on the data (PII) level.
- The user needs to be able to determine the sensitivity and context of the PII provided.

This seems a major challenge when we read about impressive tools such as CenceMe. CenceMe injects sensing presence into popular social networking applications such as Facebook, MySpace, and IM (Skype, Pidgin) allowing new levels of “connection” and implicit communication (albeit non-verbal) between friends in social networks. Sensing presence captures a user’s status in terms of his activity (e.g., sitting, walking, meeting friends), disposition (e.g., happy, sad, doing OK), habits (e.g., at the gym, coffee shop today, at work) and surroundings (e.g., noisy, hot, bright, high ozone) (Miluzzo *et al.*, 2009). An interesting solution for the context problem is DroPicks (Hosio *et al.*, 2007). The context is defined by the fact that many everyday artefacts are immobile, which implicitly restricts contents stored in a location. Such indirect, contextual sharing has advantages over direct communication mechanisms (E-mail, SMS, IM...).

- Privacy-preserving data portability can only work if the user can earmark the PII provided with individual privacy preferences.

In order to respect individual privacy preferences, the user self-control and the ease of public accessibility, further research is announced on semantic technologies for tagging data for context and purpose, transparency-enhancing technologies and Digital Rights Management (Weiss, 2009).

5. Future Trends: Privacy and Ubiquitous Computing

The deployment of ubiquitous computing casts doubt on the extent to which privacy is legally protected in public spaces (De Hert *et al.*, 2009). In case law, the European Court of Human Rights has introduced the notion of “reasonable expectation of privacy”. But ubiquitous computing is turning the reasonable expectation of privacy into an expectation of being monitored. Furthermore,
pervasive computing needs as many data as possible which clearly clashes with some of the main principles of data protection law. For instance, the data minimisation principle, collecting as little data as necessary, and the purpose specification principle only use the collected information for the purpose defined at the moment of data collection (De Hert et al., 2009).

With the emergence of Ambient Intelligence or pervasive computing, the definition of personal data needs to be reconsidered (Wright et al., 2009). Moreover, the distinction between personal and other data in a ubiquitous computing world is difficult to maintain. Perhaps it is time to data protection tout court. Indeed, the definition of personal data more and more implies the necessity of a case-by-case assessment, an approach upheld in a recent opinion from the Article 29 Data Protection Working Party on the definition of personal data. With ubiquitous computing, a new net, the Internet of things, will link many “anonymous” data to persons. Therefore, a lot of data mining and analyses will follow. As a result, all data will be personal data (Wright et al., 2009). In fact, in many cases it would not be even necessary to identify an individual in order to conduct commercially profitable operations. The unique identifier will be enough for intrusive marketing, for instance. So instead of identifiability as a criterion, privacy relevant data should rather be all those that can be used to affect our behaviour and decisions (Wright et al.).

Another important issue is the transparency of the processing. According to data protection regulation, data collectors and data processors should indicate the user which data are collected and give him or her basic information about the data processing. Nowadays, the user has not a comprehensive view of the data processing and its implications. In any event, such an information requirement might be unworkable with the generalization of the ubiquitous computing (Wright et al., 2009). What will become important in the context of SNA is the profiling knowledge, that is to say the access to the profile. This information could make comprehensible why the environment takes some actions, and could even help to prove liability in case of damage (Wright et al., 2009). On the other hand, anonymity can also be a useful tool with mobile community services (Demestichas, K. et al., 2009). PETs could provide important factual means of transparency. Transparency-enhancing technologies (TETs) could contribute to information exchange and management. An example of a TET is the so-called “sticky policies”, which stick to or follow data as they are disseminated. Sticky policies would provide clear information and indicate to data processors and controllers which privacy policy applies to the data concerned (De Hert et al., 2009).

---


The privacy concern about cloud computing can be a good analogy. As cloud services process users’ data on machines that the users do not own or operate, recommended privacy practices for developers are welcome (Pearson, 2009):

1. Minimise personal information sent to and stored in the cloud
2. Protect personal information in the cloud
3. Maximise user control
4. Allow user choice
5. Specify and limit the purpose of data usage
6. Provide feedback

Particularly, for mobile, ubiquitous social awareness applications, other complementary principles are (Raento, Oulasvirta, 2008):

1. Support lightweight permissions
2. Assume reciprocity
3. Make it possible to appear differently to different people
4. Allow for commenting, modifying and framing automatic disclosure
5. Provide for feedback
6. Allow the user to lie
7. Do not take control away from the user
8. Allow opportunistic use
9. Do not try to do everything within the system

In any case, Ambient Intelligence, or ubiquitous computing, requires a shift to privacy-by-design and PETs (Wright et al., 2009). Regulatory authorities and/or industry leaders could usefully encourage or formalise this option. Some research consortia, under the European Commission 6th Framework Programme, are good examples.

6. Conclusion

No single measure will adequately respond to the challenges to privacy posed by SNA and the ubiquitous Information Society. In fact, some combination of measures will be needed (Wright et al., 2009). Privacy principles, like Fair Information Principles (FIPs) or data protection principles are not enough. Perhaps, we are witnessing the very first stages of an important shift: the proportionality and transparency can widen the traditional data protection

principles. But, if we want to face to new possibilities of data portability or ubiquitous computing, we have also to encourage the adoption of PETs and TETs. Furthermore, this has to be done with privacy-by-design practices and not in a second moment of the implementation. We can go on indicating how these changes are transforming the privacy right, but perhaps it is time now to analyse and offer concrete solutions to concrete SNA. This should require a deep study of the privacy policies of a concrete SNA, and work together in the design of future applications, adopted with privacy-by-design practices.

References


