# Towards Privacy Propagation in the Social Web

Tyrone Grandison, E. Michael Maximilien
IBM Almaden Research Center, San Jose, CA 95120, USA

*The Social Web is one of the dominant aspects in a broader movement towards a programmable Web. A consequence of using the Web as a social substrate is that increasingly Web applications elicit and expose information that have various levels of sensitivity. Social data allows the creation of applications that are increasingly becoming vital to business users as well as individuals wanting to maintain connections with the social networks that they form. However, while social networking applications are increasingly becoming key hubs for our day to day interactions with the Web and colleagues, they are also increasingly creating a nightmare in terms of management of the privacy settings to protect the ever increasing mountain of social information.*

## 1. Introduction

The Web has transformed into a programmable platform. One of the main classes of applications that has resulted from this shift are social applications. These applications allow people to create connections to others, thereby creating a graph of human relationships. Nowadays, Web applications increasingly expose data and functions as application-programming interfaces (APIs), which allow the creation of new applications made up of combinations of the data, functions, and user interfaces (aka *mashups)*.

Indeed, mashups have accelerated the move towards a programmable Web and are increasingly showing up as components of a multitude of Web applications. This is leading to the creation of platforms for hosting Web applications components as mashups or widgets or gadgets. The Facebook application platform is an exemplar of such a platform. It already counts hundreds of millions of users and thousands of applications. Another example of this trend is OpenSocial (http://code.google.com/apis/opensocial/), which promises to enable social features and portable social graphs and social applications into any Web site.

While it is an exciting time to be a Web user or Web developer, it is also a scary proposition. Not only is Web data increasingly sensitive, it also represents a clearer mirror of real life data. Additionally, this sensitive data is also becoming sharable and reusable as part of APIs, mashups, and social platforms. An important implication of this side effect is that users should pay careful attention to privacy capabilities and settings of Web applications. However, due to the plethora of sensitive data and their usage, it also means that these artifacts are becoming more difficult for the regular user to grasp and carefully consider. The success of social applications has created an opportunity to reconsider how privacy settings and data should be configured, propagated, as well as potentially shared and reused.

## 2. Background

Privacy settings have appeared in various forms in Web applications [1, 2]. We can broadly categorize them as *mandatory* or *discretionary*. In the mandatory case, Web applications force users to accept the terms of service specified in their privacy policies. Google's search engine is an example of a Web application that employs this approach. End users have no control over privacy settings; they either can accept the policy, and its associated privacy provisions, or decide not to use the service.

In newer Web applications and services, due to the increased amount of personal data exposed, more control over the disclosure and usage of information is given to the user. That is, end-users have greater flexibility in configuring their privacy settings. This gives the end-users choices as to how much information they are willing to share.

While the discretionary approach seems to be the preferred one, it leads to a combinatorial explosion of possible policies and therefore to end-user confusion and frustration. The Facebook application platform is an example of a more discretionary approach and OpenSocial is an hybrid (or open) approach, since it requires privacy features, but leaves the details of the privacy approach to the social application development team or OpenSocial container provider.

## 3. Problems

We broadly categorize the current problems in the Social Web privacy space into three general categories:

1. **Data partitioning** - how should the user's data be partitioned into exchangeable granular pieces? This needs to reflect aspects of the data that are used by end-users as well as applications. For instance, what are the grouping of profile data which would correspond to the data that a user would like her friend to see as opposed to members of her network.

2. **Privacy settings** - what level of granularity is required for privacy settings? The settings must allow:
    a. partitioning - grouping of the settings to minimize user decisions
    b. elicitation - facilitate user decisions
    c. communication - exchanging the settings in an manner that is not ambiguous. This is especially important for communication of settings amongst applications.
3. **Management** - how are privacy settings and data managed? This involves:
    a. monitoring - data and settings changes
    b. enforcement - how are settings decisions enforced across application containers?
    c. sharing - how are settings shared amongst users of a group, networks, or friends?

Our categorization of problems is preliminary. However, we believe they are broad enough to cover the new issues specific to social applications.

## 4 The Model

As mentioned before, the explosion of applications and features and the commoditization and standardization of social application functionality means that there will be a lot more privacy questions that the end-user must answer. As with all things that require human interaction, if the complexity of the task exceeds a certain threshold, then the task will be mostly ignored or eliminated (in this case, turned off or set to the default setting).

In an effort to prevent privacy controls in social applications becoming irrelevant, either because they are too difficult to manage or too intricate to grasp, we believe that a model that enables the propagation of privacy settings based on the settings in ones personal network(s) should be defined. This model is a decision-support tool that allows a user to define base privacy settings. However, the user may customize her settings in whatever way she desires.

Our examination of the privacy capabilities in Facebook highlighted several features that the model should accommodate.

### 4.1 Rich Core Constructs

An assertion such as "I want person X, who is not a friend and who I have messaged or poked, to only see the Basic info section of my profile" highlights the need for entity qualification. Secondly, given the event-driven nature of social networks, performance of particular actions may be dependent upon the prior execution of some task or may invoke another action upon successful execution. Finally, the need to offer limited (and possibly transformed) functionality may necessitate the inclusion of pseudonymization techniques, e.g. a Facebook user stipulating that "Person X can find me in searches as ALIAS".

### 4.2 Intra-Social-Network versus Inter-Social-Network

For a specific social network, assuming a standard privacy model and propagation mechanism is a perfectly reasonable assumption. This means that within that space (intra-social-network), propagation of privacy settings can proceed unfettered. However, as society moves to social application platforms like OpenSocial, then either all the participants in the initiative must agree upon the notion of privacy and its propagation or each application can implement its own mechanism and the platform has to evolve to handle interaction between the different systems, i.e., privacy model integration.

### 4.3 Selective Bootstrapping

One may want to be selective, as your trust in the judgment of everyone in the network may not be equal, with respect to all things. You may want to segment who you trust to influence your privacy settings. This could be done based on the network they are in (e.g., you may only want your Silicon Valley friends' input when joining the *San Jose Entrepreneurs' Group*), your level of trust in them (e.g., you may only require help from highly-trusted friends) or some other criteria.

## 5 Related Work

While the literature on privacy for the Web has grown over the years, there are very few related works around the propagation of privacy settings and elections in social networks or social graphs.

## 6 Call to Action

We believe that the time is ripe for the privacy research community to start addressing privacy support in social networks and graphs. Traditional privacy research never looked at how a user's privacy settings and elections are affected or can take advantage of the groups that the user belongs to. In real life, a lot of a person's decisions for privacy is made based on the relationships that the user has with organizations, as well as other individuals. We believe that on the Social Web the same will be true and privacy propagation approaches will help untangle the privacy settings overload that we are increasingly creating.

## References

[1] W3C, Platform for Privacy Preferences, http://www.w3c.org/P3P

[2] W3C, Enterprise Privacy Authorization Language (EPAL 1.2) http://www.w3.org/Submission/2003/SUBM-EPAL-20031110/