Can Users Control their Data in Social Software?
An Ethical Analysis of Data Control Approaches

Rula Sayaf∗, James B. Rule† and Dave Clarke∗
∗Department of Computer Sciences
iMinds-DistriNet
KULeuven
Leuven, Belgium
Email: firstname.lastname@cs.kuleuven.be
†The Center for the Study of Law & Society
Berkeley Law
University of California
California, USA
Email: jbrule@berkeley.edu

Abstract—The concept of data access and data flow control has a central role in privacy preservation of users in social software. Certain approaches aim at enabling users to control who can access their data, and identifying who is accountable for misconduct. Those approaches are implemented in technological and legal frameworks.

Each approach raises many ethical issues. Furthermore, these approaches may conflict with each other. The issues are more complicated when the technological and legal frameworks are compared to each other. Such conflicts raise an open issue of what the appropriate balance would be when implementing the approaches by the two frameworks at once.

Index Terms—social software; privacy; access control; online social networks; ethics; data protection; law;

I. INTRODUCTION

Social software offers users the ability to interact with friends by sharing personal data and establish small social groups. Those features come often at a price in terms of privacy. Some commentators consider any such consequent bad results as costs of innovation, while others consider such costs unacceptable [1]. Privacy issues may dramatically affect users lives for which the loss of control over data is viewed as one of the main causes [2].

Privacy is the right of an individual to decide the uses of personal information and the circumstances under which this information is communicated [3]. According to the “privacy as control” paradigm [4], data subjects should control who access, handle, and process their data. To some extent the paradigm implies that users should be accorded veto power over the uses of their data. However, when the concept of privacy is implemented in the online world created by social software interactions, ethical issues arise. Such issues are related to determining the degree of control over data, without violating other users’ privacy. A set of relevant ethical questions arise here, such as: How can he be certain that the audience he selected and no one else will see it? And is the social software provider part of the audience? What if the data is a murder plot? Should that data be protected from uses that he disapproves of?

Various disciplines focus on approaches to enhance users privacy by facilitating more control over their data, such as access control models (ACMs). Many factors come into play when an approach aims at empowering users with as much control as possible: technical considerations, economic interest of social software providers, social norms of communities and legal aspects. The interplay of these aspects causes conflicts and ethical issues. In this paper, we address whether full control should and could be granted for social software users or not, and discuss the related ethical issues. We consider how technical approaches, namely ACMs, may assign control to users over their data to address privacy issues. We briefly discuss the main requirements of such models and the related ethical issues, the complementary accountability and audit approaches. We look at the implementation of access control and accountability in the legal framework. By contrasting the two implementations—technical versus legal—we expose conflicts and ethical issues. We conclude by stating how those conflicts affect the privacy of users where efforts and collaboration between the two frameworks are required to resolve the underlying ethical issues.

II. THE PROBLEM

Controlling data disclosure is an essential act of privacy preservation [5]. From a technical point of view, ACMs solve privacy issues that result from the lack of control by facilitating as much control over data as possible for users [6]. To ensure that the access control policies (ACP) are respected and enforced properly, the principle of accountability is required [7]. Accountability is the principle of verifying the behaviour of users in social software and identifying any user or entity that is responsible for misconduct [8, 9]. The principle of
accountability is achieved by auditing the logged actions of users [10].

Problem 1: Users cannot have absolute control over their data since ACMs cannot control and protect certain types of data. ACMs vary in the level of granularity of control they offer. However, there are certain types of data that cannot be controlled, such as relational and behavioural data.

Problem 2: Users cannot have absolute control over their data and be certain that their control is respected without having other entities such as the social software providers monitoring and controlling their behaviour. Upon enforcing ACPs, problems might occur, for example due to a conflict between positive and negative policies [11], which result in disruption of the control users assume is being enforced. Accountability and audit approaches can verify how the control users want to enforce is performed. However, such approaches require monitoring and auditing users’ behaviours and data, which is a type of surveillance and is perceived as a privacy violation. Thus, in order to control ones’ privacy, his privacy might be violated by an authority to ensure that there is no violation by other users.

Accountability and audit approaches are complementary to ACMs, but the boundary of the control offered by both approaches is open to dispute. In many cases there are conflicts in assigning the appropriate control to related parties within the social software.

In contrast to technical approaches stands the principle of accountability as implemented by the European legal framework in the provision of Directive 95/46/EC [12], which implicitly covers access control. The Directive states how data should be handled and processed and under what conditions. Whenever there is a violation, the responsible entity—defined by the Directive—would be held accountable.

Problem 3: Absolute data control by users, runs counter to what is practically possible and dictated by law. Most people would agree that a user who posts data has some sort of special interest in, and a role in deciding, its flow [13]. But it would be unusual to propose that anyone should totally control their data and approve how it is used or processed, even if this conflicts with the concept of privacy. In situations where data affects other users, or relates to criminal and illegal acts, a certain level of supervision is needed. Such issues can only be resolved by ethical and policy choices governing the appropriate amount and boundaries of control over data assigned to users. The limitation of control in the perspective of accountability conflicts in principle with the surge of empowering users with as much control as possible in the perspective of ACMs.

Organisation. Section III introduces the preliminary notions of social software. Section IV presents data access control and the related ethical issues within the technological framework. Section V presents data control via accountability and audit and the related ethical issues. Section VI presents the data control in the legal framework. Section VII compares the two frameworks. Section VIII concludes the paper with the open ethical issues that need to be addressed in the future.

III. Preliminaries

Social software users can exchange the data about themselves with others. A data subject is the user who posts data via social software and defines an appropriate ACP, which includes the constraints of who can or cannot access and for which purpose. A subject initiates an interaction to communicate with other users by sharing his data, e.g., creating a post. Each interaction is viewable by the audience. We distinguish two types of audience: potential and actual. The potential audience is the set of users, defined by the data subject who can view data such that those associated with an interaction. The actual audience is a subset of potential audience who do view data mentioned in the previous definition.

IV. Data Control via Access Control

An ACM offers users methods to preserve the privacy of the data linked to them within an interaction context. The act of taking data out of the context it was disclosed in and disseminating it in another context is considered a privacy violation [14, 15], although in certain cases it might be necessary to perform such an action for the interest of the public, e.g., reporting of a criminal. The control of data over any context is required to mitigate such violations [16] and is required for a fine-granular control [17]. In general, an ACM should address the following principles or requirements [11]:

1) Protection of Data: An access control policy should be defined for any type and form of data whether it is a single data item e.g., a photo, a set of data items e.g., comments about a photo, data about relationships [4], or data about behaviour.

2) Linkability to the Data Subject: A data subject can protect its privacy by defining an ACP over their data item, and therefore, the data need to be linked to its subject. This is achieved by enforcing the ACP of the subject and not allowing changing it unless the subject gives a permission.

3) Control of the Audience: The subject must have control over the audience, and be informed whenever a user becomes a member of the potential or actual audience.

4) Control of Context: The data subject should be able to control the context in which data is disseminated. Any change of the current context data is put in should be detected and the subject has to be informed of this.

A. Ethical Issues

Here are some key ethical issues raised by ACMs

1) Protection of Data: the protection of relational and behavioural data is not currently possible. This issue conflicts with the principle of offering fine-grained control, and results in no control over such types of data.

2) Linkability to the Data Subject: The use of sticky policies [18] guarantees that an ACP defined by the data subject is always enforced in any context the data is put. Without sticky policies a user can disclose a data item of another subject and change the data’s ACP. There is no criteria to limit such acts unless the social software provider is actively involved. The data disclosed via social software is controlled by the provider who enforces the data subjects’ ACPs. Consequently,
the provider has an important role, but ACMs do not model the provider. The control offered to users by an ACM does not constrain the provider. As a result the linkability of data to its subjects cannot be guaranteed by ACMs. In models of client-based enforcement, the aim is to decrease the involvement of the provider as much as possible, although it is the provider who runs the social software. The economic interests of the provider should not bias ACMs, rather they should model the provider and offer privacy solutions for real-world social software that are managed by providers of certain interests.

3) **Control of the Audience**: The act of viewing a data item by the audience is behavioural data that belongs to that audience who should be able to control such data. Moreover, the audience should be able to control what data and from which data subjects they would like to view. As long as this type of control is missing, the control assigned to users is not equally distributed.

4) **Control of Context**: ACMs do not offer fine-grained control over any context in which the data is put. When subjects’ data is logged to be audited, the data is put in another context that subjects. If the data subjects do not want the data to be disclosed to such context, then auditing cannot be performed. However, data subjects cannot have this level of control, and this raises the issue of who defines the contexts that users may have control over? How can such a decision be justified?

V. **Data Control by Accountability and Audit**

Accountability and audit approaches control users' behaviour to detect when a violation of terms of conduct occurs [9]. Users that are highly concerned about maintaining control over their data would benefit from audit approaches. However, those users have to accept that their data, how they behave and how other users behave would be monitored.

Audit approaches raise issues of being privacy violating by means of surveilling user data in order to protect users privacy. Those approaches also conflict with ACMs because they require users' data to be processed in a context users have no control over as discussed above.

Next, we discuss how access control and accountability are implemented within the legal framework, and contrast them to the technological framework discussed above.

VI. **Data Control by Law**

In this section, we compare the control of data possible via the legal framework through Directive 95/46/EC to that possible via the technological framework through ACMs. While the data subject is considered as the only entity that decides how the data should be accessed and processed in ACMs, the Directive distinguishes two entities instead: the data subject and the controller, which can be the social software provider or a third party application provider. The controller “determines the purposes and means” of the processing (article 2(d)) and is responsible for ensuring compliance with the Directive [19].

We are interested in investigating if data is controlled in the Directive as it is controlled in ACMs, and whether the Directive addresses some of the limitations of ACMs. We focus on the same principles of ACMs control, elaborate on them and highlight the conflicts and ethical issues as follows:

1) **Protection of Data**: A data subject controls his data for personal use; such control is inapplicable, without the consent of the data subject, if the data relates to criminal acts or security measures [12]. The issue here goes beyond the control of criminal data, rather, it is about how such sensitive and criminal data is detected. Recently [20], an incident was reported about FBI investigating some people about terrorism by analysing their data—likes, posts—on Facebook. Such investigation is arguably beneficial for society and national security, but the issue here is to which extent the data can be controlled and monitored for the early detection of such criminal and security-threatening acts.

2) **Linkability to the Data Subject**: A data subject is not considered the controller of his data according to the personal use exemption of the Directive, which directly conflicts with the perspective of granting fine-grained control to data subjects of ACMs. The article 29 working party considers that, data subjects or users can be data controllers of their friends data only under certain circumstances.

3) **Control of the Audience**: The Directive does not state anything directly related to the audience of data. However, it is the data controller—and not the data subject— who may control the audience by defining who may process the data in a way that complies with the controllers policy (article 16).

4) **Control of Context**: According to the Directive (articles 10,11), when data is not obtained from the data subject, he must be informed of at least the identity of the controller and the purpose of processing and he has the right to rectify inaccurate data [19]. By the required consenting stated in article (7), data subjects can control of whether their data can be disclosed and processed in a certain context—unless the processing is required for legal reasons or protecting the interest of the data subjects or the public.

The legal framework offers control for users over how their data is handled by social software providers and third party applications, but the problem is that users are not properly acquainted with the existence of such laws. A ‘privacy notice’ is documentation of how the social software provider controls data to satisfy data protection and privacy requirements. The Directive suggests that privacy notices can serve to acquit the data controller of his transparency obligations under articles (10,11), but it does not require the notice to be displayed to users [19]. Users have inaccurate perception of their knowledge about privacy technology and vulnerabilities [21], and without being aware of the data protection laws such a misconception could not be corrected. Moreover, the theoretical value of privacy notices is challenged on the practical level by being difficult to read, and that individuals mostly do not read or cannot understand the abstract-level of description [19].

In comparison to ACMs, the legal framework has standard definitions of what is allowed and what is not allowed. The main conflict with ACMs is that the notion of control over data is handled and assigned differently, which might result in
conflicts when the two frameworks are applied on the same social software.

VII. WHAT KIND OF APPROACHES ARE MOST ETHICAL?

The technological and legal frameworks currently work on solving privacy issues separately, which raises ethical issues and conflicts between the two frameworks. ACMs may be impractical and too complex to use. But they aim at making users capable of controlling their data disclosure and dissemination where the law has no effect such as in the personal use exemption. Nevertheless, legal frameworks facilitate more control over the behaviour and conduct of social software providers and third party applications than ACMs do.

In our perspective, the most ethical framework possible is that integrates both technological and legal frameworks, where the ethical issue may be minimized and probably other issues may arise.

VIII. CONCLUSION

In this paper we have discussed approaches to preserving privacy by means of data control mechanisms. We have presented the technical approaches of access control and accountability mechanisms and discussed some possible ethical issues. We have contrasted these approaches in the technological and the legal frameworks. By comparing the two frameworks, we have identified conflicts that raise an open ethical issue about how such frameworks could be utilised together to facilitate more privacy preservation by means of data control.

ACKNOWLEDGMENT

This research has been funded by the IWT in the context of the SBO project on Security and Privacy for Online Social Networks (SPION). The authors are also thankful to Brendan Van Alsenoy.

REFERENCES