Facebook Goes to the Doctor

Matt Bishop
Dept. of Computer Science
University of California at Davis
Davis, CA 95616-8562
bishop@cs.ucdavis.edu

Carrie Gates
CA Labs
520 Madison Avenue
New York, NY 10022
carrie.gates@ca.com

Peter Yellowlees
Dept. of Psychiatry
UCD Health System
Sacramento, CA 95817
peter.yellowlees@ucdmc.ucdavis.edu

Gabriel Silberman
CA Labs
520 Madison Avenue
New York, NY 10022
gabriel.silberman@ca.com

ABSTRACT

The use of computer-based social networks for health care changes the privacy paradigm of face-to-face treatment. For example, in an office, a patient can be reasonably sure that the physician or therapist is the only one present, and is who has been providing treatment. On a computer-based social network, communications travel over the World Wide Web, raising the possibility of eavesdropping, delay, and other problems. Further, verification of the party with whom the patient is communicating is more difficult, and to many less credible, than in-person verification. This paper describes the privacy paradigm, presents a set of requirements for effective use of computer-based social networks in health care, discusses what current technology can provide, and what gaps must be closed to meet the rest of the requirements.

1. INTRODUCTION

The property of computer-based social networks that sets them apart from ordinary social networks is the ability to connect, in real time, people who are geographically distant. People can locate old friends, make new ones, and communicate with them even if they are in other countries. Perhaps more importantly, they enable people of like interests, beliefs, and other attributes to meet and interact in a way not possible before the World Wide Web. Thus, people with a rare illness can interact with others who have the same illness in a "virtual" support group. Computer-based social networks also enable a physician or a therapist to interact with these patients and provide care to people in places that would normally be inaccessible. Similarly, in today's mobile society, psychological and psychiatric clients may have to travel as part of their job. Such networks would enable them to keep appointments with their therapist in a cost effective manner even when traveling abroad.

Existing networks allow users to provide support to other users. From the 1980s, USENET provided support for people in technical disciplines, and the flexibility of that system led to social support groups developing rapidly. Professional colleagues can keep in touch with each other's jobs and work using LinkedIn. Frequent travelers can see when they will be near friends who also travel using TripIt. Facebook and Google+ enable people to meet others with particular interests, hobbies, and life styles. Going from these to medical support groups is straightforward.

Indeed, the use of the Web to establish medical support for patients has already begun. The site "Patients Like Me"[1] allows people with medical conditions to share their symptoms and treatments with others similarly situated, and to talk with those other patients to see what they have tried, what worked, what didn't work, and what the effects of various treatment programs were. Other web sites provide information (of varying degrees of reliability) about various illnesses and possible treatments. These "socio-medical networks" are becoming more common as time passes.

They also raise several interesting questions, many involving security. What is "security" in the context of web-based treatment? What issues does the online, interactive environment described here raise that face-to-face treatment does not? How does the online environment change the issues that are in common with face-to-face treatment?

This paper explores the governance issues relating to security in extending the use of the web to provide interactive talk therapy and treatment much as it is done in a physician's office. We briefly review some of the issues. Then we discuss several requirements that arise from the issues, and explore the use of current technology to meet them. We identify gaps in the existing technology, and suggest directions for future work to close these gaps.

2. BACKGROUND

We consider three categories of support groups [6].

A *self-moderating group* is one with no officially designated leader. The individual group members act as they think best to provide support to other members. The key characteristics are that everyone in the group sees all messages, and that membership of the group is decided by the

[1] http://www.patientslikeme.com

members of the group. This may create problems, especially when the medium is available to all. We do not consider this type of group in this paper, because it is typically formed *ad hoc* and not conducive to providing sustained treatment.

A *facilitated group* is one with a facilitator who helps the group members interact, but does not direct or provide insight. The facilitator, in conjunction with members of the group, decides membership—although exactly how this is done is up to the facilitator and members. The facilitator may delay messages among group members, or to the entire group, to ensure his own messages arrive first. This enables him to deal with messages that may cause problems for members of the group.

A *moderated group* is like a facilitated group, but the moderator exercises greater control over communications within the group. Like a facilitator, a moderator can delay messages. Unlike a facilitator, the moderator may block messages, may modify messages to inject her own thoughts, comments, and ideas (of course, with attribution of any changes), return messages to the sender with suggestions or comments, or pass messages through unchanged.

In the latter two groups, the facilitator and moderator may be the primary caregiver to members of the group. They may also be secondary caregivers who interact not only with the group but also with the primary caregiver to be sure that all responsible parties are current on the condition and events affecting the patient. For clarity, the person or people who are responsible for the treatment of the patient (or patients) are called the *caregiver* and the facilitators or moderators are called the *leaders*.

In some cases, patients need to be informed and give consent. In fact, the patients may not be able to give informed consent because they are incapacitated. Then their guardian(s) and caregiver(s) may need to give consent. In this paper, we include this case when we speak of a patient or member taking some action or being informed of something.

These groups may be confined to a single governing jurisdiction, or may span multiple jurisdictions. This raises issues of governance: whose rules apply, or how are the rules composed and reconciled so that the needs and expectations of everyone involved will be met?

We begin with the following principles:

1. *Informed consent.* The members of the group must be able to learn the rules under which the group functions.
2. *Protection of privacy.* The members of the group, and the facilitator or moderator, must control what information is shared with other members of the group and with people not associated with the group (subject to legal and medical requirements).
3. *Integrity of information.* No messages may be altered except as authorized by the structure of the group. So if Anna sends a message to a moderated group, only the moderator should be able to alter it.
4. *Assurance of attribution.* All attributes of a message (such as origination) should be correct or omitted.
5. *Protection of the patient.* The purpose and actions of the group must not harm the patients.
6. *Adherence to applicable laws and regulations.* The system must be able to determine the correct jurisdiction(s) and then adhere to (or resolve) the (possibly conflicting) governing regulations.

We then define the notions of a "policy" and "secure".

Definition. A socio-medical network *security policy* defines the requirements that the design, implementation, and management of the network must meet. These requirements must balance the above principles. The network is said to be *secure* if it meets these requirements.

3. REQUIREMENTS

We focus on the first four principles, which lead to several *technical* requirements for socio-medical networks. The final two principles are non-technical, and not discussed here.

3.1 Informed Consent

To treat a patient effectively, the patient and the caregiver must know what the treatment may consist of, and how it is supposed to work. They also must know what information is recorded in medical and other records, who can access that information, and how it will be used. This leads immediately to the question of observing and recording sessions.

Anyone can enter a room and sit quietly as the session is proceeding; handling this requires proper procedures that the members and leaders follow. Similarly, any of the systems being used in the session may have vulnerabilities (including malware) that allows unauthorized observers. These vulnerabilities can be remedied—if they are known. There should be mandatory procedures for securing the workstations that members use for sessions, or the members should be made aware that no such procedures exist. A complication is that, in this environment, the end users may be unable to learn, or even understand, workstation security.

REQUIREMENT 1. *The members must know of any procedures for having an observer to the session. Members must follow any procedures for blocking unauthorized observation.*

A general rule is that the members of the group should know whether they are being recorded; applying this rule to an on-line session, they should know whether a log is being created. They should know what exactly is being saved: the actual, entire session, parts of the session, summaries, times the sessions ran, or who was present at the session.

An interesting aspect to this is assurance. What evidence is needed to demonstrate to the members that only the identified aspects of the session are (or are not) being saved?

REQUIREMENT 2. *The members of the group must know exactly what aspects of the session (if any) are being saved.*

Records or logs are made in anticipation of someone using them. Models such as the Clinical Information System Security model [1] provide bases for allowing or preventing access. Different parts of the world will have different requirements for controlling who may access the records. Here, we simply note that the patients, leaders, and caregivers must know who may access the records, and have confidence that *only* those people can access the records appropriately. Further, unless there is a compelling reason, the members participating in the session should know who accesses the session record, and when.

REQUIREMENT 3. *Access to records of the session must be restricted to those authorized to access the records, in the manner authorized.*

3.2 Protection of Privacy

A critical aspect of protecting the patient lies in protecting the patient's privacy. Specifically, the mishandling of personally identifiable information (PII) can cause problems

for the patient ranging from embarrassment to marital, occupational and legal difficulties. Information that does not include someone's name can uniquely identify an individual. An example of such a *quasi-identifier* (QID) is a social security number in the United States, because each social security number refers to one person. In some cases, what constitutes a QID is not obvious.Studies have shown that a ZIP code, date of birth, and gender can uniquely identify between 63% to 87% of the U.S. population [7, 14]. Finding these relationships is an open problem, and one that is critical to data sanitization [3].

If membership in the group is to be anonymous, the PII must be hidden from others in the group. But the purpose of the group is to treat patients, and as part of that treatment they may reveal PII such as their name. This is especially true of participants in distress, who will not be thinking clearly. For membership to be anonymous, messages must be examined and possibly redacted before transmission to the full group, a function presumably done by the leaders or their designees. We discuss this assurance issue later.

REQUIREMENT 4. *Someone must determine what information a group member can share with the group without revealing his or her identity to others.*

Another key question is to what degree the *leaders* should be anonymous. In face-to-face therapy, the patient will know leaders' names and something about their qualifications, but not information such as where they live. In virtual groups, the leaders can become largely anonymous.

REQUIREMENT 5. *A leader must be able to determine what information he or she can expose and preserve the desired degree of anonymity.*

One non-technical question is whether the individual should be able to reveal his or her identity. For some forms of support groups (e.g., some forms of mental illness), the individual may not appreciate the consequences of doing so. Countering that is the impossibility of preventing this. For example, a participant could tell someone to go to a particular website that contained the participant's identity. We regard this as out of scope of our work, because it requires an evaluation of the patient's mental state and circumstances.

3.3 Integrity

When one communicates within a group, the leaders will base their responses (and, presumably, any treatment) on the contents of the message. This means the leaders must see the raw message so they can react appropriately. Similarly, their messages must arrive unchanged.

REQUIREMENT 6. *Message integrity must be preserved between members of the group and the leaders.*

However, the leaders need *not* pass along the messages unchanged. It may be necessary to redact information or comments to have the group focus on something in particular, or to minimize distractions. If the leaders can change messages, or interfere with them, the group members should know.

REQUIREMENT 7. *If leaders can modify, delay, or delete messages, then group members must be aware of this ability.*

Assuming that the leaders can tamper with messages and the flow of messages, it is reasonable that the group be informed when this is done. In unusual circumstances this may not be appropriate. In either case, the leaders should decide the general policy and inform group members of it. Otherwise, a member may impute a delay in another member responding as being due to that member rather than a delay imposed by the leader.

REQUIREMENT 8. *If leaders modify, delay, or delete a message, the policy for informing group members must be disclosed in advance.*

Another aspect of integrity is that messages be ascribed to their sender. Complicating this is the possibility of anonymity, where members of the group or the leaders do not want to reveal their identity. In such cases, the messages may be completely anonymous or they may be sent from a pseudonym that remains consistent across messages (pseudonymous). In either case, the sender must not be able to attribute the message to another:

REQUIREMENT 9. *Each message must either be* completely *anonymous, or identify the (possibly pseudonymous) sender of the message; it may not identify anyone else.*

Finally, the group should be available when it is supposed to be, and those charged with responding to requests or facilitating the group session should do so in a timely fashion (or inform group members of their inability to do so in a timely fashion):

REQUIREMENT 10. *The communications medium, and supporting infrastructure, must be available.*

Many of the requirements described in this section are unique to the digital medium. In a face-to-face group session, it would not be possible for a "message" (someone speaking) to be modified before others in the group heard it. We believe that providing this capability is important, but that groups might opt out of using it in favor of an experience that more closely resembles a face-to-face meeting.

Underlying the informed consent, privacy of patients, and integrity and availability are assurance: how do we know the group mechanisms and supporting infrastructure will provide what is needed?

3.4 Assurance Issues

Integrity begins when information enters a system; its reliability, timeliness, and other characteristics must be assessed. This is a basis for assurance—gathering evidence that the information is indeed correct. Here, we trust the information because in a technical sense, we cannot assess how accurate it is (but note that the caregivers can). Instead, we focus on its attributes. Consider the release of information as a basis for this analysis.

Releasing information can occur in real time or after the fact. The former case occurs during the session, because people will interact with one another, releasing information to the participants and any observers. The latter case arises when the contents of the session are made available after the session is completed, for example to an insurance company. As the record will include information about *all* participants, PII of those other than the person(s) involved should be suppressed. Attempts to suppress PII require an analysis of the data being disclosed to discover not only names and other identifying information, but also relationships among the suppressed (anonymized) and unsuppressed data. The difficulty of uncovering these relationships becomes clear when one realizes that the relationships may not arise in the session data, but in a combination of external and session data—something the sanitizer may not know about [2]. As

of now, sanitizing the data in real time and providing a high degree of assurance that the data cannot be deanonymized is not feasible. The difficulty of sanitizing the data in the second case is unclear.

An interesting approach is to examine this problem using an attribution framework, because protecting privacy requires preventing the attribution of the information to an entity (or set of entities). One such framework [4] identifies nine parties, seven of which are relevant here: the sender, receiver, ISPs, backbone providers, and the political jurisdictions of the sender, receiver, and over which the messages transit. This suggests three groups of interested parties.

First are the external observers. Either they are analyzing logs of the session, which we discuss below, or they are observing the session in progress, in which case they can be treated as (passive) participants. This set includes the ISPs, backbone providers, and political jurisdictions, because they may be able to observe attributes of the session by monitoring traffic, even if the sessions are encrypted.

Second are the participants in the session, the senders and receivers. Presumably, those not wishing to identify themselves will not do so. As noted above, sanitizing the messages being sent is infeasible. It may in fact hinder the treatment that the session is intended to provide, especially in support groups such as Alcoholics Anonymous.

Third are the group leaders, also senders and receivers. If membership in the group requires the consent of the leaders, then the leaders must determine that the prospective member meets the membership criteria. Often, identity is a component of this determination, especially if the leaders must verify some specific condition (which may be viewed as an attribute) in order to determine eligibility. In the non-cyber world, the leader seeing the prospective member may supply the needed attributes. But there is no physical presence in the cyber world.

A trusted third party may verify the attribute, and generate a certificate containing the attribute and the issuer's (digital) signature. Then the certificate attests to the presence of the attribute, and the leader can determine whether to allow the prospective member to enter the group.

REQUIREMENT 11. *The leader must be able to validate the information given by the applicant to decide if the applicant should be admitted to the group.*

This provides a mechanism for reversing privacy. Leaders may need to reveal identity to comply with laws[2] or to protect the member[3] A requirement for getting a certificate to join the group could be that the *issuer* know the subject's identity. Then the leaders and the issuer could together take the appropriate action. This also has the advantage of putting a second person in the deanonymization process, thus enforcing the principle of separation of privilege [12].

Using certificates, or some other key (attribute) management system, requires support. Constructing such systems and making them easy for naïve users to use is difficult [16]; witness the dearth of such systems now. But the supporting infrastructure, and the ability to use it, must be present.

The details of this infrastructure, and its management, falls into the realm of governance. Governance speaks to policy: should there be one infrastructure for all socio-medical networks, or many such infrastructures with different properties and assurance levels? If there are many, how do you identify policy conflicts and deal with them? In fact, how do you identify the components to which the policy applies—for example, messages may transit networks over which the managers of the infrastructure have no control.

As an example of a governance issue, consider the use of originator-controlled access control (ORCON) models [8,10] for protecting session data and PII. In some sense, the "originators" of the session data are the members of the group, and the "originators" of the PII are those whom the PII identifies. Thus, an instantiation of an ORCON model supports control of that information. As noted in section 6, existing instantiations have been quite unsuccessful.

REQUIREMENT 12. *If certificates are used, there must be one or more public key infrastructures and attribution infrastructures that support the socio-medical network.*

Finally, modeling trust in these groups will highlight the assumptions of existing groups, and indicate what assumptions should (or must) be made about new groups. The requirements must reflect the importance of privacy, confidentiality, and integrity in this realm.

REQUIREMENT 13. *A trust model that facilitates analysis of dependencies and assumptions must be created for each group and infrastructure.*

3.5 Summary

Given these requirements, the question is whether the requirements can be met. In the next section we present three case studies, and use those studies to examine what can currently be done and what future work will be necessary to support all these requirements.

4. THREE CASE STUDIES

Social networks offer astonishing potential for rich healthcare interactions, many of which have not been common in the in-person world, or have not been possible at all. As a result, innumerable use cases can be developed. To examine their range, we start with the possible interactions and relationships within a social network environment, add on the reasons for those relationships, and finally take account of the differing content or types of diseases that might be of importance and of interest to the users.

The four main types of online relationships are [18]:

1. Patient to provider
2. Patient to patient
3. Provider to provider
4. Patient or provider to almost anyone else

These relationships can be 1:1, 1:many or many:many. They can involve people from different cultures and countries using different languages and with widely varying social expectations. People can communicate in real time (synchronously) using instant messaging, telephony or through videoconferencing, or in delayed time (asynchronously) using email, blogs and general website content access.

People interact about healthcare using five core paradigms:

1. A *clinical provider-patient interaction* to obtain an opinion, treatment or help. This is the most common healthcare interaction in the in-person world, and these

[2]For example, in California if a patient threatens to kill someone in the future, patient confidentiality does not apply to the threat.

[3]For example, if the member threatens to commit suicide.

interactions will undoubtedly occur over social networking as long as appropriate security and privacy are available for all concerned.

2. A *patient-driven therapeutic interaction* to obtain support and help in dealing with a medical problem. Patients support one another in these interactions, in support groups such as Alcoholics Anonymous.
3. An *educational interaction* to learn about a specific disorder, drug, therapy, operation or provider. On the Internet, 60% of all healthcare searches and online questions are by people searching for information to help a family member [17].
4. A *research interaction*, usually to collect or analyze data or information. Typically providers and researchers drove these in the past, but the rise of patient-focused research makes this type of use within social environments much more likely in the future.
5. An *administrative interaction* such as scheduling appointments and procedures, paying bills, authorizing care, ordering tests and medications.

Finally, people interact for content-related reasons. They may want a simple answer to a straightforward question, or be searching out a sophisticated opinion from a world renowned expert. They may be physically of psychiatrically unwell and in need of urgent care, or may be bored and simply surfing for therapeutic minutiae. They may be interested in cancer or in sexually transmitted diseases.

These five paradigms of interaction show that there are many variable and continuously changing interactions that could occur within a social network environment, just as there are similar ranges of interactions in the in-person world. Social network applications differ from most in-person environments in the potential for very rapid interactions in multiple sequences and configurations. The use cases that may occur in social network environments consequently vary from the very simple, to the almost impossibly complicated, as illustrated below.

Case Study 1. This is a simple patient-patient interaction as happens in many online patient support groups. Patient Anne seeks out another patient with bipolar disorder to ask some straightforward questions about her experience with the drug lithium carbonate. Anne learns useful information from the other patient, who also suggests a website for her to visit and find more information. Anne takes this new information to discuss her treatment regime when she sees her in-person doctor.

Case Study 2. This is a patient-doctor interaction that developed from an online group. Patient Bill has Tay-Sachs disease, which is an unusual disease, and is part of a social network environment that contains a number of doctors interested in that disease. Doctor Caroline answers one of his questions in a very helpful manner in a public forum, but Bill decides that he would like to ask her some more questions privately, rather than reveal too much information about himself to the rest of the group. The patient and provider are able to withdraw from the group into a HIPAA compliant private "room" that is part of the social network environment, and have a private conversation. They can choose to interact by phone or videoconference, either immediately or as a follow up telemedicine consultation.

Case Study 3. This is a group interaction that develops into a research study. A large number of patients with anorexia nervosa use the social network for mutual support, and to self-educate and interact with several therapists. Some members read about a vitamin that is supposed to be particularly therapeutic for them, but are unsure how effective it is as no formal drug trials exist. They approach a renowned researcher who agrees to assist them in designing and carrying out a therapeutic trial of the vitamin over several months. This involves the use of multiple private rooms for clinical assessments, joint writing sessions to develop the protocol, conferences and meetings to oversee the research project, and numerous differing levels of privacy and security within the social network environment. The project involves patients and researchers from multiple countries, using different languages.

5. EXISTING MECHANISMS

Computer security provides many mechanisms and policies for dealing with threats. Here, we examine our case studies to see what existing mechanisms support them. We make one key assumption:

The end point systems are not compromised.

Without this assumption, any mechanism can be defeated easily, for example by simply recording the keystrokes being entered and the data (text and images) being displayed, and photographing the users.

Case Study 1. This case study differs from Patients Like Me because we assume that the patients have an email address for one another. The email address may be an anonymous one, so the patients may not know each others' real identities. We also assume the patients are put into contact through some external mechanism.

Once in touch, the patients need some form of secure communication and secure storage for their messages. These problems are well known. Secure email systems such as PGP will provide the necessary support, as will simply storing encrypted messages. Given that the endpoints are not compromised, the cryptographic keys entered on those systems will not be compromised either.

The cryptographic mechanisms can support anonymity, through the use of Persona certificates [9]. This requires an infrastructure to support the use of certificates, which raises issue discussed in section 6.3. Other mechanisms not involving certificates use third parties to act as authentication servers. These parties need not require verification of identity before issuing (anonymous) interchange keys.

As of now there is no way to assure that all parties know whether any party is saving messages. But this evades the underlying issue of deniability. Another party may forge a message and claim it is from their peer. This is similar in nature to someone generating a letter in the "real world" and claiming that it is from someone else.

If social media like Facebook are used, then both parties know the identity of the other. Again, that identity might be a pseudonym, whether officially supported by the social network or otherwise. Presumably the two parties will use the email feature of the social network rather than, for example, posting publicly on each other's wall. So their security and privacy is limited to what is provided by that network.

Case Study 2. For this case study, we assume that the support group area requires a login. The discussion might be public or accessible to others who log into the support area. MedHelp[4] operates in this fashion. We extend this

[4]http://www.medhelp.org

notion to a support group that is focused on a single disease, and thus requires:

1. that the user have that disease or be supporting someone with that disease;
2. that the interaction be immediate (e.g., chat) rather than asynchronous (as in the case of MedHelp); and
3. that the interaction will be publicly viewable by all those either currently logged in, or historically for users who log in at a later time, but not publicly available on the Internet.

A user needs to authenticate to the system via some login (which may be a pseudonym) and the messages delivered unmodified between the user and the leader, like on MedHelp. But this system also requires that the user be authenticated as belonging to that group and that the chat not be publicly visible. To do this, an extra verification step is required. The assurance that the user has the disease, or is a caretaker or guardian for someone with the disease, can be provided technologically using attributed-based certificates (as discussed in section 3.4), which require a third party to provide the verification. This same verification would be required for the leaders to prove to the users that they have the credentials necessary to lead the session. If the leaders also must know, or in an emergency be able to determine, the identity of the patient, then the certificate signer must also maintain this information. It may be the contact information for the patient's primary caregiver rather than the patient's actual identity. The leader must also be able to contact the signing authority.

Keeping the discussions private requires configuring the software supporting the group interaction properly. However, it will not be able to control, or even detect, one of the users recording the information and then posting it to another location. Thus this requirement remains as future work if, indeed, it is even solvable.

Case study 2 also has the notion of the patient and doctor retreating to a "private" room to continue their conversation. This is similar to initiating a private chat between the two individuals—a feature that is widely supported in social media software. In this case, the communication must prevent eavesdropping, which requires the use of standard cryptographic mechanisms. A concern is the mental model that users might have of their surroundings. For example, the mental model of retreating to a private room for discussion might be more appropriate for the user than to "initiate a secure chat session". The secure chat can provide the underlying technology, while the manner in which this is presented is an important user interface consideration.

Governance issues abound. In addition to regulations that are specific to health data (e.g., HIPAA in the United States), which limit the recording of the conversations in the support room, more general laws conflict with the security requirements. For example, the use of cryptographic protocols is against the law in some countries, yet required to protect patient confidentiality. We note the existence of such conflicting policies; however, further discussion is outside the scope of this paper.

Case Study 3.

This case study is an extension of the group forum from the second case study. In particular, the group interactions are not publicly available, and the participants must have proven their appropriateness to participate in that particular group. The differences here, however, include additional conditions that the participants must meet in order to be included in the study (and that they will need to prove their suitability), and the necessity for greater "real world" coordination between the participants, their caregivers, the researchers, and government organizations (when, for example, the treatment to be tested requires approval within a particular country before its use).

The first difference could be solved using attribute-based certificates that provide the additional information, just as patients indicate that they have the condition that makes them appropriate for the support group. But perhaps it is best solved through greater coordination. The interactions have now moved from a more passive support group structure to a more active treatment, and so the patient's caregiver as well as the researcher need to be involved. This adds two more parties to the conversations. In this case, even if the patient's true identity is protected using pseudonymity, the caregiver, the leader, and the researcher will want to exchange actual credentials. Thus the pseudonym used by the caregiver would have a known mapping back to a true identity for the researcher, online leader, and the patient, but not necessarily for any of the other study participants. The advantage to this is that the researcher has the opportunity to prove his credentials and to work closely with the caregiver and leaders, as well as each party having the contact information of the other should there be some emergency.

Another requirement is the availability of private rooms for small group interactions. Again, the private room concept comes from the second case study, and the private nature of the conversations are protected through existing cryptographic protocols.

As noted above, participants being in different countries results in governance and legal issues, which we do not address in this paper. Nor do we address any issues that may arise from language and cultural differences.

6. FUTURE MECHANISMS

Although existing technology, policies, and procedures can address many of the problems raised by the therapeutic use of social networks, addressing the remaining issues requires that we advance in all three areas. Existing technologies support social networks within a societal context or contexts, and so the policies, procedures and technology must be tempered to meet the requirements within this context or these contexts. Technology applies only within the social network, and many of the requirements impose restrictions both internal and external to the social network. Meeting some requirements is simply beyond our current capabilities. Finally, meeting the requirements implies that there is agreement when a requirement is satisfied. Different people and cultures may disagree whether a particular policy, and its implementing procedures and technologies, actually satisfy requirements, or even that the risk is "acceptable".

We examine our case studies by looking at specific areas.

6.1 Protection of Privacy

A key consideration is the definition of "privacy," which we define as the inability to associate data with an individual external to the social network, unless the individual discloses the association. Suppose Anthony joins a network for the treatment of depression. He uses the pseudonym "Abel". He discloses this to the leaders, but not to anyone else. If no-one other than the leaders can determine that Abel is Anthony,

then his privacy has been protected. Otherwise it has not.

Such an association can be made in two ways. The first is by examining data on the network, looking for information that may reveal someone's identity. The second is by correlating data on the network with data external to the network. Here, we focus on the first.

Messages on the network will be associated with a pseudonymous sender. The goal is to be able to prove that the pseudonym cannot be tied to fewer than k identities [15]. For example, knowing the sender is one of 100 people may be acceptable in some circumstances, but not in others. Given the sensitive nature of the information on the network, preventing *false associations* (associating the pseudonym Abel (incorrectly) with Adrian, rather than (correctly) with Anthony) is of concern. Further, being able to prove that a pseudonym belongs to an identity beyond a reasonable doubt may be unacceptable, but being "pretty sure" of the association may be acceptable. The level of assurance needed depends upon the culture, the use to which the association will be put, and the law.

Under some circumstances, members may require privacy from the leaders. Perhaps they trust their caregiver, but do not feel comfortable giving the leaders their identity. Under these conditions, the caregiver can work with the leaders to provide the information necessary for the patient to join the network. Rather than divulging the name of the patient, though, the caregiver holds the association between the patient's identity and the pseudonym that the patient will use on the network. Then, in case of emergency, the leaders can contact the caregiver, who can take charge.

Laws and regulations control the management of PII—we will discuss *whose* laws and regulations below. For example, in the United States, the federal HIPAA regulations define 18 types of information in medical records that must be suppressed in some way when disclosing medical records. Other jurisdictions in the United States have stricter privacy laws when patients have specific conditions such as AIDS/HIV. A key question underlying all these laws and regulations is whether the messages are under the control of the patient or someone else. ORCON policies would need augmentation for critical situations. A patient who is incapacitated in some way may not be able to give consent for the leaders to provide her messages to her caregiver.

A second concern is that someone may correlate the information on the network with external information and, from that, determine the identity. The key question is what can be shared without enabling this. Traditional methods of data sanitization fail here because they assume a closed-world model, and this problem has an open world. An alternate view of sanitization is to ask what relationships must be known in order to reverse the sanitization [2, 5]; this is an open research problem.

Protecting the privacy of the patients is essentially the same in all our case studies. In particular, the pseudonyms must meet the above requirements, as must the messages. The latter is particularly difficult, as the sanitization must be done in real time and in ignorance of all possible external relationships. Particularly vexing is that a relationship enabling discovery of identity may not appear until *after* the message is sent to all members.

The third case study poses an interesting problem. The researcher will want to talk to the patients' therapists, to determine whether the patients are good subjects for the research. In order to make this assessment, the researcher must be convinced of the credentials of the leaders—interestingly, he need not know their identities. He will also need to know how to contact those leaders in case of an emergency, so some contact information must be provided.

6.2 Legal and Cultural

We return now to a problem identified earlier. Suppose the group spans multiple political or legal jurisdictions. What rules govern the network, its participants, and its data?

Suppose one government requires that the identities of patients with a particular disease be disclosed to its health authorities, so those authorities can prevent epidemics. A second jurisdiction requires that patient privacy be respected. The network spans both jurisdictions, with leaders and patients in both. The rules that control how this situation is handled are unclear.

Similar issues arise from clashes of cultures. The nuances of words for emotions in different languages mean that the emotional state of a group member could be easily misunderstood. Different customs may mean a patient considers unacceptable a treatment that a caregiver feels is proper. Those involved in such a multinational group must be prepared to handle such conflicts, even when they are completely unexpected. Similarly, members of the group must accept that others in the group may act in ways their societies would consider unacceptable, but the societies of others do accept. The leaders will play an important role in this.

These considerations affect all our case studies. Case study 3 show best the complexity of the issues involved. The researcher collects data as part of clinical assessments, so the protocols for the collection and handling must comply with the requirements of the legal jurisdictions involved, and be acceptable to the patients. As the patients are from different cultures and languages, mechanisms must be developed to ensure that all patients can give informed consent and understand the therapeutic trial. The trial protocols must take the differing time zones into consideration—a seemingly minor point, but one easily overlooked.

6.3 User Interface

Considerations of society, culture, language, and human factors all affect the design of the user interfaces. It begins with the ability to find the appropriate group in the glut of information on the Internet. Combine that with patients whose access to the Internet is severely limited, for example to library or school computers (which typically have an enforced time limit per user). Where no such group exists, but someone has a list of addresses of people he trusts, the user interface should set up such a secure group automatically. This minimizes the human effort—and with it, the chances that the human setting up the group will make a mistake that endangers the privacy and security of the group.[5]

A second consideration is the ability to generate, provide, and validate credentials easily. As noted in section 5, PKI is one possible solution to managing credentials. But PKI systems have proven effective only in specialized environments. In more general environments, they have been compromised when the issuers of credentials are not careful [11, 13]. In the therapeutic environment discussed in this paper, such an event would eliminate the trust critical to the success of

[5]Note this will be useful in contexts other than therapy, for example in education.

this approach. Also, users must understand not only how to validate a credential, but the importance of doing so. Hence, the validation mechanism must also be simple.

Finally, patients accessing the network will likely do so from many computers—perhaps they are traveling, or they use computers at school. Thus, the user interface must present an environment that the user expects regardless of where they access the network. In most cases, this simply means that the environment is established on a per-user basis, but some users may want different features on different machines (for example, if they use a Windows system at work and a Mac at home). A similar issue arises when they must break off during a session or an exchange of some kind. Should what they are working on migrate to where they next connect, or should it be lost? The answers to these questions require research, as do the mechanisms to make them clear to users who are not technologically sophisticated.

Case studies 2 and 3 have the idea of private rooms in which a set of people (patient and doctor in case study 2, and the researcher, leaders, and several patients in case study 3) can communicate securely and privately, without interfering with the rest of the network. This requires a simple interface to set them up, and to enter and leave.

7. CONCLUSION

People use social media to find and keep in touch with friends. Through social media they follow celebrities and events. They use the web to learn about medicine and medical conditions, and educate themselves. It seems only natural to use social media for therapeutic purposes such as mental or physical treatment. Such use raises many security questions because of the collaborative and social nature of the medium.

This paper described an online environment that can support group therapy and group interaction in a medical context. It presented the security requirements for such an environment, followed by three case studies to illustrate how such an environment might be used. Based on these, we described existing security mechanisms to support the effort, along with where additional research is required. There are substantial issues, particularly in the definition and protection of privacy, in supporting multiple and potentially conflicting government rules and regulations, and in providing user interfaces that are appropriate for the target audience.

Acknowledgements. Carrie Gates and Matt Bishop were supported by the GENI Projects Office project 1776 (in turn supported by NSF Grant No. CNS-0940805). Matt Bishop was also supported by NSF Grants No. CCF-0905503 and CNS-1049738. Any opinions, findings, and conclusions or recommendations expressed in this material do not necessarily reflect the views of BBN Technologies, the GENI Project Office, or the National Science Foundation.

8. REFERENCES

[1] R. J. Anderson. A security policy model for clinical information systems. In *Proceedings of the 1996 IEEE Symposium on Security and Privacy*, pages 30–43, May 1996.

[2] B. Bhumiratana and M. Bishop. Privacy aware data sharing: Balancing the usability and privacy of datasets. In *Proceedings of the 2nd International Conference on Pervsive Technologies Related to Assistive Environments*, pages 73:1–73:8, June 2009.

[3] M. Bishop, J. Cummins, S. Peisert, A. Singh, B. Bhumiratana, D. Agarwal, D. Frincke, and M. Hogarth. Relationships and data sanitization: A study in scarlet. In *Proceedings of the 2010 Workshop on New Security Paradigms*, pages 151–164, Sep. 2010.

[4] M. Bishop, C. Gates, and J. Hunker. The sisterhood of the traveling packets. In *Proceedings of the 2009 Workshop on New Security Paradigms*, pages 1–12, Sep 2009.

[5] R. Crawford, M. Bishop, B. Bhumiratana, L. Clark, and K. Levitt. Sanitization models and their limitations. In *Proceedings of the 2006 Workshop on New Security Paradigms*, pages 41–56, Sep. 2006.

[6] C. Gates and M. Bishop. The security and privacy implications of using social networks to deliver healthcare. In *Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments*, pages 29:1–29:6, June 2010.

[7] P. Golle. Revisiting the uniqueness of simple demographics in the US population. In *Proceedings of the 5th ACM Workshop on Privacy in Electronic Society*, pages 77–80, 2006.

[8] R. Graubert. On the need for a third form of access control. In *Proceedings of the 12th National Computer Security Conference*, pages 296–304, Oct. 1989.

[9] S. Kent. Privacy enhancement for internet electronic mail: Part ii: Certificate-based key management. RFC 1422, Feb. 1993.

[10] J. Park and R. Sandhu. Originator control in usage control. In *Proceedings of the Third International Workshop on Policies for Distributed Systems and Networks*, pages 60–66, June 2002.

[11] R. Richmond. An attack sheds light on internet security holes. *The New York Times*, page B3, Apr. 7, 2011.

[12] J. H. Saltzer and M. D. Schroeder. The protection of information in computer systems. *Proceedings of the IEEE*, 63(9):1278–1308, Sep. 1975.

[13] S. Sengupta. Hacker rattles security circles. *The New York Times*, page B1, Sep. 12, 2011.

[14] L. Sweeney. Uniqueness of Simple Demographics in the U.S. Population. Technical Report Data Privacy Working Paper 3, Laboratory for International Data Privacy, Carnegie Mellon University, Pittsburgh, PA, USA, 2000.

[15] L. Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5):557–570, 2002.

[16] A. Whitten and J. D. Tygar. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In *Proceedings of the 8th USENIX Security Symposium*, 1999.

[17] P. Yellowlees. *Your Health in the Information Age: How You and Your Doctor Can Use the Internet to Work Together*. iUniverse, Nov. 2008.

[18] P. Yellowlees and N. Nafiz. The psychiatrist-patient relationship of the future: Anytime, anywhere? *Harvard Review of Psychiatry*, 18(2):96–102, 2010.