SNS におけるユーザの投稿内容に応じた プライバシ保護に関する研究

町田　史門

博士（情報学）

総合研究大学院大学
複合科学研究科
情報学専攻

平成 28 年度
(2016 年)

2016 年 9 月

本論文は総合研究大学院大学複合科学研究科情報学専攻に
博士（情報学）授与の要件として提出した博士論文である.

審査委員：

主査	越前 功	教授	国立情報学研究所／総合研究大学院大学
	相原 健郎	准教授	国立情報学研究所／総合研究大学院大学
	小舘 亮之	教授	津田塾大学
	岡田 仁志	准教授	国立情報学研究所／総合研究大学院大学
	曽根原 登	教授	国立情報学研究所／総合研究大学院大学

（主査以外はアルファベット順）

Privacy Protection Corresponding to User's Contents in Social Networking Services

Shimon Machida

DOCTOR OF
PHILOSOPHY

Department of Informatics,
School of Multidisciplinary Sciences,
SOKENDAI (The Graduate University for Advanced Studies)

September 2016

A dissertation submitted to
the Department of Informatics,
School of Multidisciplinary Sciences,
SOKENDAI (The Graduate University for Advanced Studies)
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Advisory Committee

Prof. Isao Echizen (Chair)	National Institute of Informatics/ The Graduate University for Advanced Studies
Assoc. Prof. Kenro Aihara	National Institute of Informatics/ The Graduate University for Advanced Studies
Prof. Akihisa Kodate	Tsuda College
Assoc. Prof. Hitoshi Okada	National Institute of Informatics/ The Graduate University for Advanced Studies
Prof. Noboru Sonehara	National Institute of Informatics/ The Graduate University for Advanced Studies

(Alphabet order of last name except chair)

要旨

Twitter や Facebook に代表される Social Networking Service (SNS) は，世代を問わず人々の日常生活に溶け込み，オンラインコミュニケーションツールとして欠かせないものとなっている．その一方，SNS ユーザの不用意な投稿により，個人に関するセンシティブデータの漏洩が問題となっている．投稿の結果，自身や他人のセンシティブデータが友人関係性の弱い知人レベルの友人や面識のない他人にまで拡散し，友人関係のもつれや職を失うユーザが出るなど，何気ない投稿が予期しないトラブルへと発展している．このような漏洩の原因の 1 つに SNS ユーザが持つ投稿に対する主観的な判断基準がある．各ユーザの判断基準は，個人の主観に依存しているため，その投稿判断を誤った場合に，センシティブデータの漏洩が発生する可能性がある．そのため，SNS ユーザにとって，どのような情報がセンシティブデータとなりえるのかを明確にし，統一的な判断基準のもとに投稿判断が行えることが理想である．また，SNS 提供者側では，ユーザのセンシティブデータ漏洩を防ぐために，投稿に対する公開ポリシや公開対象者を選択する公開範囲の設定機能などのプライバシポリシ管理機能を提供している．しかし，SNS におけるプライバシポリシ管理は複雑性があるだけでなく，その維持にも多くの労力が必要とされるため，簡易に対処する手段が必要である．

本論文では，センシティブデータの漏洩を無理に防止するのではなく，SNS の特徴である他ユーザとのオンラインコミュニケーションの楽しさを極力維持しながら，SNS 投稿ユーザの主観的な判断基準のみに依存せずに，客観的なセンシティブデータの判断基準，および投稿内容に含まれる他ユーザのプライバシポリシに基づいて，投稿メッセージを適応的に保護する手法の確立を目的とする．具体的には，SNS 投稿時に，テキスト，写真を含む投稿予定メッセージにセンシティブデータを含むか否かを検知し，その検知結果により，限定的な公開範囲・公開対象者を自動提案する．さらに，投稿される写真に写る/写り込む被写体のコミュニティや状況に応じたプライバシポリシに基づき，当該人物の顔領域を適応的に保護する．

本目的に向けて，1 つめの課題として，「SNS におけるセンシティブデータの客観的な判断基準の定義」に取り組む．SNS ユーザの主観的な判断基準に依存した SNS 投稿により発生するセンシティブデータ漏洩を防止するために，ユーザにとって，どのような情報がセンシティブデータとなりえるのかを明確化し，客観的な判断基準を定義する．本課題では，最初に従来研究で言及されるセンシティブデータの漏洩が実際に発生していることを確認するため，過去 1 年間分の Twitter アーカイブを用いて，センシティブデータ漏洩の発生有無を試算評価するとともに，投稿予定メッセージからセンシティブデータを検知するための手法を検討した．次に，この評価結果を受け，SNS 投稿前の投稿予定メッセージに含まれるセンシティブデータの自動検知を実現するために，当該メッセージにセンシティブデータを含むか否かの客観的な判断基準の定義の第一歩として，公文書におけるプライバシの取り組みを参照し，

非公開とすべき情報の内容分類と公開範囲を表す開示レベルを対応付けた,“SNS におけるプライバシ侵害情報分類表”を提案した．さらに，本分類表のシステム適用・実装に向けて，分類表の妥当性を評価調査により示した．本分類表をシステムに適用することにより，SNS ユーザは, 投稿を行う前に客観的な判断基準のもと, 投稿判断を可能とする第一歩を示した.

次に 2 つめの課題として,「センシティブデータの漏洩検知に基づき投稿ユーザへ指摘・通知」に取り組む．投稿ユーザが投稿メッセージにセンシティブデータを含んでいることに気付かずに投稿してしまい，その結果，ユーザがその投稿を後悔している課題に対して，1 つ目の課題で提案した分類表が持つ，SNS 投稿メッセージに含まれるセンシティブデータの有無とその開示レベルを客観的に判断可能とする特性を活かし，センシティブデータの漏洩有無の検知・通知と，その情報の重要度に沿った公開範囲の自動設定を提供する“センシティブデータの漏洩検知に基づく公開範囲の設定方式”を提案した．また，その実現化として，アプリケーション：Adaptive Disclosure Controller for Facebook を実装し，検知したセンシティブデータの重要度に応じた開示レベルを投稿ユーザへ視覚的に示し，センシティブデータが含まれていることを認識できるよう通知した．さらに， SNS ユーザが持つ，投稿内容に応じた特定の個人や興味を持つコミュニティへのメッセージ公開要望に対して，容易に任意の公開範囲・公開対象者を定義可能とすることで，特定コミュニティ・グループにとらわれることなく，投稿メッセージを公開できることを示した.

最後に 3 つめの課題として,「コミュニティに応じた被写体のプライバシポリシの反映」に取り組む．投稿ユーザは写真に写る・写り込む被写体の公開・非公開などのプライバシポリシが分からないため，自身の主観的な判断基準に依存した SNS 投稿を行わざるをえない課題に対し，投稿ユーザに被写体のコミュニティや状況に応じたプライバシポリシを参照させる手法として,“被写体のコミュニティベース・プライバシポリシの設定方式”を提案した．本方式では，被写体が属するコミュニティ内外におけるプライバシの振る舞いをポリシとして埋め込んだタグ：PrivacyTag を用いて，コミュニティ内外で当該人物の顔領域を適応的に保護する．また，本方式の実現化として，被写体から PrivacyTag の検知・解析を可能としたアプリケーション：Photo Privacy Realizer for Facebook の実装を行った．これらにより，SNS 投稿ユーザの主観的な判断基準のみに依存せずに，被写体のコミュニティや状況に応じたプライバシポリシを反映させることを可能とした.

本論文では，SNS ユーザの主観的な判断基準に依存した投稿により発生するセンシティブデータ漏洩を防止するために，客観的な判断基準の定義の第一歩として，SNS におけるセンシティブデータの分類を提案・評価した．そして，本分類表をベースに，投稿前のユーザへ視覚的に指摘・通知をする設定方式を提案・実装した．また，軽視されやすい写真に写る被写体のプライバシ保護のために，被写体のポリシを撮影者・投稿ユーザに参照させる方式を提案・実装した．本論文で提案したこれらの方式は，現在問題となっている SNS ユーザのセンシティブデータの漏洩に対して，その防止に寄与するものと考える.

Abstract

Social networking services (SNSs), such as Facebook and Twitter, have become popular among people of all ages, and online communication with friends and acquaitances via messages that include photos and videos has become very common. While it has become very easy for users to post messages, inadvantent disclosure of sensitive information through unintentional posting has become a problem. A message containing sensitive information can be passed along by userself or friends to acquaintances and strangers. Disclosure of such information can trigger unexpected problems such as loss of credibility, loss of one's job, and problems at school. A major factor in the uninitended disclosure of sensitive information is that SNS users subjectively judge whether information is sensitive or not. The judgement criteria differ among users, so a user may judge information that the posting user considers to be sensitive as not sensitive and pass it along. Ideally, SNS users would make judgements on the basis of common criteria, but this requires a clear definition of what information is sensitive for users. While SNS prividers have implemented privacy management mesures such as publishing a privacy policy and providing a function for setting message access control, managing privacy is complex and requires much effort. An easier way to protect the privacy of SNS users is thus needed.

In this paper, I present a method for preventing the leakage of sensitive information due to differences in judging the sensitivity of information. It is based on objective judgement criteria and a privacy policy that is contained posted message and therefore does not rely on the subjective judgement criteria of SNS users. In particular, when a message is posted to an SNS, it detects whether the message contains sensitive information. If it does, the method proposes a limitation of the scope and targets for disclosure. I also present a method for adaptively protecting the privacy of people appearing in photos for disclosure within and outside the communities to which they belong. It uses tags embedded with community-based privacy policies to cover the face of people.

My first challenge was defining criteria for objectively judging whether information in a message posted on an SNS is sensitive for use in preventing the leakage of sensitive information due to differences in judging the sensitivity of information. I defined objective judgement criteria by clarifying what sensitive information is for users. To establish a baseline for the actual leakage of sensitive information on an SNS, I investigated the leakage of sensitive information using the Japanese Twitter archives and developed a method for efficiently detecting the leakage of sensitive information in a message. Next, using the results of this investigation, I defined a classification table for use in automatically identifying information that should be kept private and objectively assigning a level for the scope of disclosure. This table was validated by user evaluation.

The next challenge was developing a method for notifying SNS users of privacy leaks on the basis of

detection of sensitive information. If a user does not detect that a message to be posted contains sensitive information, the user may experience regret after posting it. The method I developed helps prevent such regret by using the classification table to detect sensitive information in messages to be posted. It was implemented in a Facebook application that adaptively controls information disclosure. The disclosure level is visualized on the basis of Dunbar's circle as the weight of the detected sensitive information, enabling the user to recognize that the message contains sensitive information. This application also enables SNS users to efficiently and effectively publish a message to specific friends or a specific community in accordance with the contents. The application can be easily defined to "think outside the box" and target a specific community or group such as friends and friends and friends.

The third and final challenge was developing a method for adaptively protecting the privacy of people appearing in photos to be posted within and outside the communities to which they belong. SNS posters are forced to post messages on the basis of their own subjective judgement criteria because they do not know whether the people in a photo agree to having their photograph published. The method I developing for avoiding this situation is community-based. It adaptively protects the face area of persons appearing in photos to be posted within and outside the communities to which they belong by using tags embedded with community-based privacy policies. It was implemented in a Facebook application that enables protection of privacy on the basis of the privacy policies of the people in a photograph and thefore does not only rely on the subjective judgement criteria of the poster or the photographer.

This research is aimed at avoiding privacy leaks due to SNS users relying on subjective judgement criteria when posting a message or photograph. The table developed for classifying sensitive information in SNS messages is a first step towards defineing objective judgement criteria. The method developed for notifying SNS users of pending privacy leaks will help reduce instances of post-posting regret. Finally, the method developed for taking the privacy policies of the people in photograph into account when posting it will help protect their privacy. This research is a first step toward efficiently and effectively protecting the privacy of SNS users.

目次

図目次

表目次

第1章　序論

1.1本研究の背景

1.1.1 不用意な投稿による個人に関する情報の漏洩

Twitter や Facebook に代表される Social Networking Service (SNS)が幅広い年代に普及し[1], 友人・知人と写真や動画を含むメッセージを介したオンラインコミュニケーションが日常化している. これら SNS において 1 日あたり, Twitter は 3.4 億ツイートのメッセージ[2], Facebook は 3 億 5000 万枚[3], Instagram には 7000 万枚[4]もの写真が投稿されており，SNS ユーザは日常の出来事をスマートフォン等のカメラ付きデバイスを用いて手軽に撮影し，メッセージを添えて SNS へ投稿している．手軽に撮影・投稿できる一方，ユーザの不用意な投稿による個人に関する情報の漏洩が問題となっており，その結果，友人関係のもつれや職を失う SNS ユーザが出る等[5]，何気ない投稿が予期しないトラブルへと発展している．SNS での投稿はインターネット上に瞬時に流れ，友人関係性の弱い知人レベルの友人や面識のない他人を含めた広範囲に拡散していくため，長期間残り続ける可能性が高い．投稿に含まれる現在の所在や状況が推測されやすい情報が漏洩した場合, 犯罪につながる恐れも指摘されているため[6], 投稿内容には細心の注意を払う必要がある.

この原因の 1 つとして，自分自身により，個人に関する情報を漏洩させていることがあげられる．Facebook ユーザを対象とした調査[7]では，81%の調査対象者（2,383 名）が家族の名前や生年月日，写真などの家族に関連した情報の投稿経験があった．また，日々このように家族の情報が投稿される Facebook を“modern day baby book”[8]と評し，子や孫を持つユーザは，成長軌跡として子供の可愛い写真を家族・友人等のコミュニティへ共有したい願望，良い母・幸せな家族を演じたい願望を持つとされる．このような投稿の結果，自身の情報のみならず，写真の写り込みを含む，他ユーザの情報も漏洩する可能性がある．しかし，ユーザは，自身が写った写真の許可を得ていない無断投稿やタグ付けをされた場合，このことを不快に感じている[9]にも関わらず，情報セキュリティの倫理に対する意識調査[10]では，70%の調査対象者（8,500 名）が他人の写った写真を SNS に公開することに対し，問題意識がない傾向にあった．このような状況を受け，SNS 投稿時の不要なトラブルを避けるための注意喚起[11]として，投稿前に (1) 位置情報等のメタ情報の削除，(2) 写真の被写体から事前に投稿許可の取得，(3) 不要な写り込みに対して特定ができないよう加工 が推奨されている．しかし，投稿の都度，適切な対応を施すことは困難であるため，簡易に対処する手段が必要である.

現在，各 SNS では，ユーザのプライバシを保護するための機能として，投稿に対する公開

ポリシや公開対象者を選択する公開範囲の設定機能など，ユーザのプライバシポリシ管理機能を提供している[12]．しかし，SNS ユーザは度々これら管理設定を間違える傾向にある[13, 14, 15]．また，SNS におけるプライバシポリシ管理は複雑性があるだけでなく，その維持にも多くの労力が必要[16]であるため，SNS ユーザの中には，(1) 面識のないユーザが参照可能であるデフォルト設定のまま使用し続ける，(2) 過去に友人等の特定の人のみに公開できるよう設定変更をしていたが，SNS 側の機能拡張・改変時にデフォルト設定へ戻ってしまったことに気付かず，自身の投稿を面識のないユーザが参照可能な状態で使用し続けるなどの問題[17]が発生している．このようにユーザの IT リテラシの低さから，意図していない他ユーザへ情報が漏洩する場合もあるため，現在提供されているこれらのプライバシポリシ管理機能のみでは十分とは言えない状況である．

1.1.2 SNS ユーザのプライバシ理解

総務省主催 パーソナルデータの利用・流通に関する研究会のプライバシ保護等に配慮したパーソナルデータ（個人に関する情報）のネットワーク上での利用・流通の促進に向けた方策について検討の報告書[18]では，個人に関する情報を“パーソナルデータ”と定義し，図 1-1 に示すように，(1) 一般パーソナルデータ，(2) 慎重な取扱いが求められるパーソナルデータ，(3) センシティブデータの 3 種類に分類している．

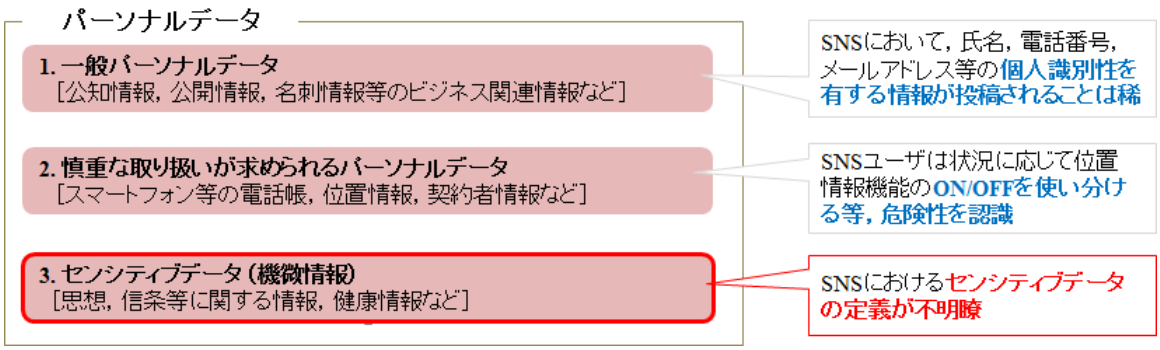


図 1-1　パーソナルデータの 3 分類と SNS ユーザの理解

“一般パーソナルデータ”とは，氏名等の公知情報，本人の明確な意図で一般に公開された情報，名刺に記載されている情報等のビジネス関連情報を指す．“慎重な取扱いが求められるパーソナルデータ”とは，GPS 等の位置情報や継続的に収集される購買・貸出履歴などを指す．“センシティブデータ”とは，思想・信条などに関する情報，健康情報等のさらに慎重に扱われるべき情報を指す．次に，これら 3 種類のパーソナルデータの SNS における扱われ方であるが，一般パーソナルデータは，氏名等保護される情報が定まっているため，パーソナルデータに該当するか否かの判断がつきやすく，SNS においてメールアドレス等の個人識別性を有する単純な情報が投稿されることは稀である[19]．また，慎重な取扱いが求めら

れるパーソナルデータの中でも, SNS 投稿に付帯する情報として知られる GPS 等位置情報は, Foursquare や Google Maps 等の位置情報サービスで主に利用されているが, ユーザは状況に応じて位置情報機能の ON/OFF を使い分ける等，その危険性を認識している[20]．しかし，最後のセンシティブデータにおいては，保護すべき情報の境界定義が曖昧であることから，パーソナルデータであるかの判断が難しくなる．このセンシティブデータは，2015 年 9 月に成立した個人情報保護法改正[21]において，“要配慮個人情報”*1として定義され，その対象となる情報については，「本人の人種，信条，社会的身分，病歴，犯罪の経歴，犯罪により害を被った事実その他本人に対する不当な差別，偏見その他の不利益が生じないようにその取扱いに特に配慮を要するもの」[22]とし，配慮をすべき具体的な個別情報については，今後政令で定義される予定である．しかし，これまで実社会において，個人情報保護に対する誤解や理解不足が指摘[23]されてきたように，SNS ユーザが SNS へメッセージを投稿する際に，これらのセンシティブデータが含まれていることに気付きにくい．実際にセンシティブデータを漏洩してしまった SNS ユーザは，投稿後に自ら漏洩に気付く，または，友人や他ユーザからの指摘によって漏洩を認識し, その不用意な投稿行為に後悔している[24]. このことから, 本論文では, SNS で漏洩している個人に関する情報とは，主にセンシティブデータであると仮定する．次項では，このようなセンシティブデータ漏洩が発生する理由の 1 つである SNS ユーザごとに異なるプライバシポリシについて述べる.

1.1.3 個人やコミュニティ・状況により異なるプライバシポリシ

実生活において，人々は属する地域や家庭における文化・その背景等から，プライバシの理解に多様性があるため，それぞれ異なるプライバシポリシを持つ．また，同一個人においても，自身が属するコミュニティや状況に応じた異なるポリシを持っている[25]．そのため，全ての人々にプライバシに関する共通認識を持たせることがきわめて難しい.

同様に SNS においても，SNS ユーザは投稿メッセージに対して，アクセス可/不可とする友人をサブセット化したリスト（Social Access Control List）を用いて，自身が属するコミュニティを複数定義し，投稿内容や状況に応じて，これらコミュニティを使い分けている[16]ように，SNS ユーザもコミュニティ等の状況に応じた異なるプライバシポリシを持っている．しかし，不用意な投稿により，自身が意図しないコミュニティに投稿が公開された場合，投稿内容に含まれる，誰が何をしているかといった行動内容や写真を他人に知られることを不快に感じている[8, 26]．一方，自身が他ユーザの投稿メッセージから他ユーザの行動を知ることは不快に感じず[27]，むしろ，その内容を詳細に知りたいと考えている[28]．このような背景もあり，SNS ユーザは投稿内容に応じて，特定の個人や興味を持つコミュニティへメッセージを公開することで，自身の投稿を安全に効果的に共有したい願望を持っている[29, 30].

*1 本論文では，“要配慮個人情報”ではなく，“センシティブデータ”の表現を使用する.

1.2 本研究の目的と課題

本研究では，これまで述べてきたようなセンシティブデータ漏洩を無理に防止するのではなく，SNS の特徴である他ユーザとのオンラインコミュニケーションの楽しさを極力維持しながら，SNS 投稿ユーザの主観的な判断基準のみに依存せずに，客観的なセンシティブデータの判断基準，および投稿内容に含まれる他ユーザのプライバシポリシに基づいて，投稿メッセージを適応的に保護する手法の確立を目的とする．具体的には，SNS 投稿時に，テキスト，写真を含む投稿予定メッセージにセンシティブデータを含むか否かを検知し，その検知結果により，限定的な公開範囲・公開対象者を自動提案する．さらに，投稿される写真に写る/写り込む被写体のコミュニティや状況に応じたプライバシポリシに基づき，当該人物の顔領域を適応的に保護する．本目的に向けて，以下の課題に取り組む．

1. SNS におけるセンシティブデータの客観的な判断基準の定義

2. センシティブデータの漏洩検知に基づき投稿ユーザへ指摘・通知

3. コミュニティに応じた被写体のプライバシポリシの反映

「1. SNS におけるセンシティブデータの客観的な判断基準の定義」では，SNS ユーザの主観的な判断基準に依存した SNS 投稿により発生するセンシティブデータ漏洩を防止するために, SNS ユーザにとって, どのような情報がセンシティブデータとなりえるのかを明確化し，客観的な判断基準を定義する．これにより，誰もが統一的な判断基準に基づき，投稿判断が行えることを目指す．

「2. センシティブデータの漏洩検知に基づき投稿ユーザへ指摘・通知」では，投稿ユーザがメッセージにセンシティブデータを含んでいることに気付かずに投稿してしまい，その結果，後悔している課題に対して，システム側で投稿予定メッセージからセンシティブデータの漏洩を自動検知し，その結果に基づき，当該投稿の公開範囲・公開対象者を自動提案する手法を検討する．

「3. コミュニティに応じた被写体のプライバシポリシの反映」では，投稿ユーザは写真に写る・写り込む被写体の公開・非公開などのプライバシポリシが分からないために，自身の主観的な判断基準に依存した SNS 投稿を行っている課題に対し，写真の被写体のプライバシ保護を目的として，投稿ユーザに被写体のコミュニティや状況に応じたプライバシポリシを参照させる手法を検討する．

1.3 本論文の構成

本論文は，全 6 章から構成される．序論では，本研究の背景，目的，本目的に向けた課題

について述べた.

第 2 章「SNS ユーザのプライバシ保護」では，SNS ユーザのプライバシを保護するために SNS 側から提供されるプライバシ管理機能の 1 つである，投稿メッセージの公開範囲設定機能とその課題について述べる．次に，その課題により引き起こされる SNS ユーザの投稿後の後悔について説明する．そして，このようなユーザの後悔を防ぐために従来研究で提案されている SNS ユーザのプライバシ保護手法として，SNS 投稿ユーザを保護主体とした手法，写真の被写体を保護主体とした手法について述べながら，これら従来手法の課題と本研究の位置付けを述べる.

第 3 章「SNS におけるセンシティブデータの分類」では，最初に，従来研究で言及されるセンシティブデータの漏洩が実際に SNS で発生していることを確認するため，過去 1 年間分の 1 億 4,000 万件の Twitter アーカイブを用いて，センシティブデータ漏洩の発生有無を試算評価するとともに，投稿予定メッセージからセンシティブデータを検知するための手法を検討する．次に，この評価結果を受け，SNS 投稿前の投稿予定メッセージに含まれるセンシティブデータの自動検知を実現するために，当該メッセージにセンシティブデータを含むか否かの客観的な判断基準の定義の第一歩として，非公開とすべき情報の内容分類と公開範囲を表す開示レベルを対応付けた，“SNS におけるプライバシ侵害情報分類表”を提案する．さらに，本分類表のシステム適用・実装に向けて，分類表の妥当性を評価調査により示す.

第 4 章「センシティブデータの漏洩検知に基づく公開範囲の設定方式」では，3 章で提案するプライバシ侵害情報分類表が持つ，SNS 投稿メッセージの内容からセンシティブデータ漏洩の有無と開示レベルの客観的な判断を可能とする特性を活かし，SNS 投稿予定メッセージからセンシティブデータの漏洩検知とその公開範囲・公開対象者の自動設定を提供する手法を提案する．そして，本提案手法の実現化として，Facebook を対象 SNS とした“Adaptive Disclosure Controller for Facebook”の実装を示す.

第 5 章「被写体のコミュニティベース・プライバシポリシの設定方式」では，3 章で提案する SNS におけるプライバシ侵害情報分類表において，投稿の公開範囲を表現した開示レベル 2 に該当する情報：“個人特定可能な写真”に焦点を当て，写真に写る/写り込む人物のプライバシに注目する．写真の被写体のプライバシを保護するために，当該人物が属するコミュニティ内外におけるプライバシの振る舞いをポリシとして埋め込んだタグ“PrivacyTag”を用いて，コミュニティ内外で当該人物の顔領域を適応的に保護する手法を提案する．また，この PrivacyTag を定義するために，まず予備評価を実施し，その評価結果を基にタグのデザイン検討を行う．そして既存手法で用いられる QR コードベースタグと提案タグの評価実験を行う．さらに，本提案手法の実現化として，Facebook を対象 SNS とした“Photo Privacy Realizer for Facebook”の実装を示す.

第 6 章「結論」では，本研究全体を総括し，SNS ユーザのプライバシ保護における本研究の成果と今後の課題を示す.

第2章　　SNS ユーザのプライバシ保護

本章では，最初に SNS から提供されるプライバシ管理機能として，投稿の公開範囲設定機能とその課題について述べ，その結果引き起こされる SNS ユーザの投稿後の後悔について説明する．次に，従来研究で提案されている SNS ユーザのプライバシ保護手法として，SNS 投稿ユーザを保護主体とした手法，写真の被写体を保護主体とした手法について述べながら，これら従来手法の課題と本研究の位置付けを述べる．

2.1　プライバシ管理機能：公開範囲設定

SNS では，ユーザのセンシティブデータの漏洩を防ぐために，様々なプライバシ管理機能を提供している．そのうちの 1 つに投稿の公開範囲設定がある．本設定は，ユーザが行うメッセージや画像の投稿，メッセージの共有などの行動に対して，どの範囲まで情報を共有するかを公開範囲として設定する．2016 年 6 月現在，Facebook では投稿時の公開範囲の設定として，“公開”，“友だち”，“自分のみ”，“SNS 側で作成される特定グループ”，“ユーザによりカスタマイズ可能なグループ”を提供している．Twitter では，“公開”・“個別ユーザに対する”ダイレクトメッセージ機能を提供している．そして，これらの選択肢から，多くの SNS ユーザは公開範囲：“友だち”を設定している[17]とされる．しかし，公開範囲：“友だち”の中には，関係性の弱いユーザが含まれており，公開範囲を狭めたとしても，依然としてプライバシ漏洩の危険性がある[12]．また，Facebook における機能変更過程において，ユーザが知らぬ間にデフォルト設定である公開範囲：“公開”に変わっていることがあり，その結果センシティブデータが漏洩することがある[17]．次に，“ユーザによりカスタマイズ可能なグループ”では，投稿に対するアクセス可/不可とする友人をサブセット化したリスト（SACL: Social Access Control List）を用いて，ユーザ定義の公開範囲リストを複数定義することができる．

図 2-1　Facebook における SACL の作成 [16]

本リストは，事前定義，または投稿の都度定義が可能であり，ユーザの中には投稿内容によって，公開対象者として望ましいユーザの選択や望ましくないユーザの除外を行っている[31]が，IT リテラシの低いユーザも含め，全てのユーザが投稿の都度設定を行っていくことは困難である．また，通常これらのリストは手動作成され，作成後の維持管理に多くの労力が必要となる．（図 2-1 を参照）

2.2　投稿後の SNS ユーザの後悔

前節で説明した投稿の公開範囲設定は，SNS ユーザのプライバシ保護機能として有効であるものの，設定の複雑さや維持管理の難しさなどもあり，現状は十分とは言えない状況である．その結果，SNS ユーザは不用意な投稿により，その投稿後に後悔している．従来研究では，SNS ユーザの投稿後の後悔理由などから，SNS におけるユーザの行動傾向，保護すべきユーザ属性情報，投稿すべきでない情報の分類を得ようとしている．Twitter ユーザへのユーザ評価[32]では，自身が意図していた公開範囲よりも広範囲に投稿が共有された時に後悔する傾向にあるとし，後悔する可能性がある投稿は，その投稿を行う前にユーザへ気づきを与えるべきとした．そして，メッセージ投稿前の投稿ユーザへ気づきを与えるため，ユーザが投稿後に削除した Tweet アーカイブからユーザの後悔理由を分析すると共に，その Tweet アーカイブをもとに教師データを生成し，投稿予定メッセージに後悔する可能性が高い内容が含まれていることを自動検知する手法を提案している[33, 34]．また，削除された Tweet には，感情が高まった時に使用されやすい罵倒語が多く含まれる傾向にあることが分かっている[34, 35]．さらに，このようなユーザの感情が高まった場合に投稿する可能性があるケースとして，休暇中の投稿，飲酒中の投稿，病気に関連した投稿に注目し，Twitter におけるセンシティブデータ漏洩の抽出とその要因分析がある[36]．そして将来的には，この分析結果を検知システムに活かし，ユーザの投稿前に再考させるべきとした．同様に，Facebook を対象としたユーザ評価[24]では，ユーザの後悔は主に，(1) アルコールや性，宗教・政治などのセンシティブデータの投稿，(2) 感情的になった投稿，(3) 意図していないユーザへの投稿共有 に集約されるとし，これらの投稿が行われる前に，ユーザに気づきを与えることが重要であるとした．

このように，従来手法では SNS 投稿後のユーザの後悔理由の分析により，SNS ユーザが後悔する投稿内容の傾向を把握すると共に，投稿前にユーザへ気づきを与える指摘機能が必要として，投稿前の自動検知を試みている．しかし，投稿前にメッセージから自動検知するためには，SNS ユーザにとってどのような情報がセンシティブデータとなり得るのかを明確に定義する必要があり，従来手法では十分に検討がされていない状況である．

2.3　SNS 投稿ユーザを保護主体とした手法

本節では，このような SNS ユーザの投稿後の後悔を防ぐために，従来手法で提案されている投稿ユーザを保護主体としたプライバシ強化手法を述べる．

Palen らは[37]，IT 時代におけるプライバシ管理の主要事項として，Altman の Privacy Regulation Theory [38]をベースに，(1) 情報の開示境界，(2) 情報の識別境界，(3) 情報の時間境界の 3 境界を提案している．“情報の開示境界”とは，その情報が一般に公開可能事項であるかの境界を指し，“情報の識別境界”とは，その情報における特定の個人や事象の識別性の境界を指す．“情報の時間境界”とは，過去，現在そして未来等の時間経過における情報の境界を指す．以降では，これら 3 つの情報の境界をもとに従来手法について論じる．

2.3.1 情報の開示境界

投稿の公開範囲

SNS ユーザのプライバシ保護の従来研究として，最も活発に提案されているのは，情報の開示境界に関する手法である．SNS では，先に述べたように，投稿の公開範囲をコントロールする機能として，SACL: Social Access Control List を提供している．作成した SACL 数や友人数が増えるにつれ，実際に友人がどのリストに含まれているのかを認識しづらくなり，結果的に意図していない他ユーザへセンシティブデータが漏洩してしまう課題がある．この課題に対して，Circle によって複数の集合の関係を図式化するベン図をベースとした公開範囲の設定用インターフェースが提案されている[39]．本手法では，友人リスト内に存在する複数のグループやコミュニティにまたがった公開範囲の設定が可能となる（図 2-2 (a)を参照）．また，ユーザの現在のプライバシ設定の確認を目的として，友人リスト内に存在するサブグループとそのグループに所属するユーザを Circle の包含関係によって可視化し，プライバシ設定の見直しを促す手法が提案されている[40] （図 2-2 (b)を参照）．

次に，通常 SNS ユーザはこの SACL を手動で作成していくが，ユーザにとってこの作業は非常に労力の掛かる作業となる．そこで，この労力の掛かる SACL 作成を，ユーザ自身や友人などの他ユーザが持つ属性情報からコミュニティを自動検知し，SACL を自動作成する手法[16] （図 2-3 を参照）や，似た趣向を持つユーザは同一グループである可能性が高いとし，新たなユーザとの関係が確立した際に，その当該ユーザの所属グループを自動提案する手法[41]がある．これら提案手法の共通した結論としては，通常 SNS ユーザが意識し難い投稿の公開範囲を可視化することによって，ユーザに気づきを与えることの重要性である．

(a) ベン図による公開範囲の可視化　　(b) 友人関係性の可視化

図 2-2　公開範囲の可視化

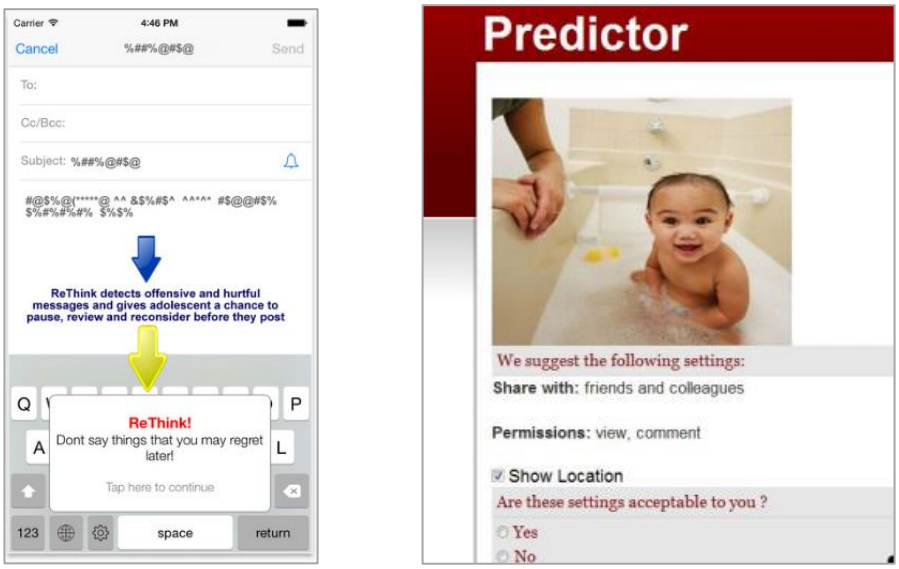
図 2-3　コミュニティベースの友達リストの自動提案

友人の可視化

次に，SACL により投稿の公開範囲を絞り込んだ後，SACL に含まれる個別の友人を識別する手法を述べる．Privacy Nudges for Social Media[42] （図 2-4 を参照）では，当該投稿が誰に対して公開予定であるのかを，対象ユーザのプロフィール画像とともに表示することで，ユーザに注意喚起を促し，センシティブデータの漏洩を軽減できることをユーザ評価により示した．しかし，この手法では，投稿の公開範囲を可視化し，ユーザに注意を促すことはできるものの，公開対象者を容易に選択できないため，SNS ユーザが望むコミュニティを横断した特定の個人や興味を持つコミュニティへの投稿ができない課題が残る．また，Ego-net digger[43] （図 2-5 を参照）では，友人の紐付き強度を手動評価するために，評価軸上に友人画像を配置・可視化し，その関係性を簡易に計れるようにした．しかし，事前に全友人の手動評価が必要になること，新たに友人が友達リストに加えられるたびに評価が必要となるため，維持管理が困難である．

図 2-4　公開対象者の可視化

図 2-5　友人の紐付き強度の可視化



(a) 検知に基づく再考提案　　(b) プライバシポリシの動的提案

図 2-6　投稿内容に応じた提案

投稿内容に応じたユーザへ指摘・提案

上述までの従来研究を通じて，センシティブデータを含んだメッセージが投稿されることを防止するために，SNS ユーザが投稿を行う前に，当該ユーザへ投稿内容に応じた指摘・通知することの重要性を述べた．ここではその具体的な実現手法として提案されている従来研究を述べる．ReThink[44] （図 2-6-(a)を参照）は，SNS 上で発生するいじめを防止するために提案された手法であるが，投稿メッセージに誹謗中傷に関連したワードが含まれる場合，投稿直前にポップアップを表示し，投稿ユーザに再考を要求する非常にシンプルな手法を提案した．ユーザ評価の結果，ユーザによる侮辱的なメッセージ投稿を大幅に減少させたように，投稿前にユーザに気づきを与えることの有効性を示した．その他にも，ユーザが投稿するテキスト情報から投稿内容のトピックを示すラベルを推定し，そのラベル毎に公開対象グループを提案する手法[13]，画像を含む投稿予定メッセージの内容に応じて，プライバシポリシを動的に生成し，公開範囲などをユーザへ提案することで，不用意なセンシティブデータ漏洩を防ぐ手法が提案されている[45] （図 2-6-(b)を参照）．しかし，これらの手法による公開範囲の提案は，投稿ユーザに気づきを与えることの重要性は示しているものの，SNS 側が提供する友だちや特定グループを対象としており，依然として公開対象に関係性の弱い友人が含まれ，かつ，公開対象が特定グループのみに制限されるという課題がある．

2.3.2 情報の識別境界

投稿内容に応じた識別境界

次に，情報における特定の個人や事象の識別性の境界を基にした従来手法から，投稿内容に応じた識別境界を論じる．本境界では，投稿メッセージに含まれるセンシティブデータを検知・匿名化することにより，センシティブデータの漏洩を防ぐ．具体的には，家族・友人などの公開対象者との関係性から，投稿メッセージに含まれるセンシティブデータを適応的な匿名化処理を行う手法[46]（図 2-7 参照）や，メッセージ内に含まれる時間に関連したフレーズをルールベースにより暈す手法[47]などが提案されている．本論文では，SNS の特徴である他ユーザとのオンラインコミュニケーションの楽しさを極力維持しながら，SNS ユーザのプライバシ保護することを目的とするため，投稿内容の匿名化によるプライバシ保護手法は未検討である．

2.3.3 情報の時間境界

投稿の時間経過に応じた境界

最後に，過去，現在そして未来などの時間経過における情報の境界として，投稿の時間経過に応じた境界を論じる．SNS に投稿されたメッセージはインターネット上に長期間残り続ける懸念があるが，Facebook ユーザを対象としたユーザ調査[48]では，SNS ユーザは過去に

投稿したメッセージを削除または非表示にしたいとは考えていなく，時間経過と共に公開範囲を狭めたいと考える傾向にあることが分かっている．また，10～20 代の若者に人気の Snapchat[49]では，テキストや写真などを友人と共有後，一定時間経過後に削除される仕組みとなっており，自撮り画像を中心に，その瞬間を共有する気軽さが受け入れられている[50]．このような背景もあり，Ayalon らは[51]，卒業や引っ越しなど SNS ユーザのライフイベントの契機に合わせて投稿メッセージの公開を制限すべきとし，時間経過による SNS 投稿メッセージの公開期限手法を提案している．（図 2-8 参照）さらに，Vanish[52]は，ユーザが設定した任意時間の経過後，自動的にメッセージを暗号化し，判読不可とする手法を提案している．本論文では，情報の識別境界と同様に，時間境界に応じた保護については未検討である．

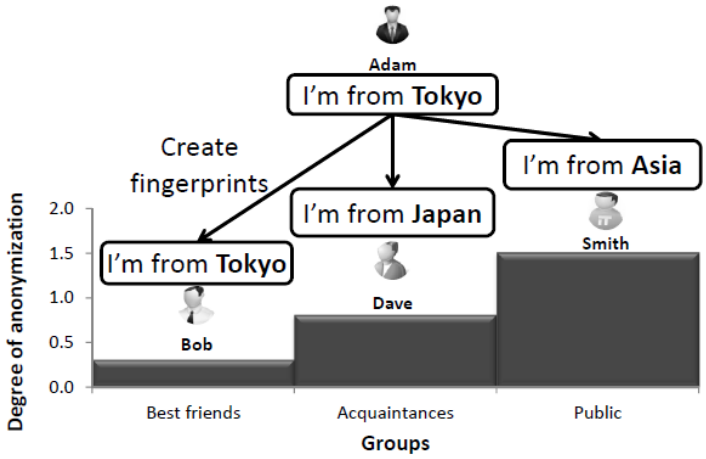


図 2-7　公開対象者との友人関係性に応じた投稿の匿名化

図 2-8　投稿メッセージに対する公開期限設定

2.4　写真の被写体を保護主体とした手法

これまで述べた従来研究の多くは，投稿ユーザのプライバシ保護を主目的としており，投稿内容に含まれる他ユーザのプライバシまで注目できていない．実際に，SNS ユーザの投稿後の後悔事例[24, 29, 32]において，個人特定可能な写真を無断投稿した結果のトラブルが報告されているように，SNS へ投稿する写真に写る/写り込む人物を含む被写体のプライバシを考慮する必要がある．そこで本節では，写真の被写体のプライバシ保護手法について説明する．写真撮影時・SNS 投稿時に被写体のプライバシポリシを反映させるためには，被写体のポリシを撮影者・SNS 投稿ユーザに参照させる必要がある．その際，考え得る手法としては，(1) 顔認識を用いる手法，(2) 電波認識を用いる手法，(3) タグ認識を用いる手法が挙げられる．

2.4.1 顔認識を用いた手法

顔認識を用いる手法では，予め顔の特徴量とプライバシポリシを紐付けて登録しておき，他ユーザによる写真投稿時に顔認識を行い，被写体のポリシを適用・通知を行う[28, 53]．（図 2-9 参照）しかし，これらの手法の場合，タグを常に身に付ける煩雑さはないものの，システム上に顔の特徴量等の人体特性を登録する必要があるため，プライバシ懸念を理由にユーザから利用を拒否される可能性がある[54, 55]．

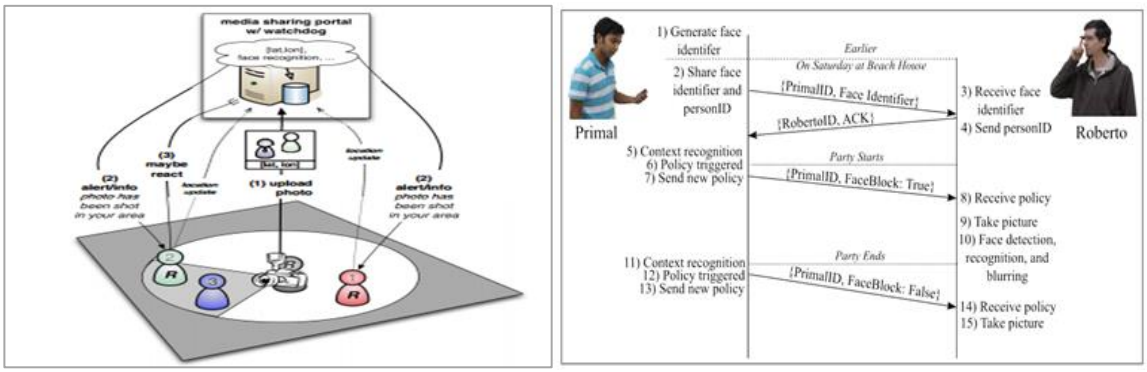
2.4.2 電波認識を用いた手法

次に，電波認識を用いる手法では，主に RFID を用いた手法が提案されている．Cheung ら[56]は，プライバシ保護対象となる人物に RFID タグを持たせ，RFID 追跡システムでタグを検知後，当該タグが持つプライバシポリシに従い，人物の非特定化を行う．同様に，予め RFID タグと個人のプライバシポリシの紐付きをシステム登録しておき，RFID タグを検知した場合に，ポリシに従い背景画像と RFID を持つ人物を置き換える手法[57]が提案されている．（図 2-10 参照）しかし，これらの手法の場合，タグを認識する RFID リーダを写真撮影時に利用されるデジタルカメラなどのデバイスに，統一的な仕様のもと組み込む必要があることが課題となる．

2.4.3 タグ認識を用いた手法

最後に，タグ認識を用いる手法では，RFID タグのように電波を介してポリシ情報の読み取りを行うのではなく，写真に映るタグから直接ポリシを読み取る．そのため，自身のポリシを埋め込んだタグを上半身などに身に付け，写真の投稿時にタグ解析を行い，被写体のポリシを適用する．従来研究では，機械のみでなく，人間も容易に判断可能な単純化されたタグを身に付け，周囲にアピールする手法[58, 59]や，QR コードに多くのポリシを埋め込み，プライバシをコントロールする手法[25, 60]が提案されている．（図 2-11 参照）しかし，単純化

されたタグは多くの情報を保持できないため，コミュニティに応じたプライバシの振る舞いを表現できない．また，QR コードのように多くの情報を含む複雑なタグは，撮影者と被写体までの距離によっては検知・解析精度に課題が残る．



(a) 顔情報をシステム登録する手法　　(b) 顔情報をデバイスに登録する手法

図 2-9　顔認識を用いた手法

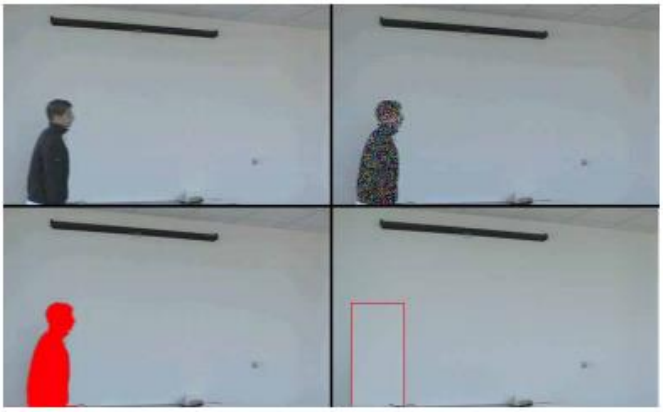
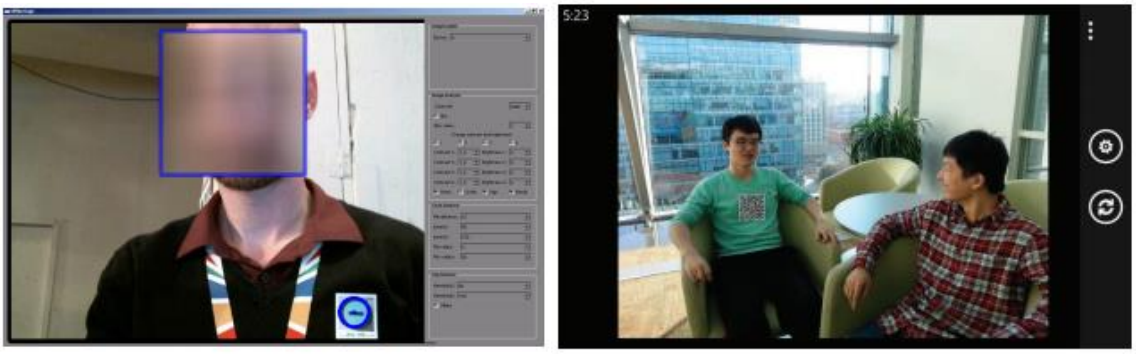


図 2-10　電波認識を用いた手法



(a) 単純化されたタグ　　(b) QR コードベースの複雑なタグ

図 2-11　タグ認識を用いた手法

2.5　まとめ

本章では，最初に SNS 側が提供するユーザのプライバシ保護機能の 1 つである公開範囲の設定機能とその課題について説明した．そして，その結果引き起こされる SNS ユーザの後悔事例を通じて，SNS に投稿すべきでない内容が含まれる場合，投稿前のユーザへの指摘・通知機能の重要性を述べた. そして, SNS に投稿すべきでない内容を正しく判断するためには，客観的な判断基準が必要とされることを述べた.

次に，SNS ユーザのプライバシを保護するための従来手法として，SNS 投稿ユーザを保護主体とした手法，写真の被写体を保護主体とした手法を述べ，これら手法の課題を挙げた．SNS 投稿ユーザを保護主体とした手法では，情報の開示境界を中心に説明し，(1) 投稿の公開範囲，(2) 公開対象となる友人の可視化，(3) 投稿内容に応じたユーザへ指摘・提案に関連した従来手法とその課題について述べた．また，写真の被写体を保護主体とした手法では，SNS へ投稿する写真に写る/写り込む人物を含む被写体のプライバシを保護するために，(1) 顔認識用いた手法，(2) 電波認識を用いた手法，(3) タグ認識を用いた手法を述べ，その課題を説明した.

このような背景から，SNS ユーザが投稿する前にメッセージに含まれるセンシティブデータを自動検知するためには，SNS ユーザにとってどのような情報がセンシティブデータとなり得るのかの判断基準を明確にする必要がある．この判断基準を基にセンシティブデータを検知後，ユーザへ指摘・通知を行うことができるようになる．また，ユーザが持つ，投稿内容に応じた特定の個人や興味を持つコミュニティへのメッセージの公開要望を実現するために，特定コミュニティ・グループにとらわれることなく，投稿メッセージを公開できるようにする必要もある．そして，写真撮影時・SNS 投稿時に被写体のプライバシポリシを反映させるためには，被写体のポリシを撮影者・SNS 投稿ユーザに参照させる手法の検討が必要である.

そこで本論文では，これらの課題対して，(1) SNS におけるセンシティブデータの分類，(2) センシティブデータの漏洩検知に基づく公開範囲の設定方式, (3) 被写体のコミュニティベース・プライバシポリシの設定方式 を提案する．図 2-12 に，本論文におけるこれら提案手法の全体概要を示す.

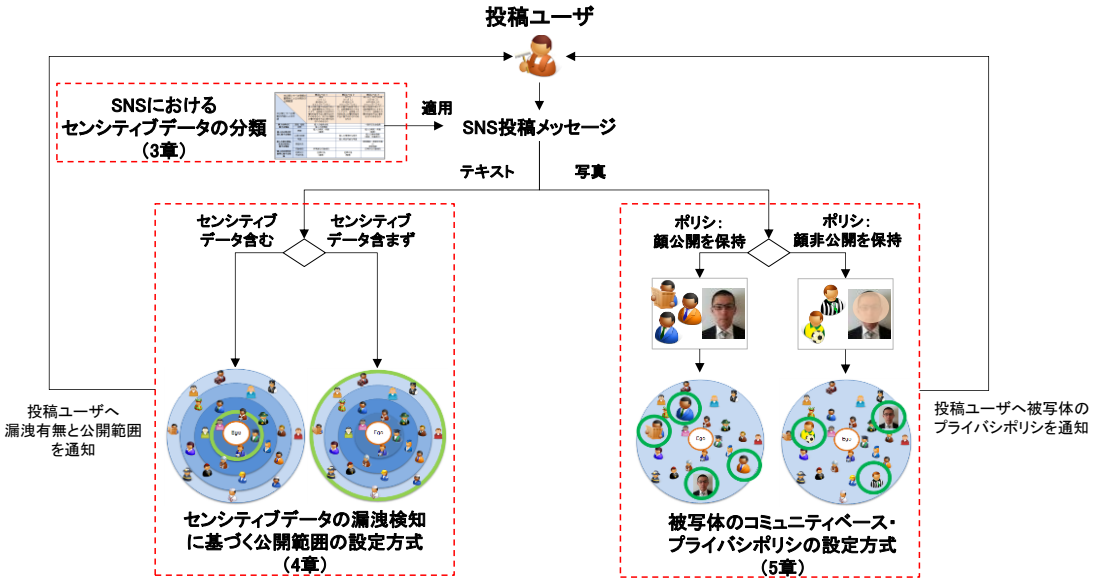


図 2-12　提案手法の全体概要

第3章　　SNS におけるセンシティブデータの分類

各 SNS ユーザは，投稿内容に応じた様々な投稿可否の判断基準を持っているとされる[29]．そして，自身の判断基準は適切であると捉え，どのような状況であっても，常に正しい投稿判断ができると考えるユーザが存在する．実際に，Westin のプライバシセグメンテーションインデックスを用いた評価[61]によると，92%の調査対象者が状況に応じて，適切なプライバシポリシを選択した結果が出ているが，状況に応じて常に正しい選択をすることは容易ではない．また，正しい選択ができたとしても，その各人の判断基準が，個人の主観に依存していることが課題となる．本章では，実際に SNS でセンシティブデータの漏洩が発生していることを確認するため，過去 1 年間分の Twitter アーカイブを用いて，センシティブデータの漏洩有無とその検知手法を試算評価する．また，この評価結果を受け，SNS 投稿予定メッセージに含まれるセンシティブデータを自動検知するための客観的な判断基準の定義の第一歩として，非公開とすべき情報の内容分類と開示レベルを対応付けた，SNS におけるプライバシ侵害情報分類表を提案する．さらに，本分類表の妥当性について，一対比較法[62, 63]を用いた評価調査により示す．

3.1 センシティブデータ漏洩の試算評価

SNS におけるセンシティブデータ漏洩の先行研究において，喜怒哀楽などの感情の変化がある時・アルコール等の影響がある時の SNS 投稿は，その投稿判断を誤ることが多く，自らセンシティブデータを漏洩させてしまい，後日後悔するとしている[24]．また，必要以上にセンシティブデータを公開してしまったうえに，意図していないユーザにまでその情報が漏れていたことに気付き，さらに深く後悔している[32]．SNS ユーザは投稿予定の内容に応じた様々な投稿可否の判断基準を持っているが[29]，このような状況下では，内容に応じた正しい選択をすることが難しい．そこで最初に，SNS ユーザの後悔を引き起こす，ユーザの不用意な投稿によるセンシティブデータの漏洩が実際に SNS で発生していることを確認するために，前述の SNS ユーザの後悔事例[24, 29, 32]の後悔理由もとに，Twitter Streaming API[64]で収集した Twitter アーカイブを用いてセンシティブデータの存在率の試算評価を行った．対象データは，過去 1 年間（2012/6/19〜2013/6/13）に蓄積された日本語 Twitter アーカイブ 137,877,745 件とした．この Twitter アーカイブに対して，これらの後悔事例より，(1) 個人の内心に関する情報として “宗教”，(2) 個人の心身の状態に関する情報として“病歴”と“心身の記録”，(3) 個人の基本情報・生活状況に関する情報として“行動傾向”および“写真”に関連した情報の含有を分析し，センシティブデータの漏洩が実際にどの程度発生しているのかを確認した．

3.1.1 個人の内心に関する情報

個人の内心に関する情報として，“個人の宗教観”を検出した．キーワードマッチング処理後のツイートに対して，SNS ユーザ自身が肯定的な意見を持つ，または否定的な意見を持つ場合に，個人の宗教観が表現された投稿であると判断した．また，投稿が肯定・否定のどちらでもない場合は，中立的な意見として一般的な宗教観とし，個人の宗教観には含めないものとした．

処理手順は以下である．

(1) 137,877,745 件の Twitter アーカイブから Wikipedia:宗教一覧[65]より，キリスト教や仏教の各宗派名，新興宗教名等をキーワードとして，キーワードマッチングを行う．
(2) SNS 投稿メッセージ内に投稿ユーザ自身の意見を含まないことが多いと考えられるハッシュタグ・リツイート・URL 付きのツイートを除去する．
(3) 抽出された 17,123 件のツイートから，ランダムに 1,000 件をサンプル抽出する．
(4) 抽出された 1,000 件のツイートをユーザの意見が含まれるか否かで判別し，(1)肯定的な意見，(2)否定的な意見，(3)中立的な意見，(4)その他の 4 値で分類する．調査の確度を高めるために，2 人により上記分類を行い，2 人が同値で判別したツイートを正しい判定と見なす．なお，複数ツイートからの判別ではなく，1 ツイートから含まれるか否かを判断することとする．

表 3-1 は，上述の 4 分類に対する検出率を示している．個人の宗教観は，肯定と否定の合計より，14%が該当した．

表 3-1　分類結果：個人の内心に関する情報

分類	検出率
肯定	1%
否定	13%
中立	10%
無関連	76%

3.1.2 個人の心身の状態に関する情報

個人の心身の状態に関する情報として，“個人の病気・疾患（重度）”，“個人の精神的な症状”，“個人の病気・疾患（軽度）”を検出した．キーワードマッチング処理後のツイートに対して，SNS ユーザ自身，または周囲の情報を含むか否かを判断した．

処理手順は以下である.

(1) Yahoo! 家庭の医学[66]より，癌，鬱，エイズ等の病気・疾患に対して，医療の要・不要により，重度（不治の病，医療者がかかわらないと治癒に至らない）・軽度（医療者にかかわらなくても自然治癒力や本人の努力次第で改善する）の分類[67]を行う．軽度に分類された病気・疾患の中でも精神疾患については，精神症状として分類する.

(2) 137,877,745 件の Twitter アーカイブから，医療の要・不要により分類した病気・疾患名をキーワードとして，キーワードマッチングを行う.

(3) SNS 投稿メッセージ内に投稿ユーザ自身の意見を含まないことが多いと考えられるハッシュタグ・リツイート・URL 付きのツイートを除去する.

(4) 抽出された 96,318 件のツイートから，ランダムに 10,000 件をサンプル抽出する.

(5) 抽出された 10,000 件のツイートを，ユーザ自身・周囲の情報を含むか否かを判別し，(1)ユーザの情報を含む，(2)なしの 2 値で分類する．なお，複数ツイートからの判別ではなく，1 ツイートから含まれるか否かを判断することとする.

3.1.1 項と同様に調査の確度を高めるために，2 人による分類を行い，2 人が同値で判別したツイートを正しい判定と見なし，各分類に対する検出率を算出した．表 3-2 は，各分類に対する検出率を示している．サンプル抽出全体の 4.23%のツイートが，いずれかの分類に該当する結果となった．個人の病気・疾患（重度）は，癌・脳梗塞等の合計から 1.63%，個人の精神的な症状は，鬱・パニック障害等の合計から 0.67%，個人の病気・疾患（軽度）は 1.93%であった.

表 3-2　分類結果：個人の心身の状態に関する情報

ユーザ情報有無	分類	病気・疾患名	検出率
あり	個人の病気・疾患（重度）	癌	0.97%
		脳梗塞	0.21%
		その他	0.45%
	個人の精神的な症状	鬱	0.5%
		その他	0.17%
	個人の病気・疾患（軽度）	血尿	0.25%
		その他	1.68%
なし			95.77%

3.1.3 個人の基本情報，生活状況に関する情報

個人の基本情報，生活状況に関する情報として，“非常識な行動傾向”，“個人特定可能な写真”，“日常的な行動傾向”を検出する．SNS ユーザが喜怒哀楽など感情に起伏があるときの

SNS 投稿は，自らセンシティブデータを漏洩させやすいといわれている[24]ことから，ツイートがポジティブ（P）・ネガティブ（N）であるかの感情極性を判別し，感情の高ぶりを検知した．感情極性の判別には，日本語評価極性辞書[68]を利用した．

処理手順は以下である．

(1) Twitter アーカイブに対して，形態素解析器 Mecab[69]を用いて形態素解析し，日本語評価極性辞書を基に単語ごとの P/N 判定を行う．
(2) 単語ごとの P/N を集計し，ツイートごとの P/N スコアを 1.0 ～0.0 で算出する．ポジティブ傾向であれば，より 1.0 に近づき，ネガティブ傾向であれば，より 0.0 に近づいた値となる．
(3) ポジティブ傾向・ネガティブ傾向にあるツイートを取得するため, P/N スコアに閾値（0.75 以上，または 0.25 以下）を設定し，対象ツイートを絞り込む．
(4) 感情の高ぶりがあるときの画像付きの SNS 投稿は，平常時と比較して，センシティブデータが含まれるか否かを判断することが，より難しいと仮定し，Instagram や Twitpic などの画像共有サービスの URL 付きツイートに絞り込む．
(5) 抽出された 876,976 件のツイートから，ランダムに 7,000 件をサンプル抽出する．
(6) 抽出された 7,000 件のツイートを, 2 人により, 投稿テキストに紐づく画像の確認を行い, (1)非常識な行動傾向，(2)個人特定可能な写真，(3)日常的な行動傾向，(4)その他の 4 値で分類する．

本項も同様に，2 人による分類を行い，双方が同値で判別したツイートを正しい判定と見なし，各分類に対する検出率を算出した．表 3-3 は，各分類に対する検出率を示している．非常識な行動傾向は検出できなかったが，個人特定可能な写真，日常的な行動傾向は計 0.54%の検出率であった．

表 3-3　分類結果：個人の基本情報，生活状況に関する情報

分類	検出率
非常識な行動傾向	0%
個人特定可能な写真	0.35%
日常的な行動傾向	0.19%
その他	99.46%

3.1.4 検知結果と考察

これまでのサンプルツイートに対する検知結果を，137,877,745 件の全ツイートに適用すると換算した場合，センシティブデータの漏洩として，計 0.008%の検知となった．センシティ

ブデータの漏洩検出率は，いずれの分類も低い結果であったが，一度漏洩してしまった情報をなかったことにすることは難しく，また，失職や損害賠償など社会的な制裁を受ける可能性を考慮すると，ごく少数であっても，SNS 提供者側に未然にその漏洩を検知・指摘する仕組みが必要である．

3.2 プライバシ侵害情報の分類

3.2.1 公文書におけるプライバシ情報の取り扱い

前節の試算評価を通じ，SNS においてユーザ本人がセンシティブデータを漏洩させてしまうことが確かめられた．したがって，投稿前にユーザのメッセージ内容を検査して，情報漏洩を防止するシステムを構築することは重要である．当該システムの構築にあたり，ユーザの投稿予定メッセージの中にセンシティブデータを含むか否かを判定するための客観的な判断基準が必要である．実社会において同様の判断基準を必要とする文書として，今日までの歴史的な史料や公的文書である公文書があげられる．公文書の管理・公開の役割を担う機関として，国・都道府県・市町村等が管轄する公文書館[70]が存在する．公文書館では，個人に関して影響がない場合，所蔵資料を可能な限り一般に公開することを理想しているが，公文書等の公開にあたり公開基準を設けており，その基準の 1 つにプライバシ侵害がある[71]．これは公開対象となる公文書に，プライバシ等の人権侵害，個人・法人の権利権益を不当に害する恐れのある情報が含まれている場合に考慮が必要なためである[72]．また，文書作成後の時の経過への考慮[73]として，一般に時の経過とともに，非公開とすべき理由が減少するとされる．文書公開に関する国際アーカイブズ評議会の 30 年原則[74]では，文書の作成から公開までの国際的な標準を設定し，一般的な資料は原則，一定期間の経過後に公開すべきとした．これらをふまえた日本国内における公文書館の公開基準は，主に，情報公開条例方式と国立公文書館方式の 2 種類に分類ができる．前者の情報公開条例方式は，非公開とすべき情報を詳細に分類し，その分類ごとに非公開期間を設定する．後者の国立公文書館方式では，非公開とすべき情報の詳細分類はせずに重み付けを行い，状況に応じた柔軟な対応をとる．そして，この両方式の利点を取り入れた公文書の公開基準として，戸嶋によるプライバシ等侵害情報分類表（私案）[72]があげられ，プライバシ等侵害の度合いを区分しつつ，保護期間を定めている．（付録 A.1 を参照）．

このプライバシ等侵害情報分類表では，縦軸は情報公開条例をベースとした「非公開とすべき情報の内容による分類」を表し，6 つの分類が定義されている．一方，横軸は「非公開とすべき情報の重要度による分類及び非公開期間」を表し，国立公文書館利用規則[75]の個人プライバシ等侵害の重さをベースとした 3 つの分類と，分類に応じた非公開期間（30 年，50 年，120 年）が定義されている．本分類表を SNS への適用を検討するが，本分類表は国家に関連した文書を含む公的な文書への適用を目的としているため，本来 SNS に直接適用することが難しい．そこで，本論文では憲法の私人間効力[87]を参考に，SNS を社会的影響力の大

きい私人として見立て，本分類表を適用することする．しかし，適用するためには，(1) 上述の分類表の縦軸にあたる，非公開とすべき情報の内容による分類に，SNS において発生し難い分類が含まれていること，(2) SNS 投稿では原則，投稿後に非公開期間を待たずに即時公開となるため，分類表の横軸が表す，分類の重要度に応じた非公開期間に置き換わる指標が必要であることが課題となる．本論文では，この公文書の分類表をもとに，SNS 投稿時に発生する事実に基づいたプライバシ侵害情報分類表を提案する．上述の 2 つの課題を解決するために，(1) SNS 投稿後に感じた後悔理由をもとに，プライバシ侵害情報分類表の縦軸の「非公開とすべき情報の内容による分類」を定義し，(2) 非公開期間に代わる指標として，「重要度に応じた情報の公開範囲」を定義する．

3.2.2 非公開とすべき情報の内容による分類

SNS におけるプライバシ侵害情報分類表の，縦軸の「非公開とすべき情報の内容による分類」を検討するにあたり，SNS 投稿時に発生する事実に基づいて修正するため，ユーザが SNS 投稿後に感じた後悔事例から SNS で発生するセンシティブな事象を分析する．本論文では，3.1 節の Twitter アーカイブにおけるセンシティブデータの存在率の試算評価で使用した既存研究[24, 29, 32]より，SNS 投稿後の後悔理由を抽出し，SNS 投稿時の非公開とすべき情報の内容による分類を行った．表 3-4 は，SNS における非公開とすべき情報の内容による分類と，その後悔例を示している．

表 3-4　SNS における非公開とすべき情報の内容による分類

分類	分類説明	後悔例
個人の内心に関する情報	個人の思想や信仰宗教の情報	自身の信仰宗教に関することを投稿後，仕事上で好ましく思われなかった
個人の心身の状態に関する情報	個人や家族の精神症状・身体情報を含む病歴情報	個人・家族の病歴や病状を軽々しく公開した
個人の基本情報，生活状況に関する情報	自身や他者が特定される可能性のある写真，個人や家族に関する情報	友人が写り込んだイベントの写真を無断で投稿した
	非常識な行動，日常生活における個人の活動状況や趣味趣向などの情報	感情に任せて，ネガティブ/攻撃的なコメントを投稿した
個人の社会的活動等に関する情報	自身が気付いていない場合を含む犯罪行為などの情報	未成年にも関わらず，アルコールを飲み，それを投稿した

3.2.3 重要度に応じた情報の公開範囲

公文書では，意図していないユーザに情報が公開されることを防ぐために，文書の内容による非公開対象者を設定する基準はなく，一定の非公開期間と文書の部分公開によりプライバシ情報が漏れることを防ごうとしている．これは原則，即時公開かつ投稿内容の部分公開をなしとする SNS 投稿メッセージには適用できない公開基準となる．そこで，横軸の非公開とすべき情報の「重要度による非公開期間」に置き変わる指標を検討した．2 章で論じたように，IT 時代におけるプライバシ管理の主要事項の提案として，開示境界，識別境界，時間境界の 3 種類を提唱している[37]．これらのうち，時間境界とは，過去，現在そして未来等の時間経過における投稿の開示境界を指す．本境界に関する従来研究として，メッセージ公開後の公開期間に期限を設け，投稿メッセージの公開を制限する手法[51]などがあるが，前述のように SNS では原則，メッセージを即時公開とすることから，メッセージ公開前の非公開期間を設ける手法は提案されていない．そこで本論文では，「重要度による非公開期間」に置き変わる指標として開示境界に着目し，情報の「重要度による公開範囲」を用いて，当該情報が公開可能な事項であるかの境界として定義した．

SNS ネットワーク内の公開範囲に関する既存研究として，La Gala ら[43]による SNS エゴネットワークにおける友人分類に関する研究があげられる．エゴネットワークとは，社会ネットワーク分析において，自分自身を中央に配置し，他ユーザとの関係構造を表現したネットワーク構造を指す．たとえば Facebook や LinkedIn における友だちリストは自分自身を中心としたエゴネットワークを構成していると見なされる．先行研究では，SNS エゴネットワークに対して，Dunbar’s number [76, 77]を適用し，SNS 内のコミュニケーション頻度を対応付けることにより，SNS エゴネットワークと実社会における友人分類が似ていることが示されている．Dunbar’s number では，認知科学の観点から人間にとって，それぞれの人間と安定した関係を維持可能な上限数は平均 150 人程度としたが，SNS においてもこの維持可能な友人数が変わらないとされている[78]．たとえば，Facebook において友だちが 150 人以上存在する場合，自身の友だちリスト内に友人関係が非常に弱いユーザが含まれていることを意味し，誤ってセンシティブデータを投稿してしまった場合に，これら紐付きの弱いユーザにも情報が漏洩することになる．

本論文では，この Dunbar’s number と La Gala らの SNS 内コミュニケーション頻度による分類をベースに，非公開とすべき情報の重要度による公開範囲として，新たに SNS エゴネットワークにおける開示レベルを定義した．表 3-5 は，開示レベルと人数・コミュニケーション頻度の対応を示している．

最も厳しい開示レベルを示す開示レベル 0 では，公開対象を“家族”と定義し，コミュニケーション頻度とその人数は未設定とした．開示レベル 1 では，公開対象は，“親友”と定義した．親友とは週に 1 回以上のコミュニケーションがあり，その人数は 1〜5 人とした．開示レベル 2 の公開対象は，“友だち”とし，月 1 回以上のコミュニケーションがある 6〜15 人と定義した．開示レベル 3 は“知人以上 友だち未満”とし，半年に 1 回以上のコミュニケー

ションがある 16～50 人と定義した．開示レベル 4 は“知人”とし，その人数は 51 人～150 人とした．最も公開範囲が緩い開示レベル 5 は，まったくコミュニケーションがない，もしくは，1 年以上音沙汰がないケースを指す“他人”とした．なお，下位の開示レベルは，上位の開示レベルの公開対象も含むものとし，分類表において複数の分類が該当した場合は，より厳しい開示レベルを採用する．図 3-1 は，Dunbar’s number をサークル上に表現した Dunbar’s circle と開示レベルの対応を示している．

表 3-5　SNS エゴネットワークにおける開示レベル

開示レベル	公開対象	人数（人）	コミュニケーション頻度
0	家族	–	–
1	親友	1～5	週 1 回以上
2	友だち	6～15	月 1 回以上
3	知人以上友だち未満	16～50	半年 1 回以上
4	知人	51～150	年 1 回以上
5	他人	151～	年 1 回未満

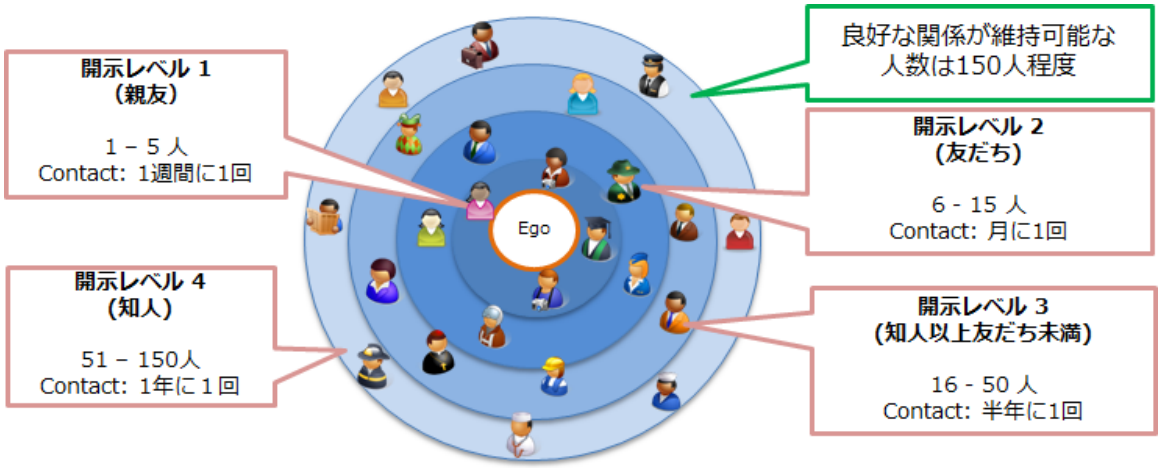


図 3-1　Dunbar’s circle と開示レベルの対応

3.2.4 SNS におけるプライバシ侵害情報分類表の提案

戸嶋により提案された公文書におけるプライバシ等侵害情報分類表[72]に，3.2.2 項で定義した縦軸の非公開とすべき情報の内容による分類と，3.2.3 項で定義した横軸の重要度による公開範囲の定義を適用することによって，SNS におけるプライバシ侵害情報分類表を提案し

た．表 3-6 は，本論文において提案したプライバシ侵害情報分類表を示している．適用にあたり，縦軸の非公開とすべき情報の内容による分類では，SNS 投稿後の後悔事例の確認ができなかった分類に関して，SNS では発生し難い分類として，本分類表の対象から除外した．また，横軸の重要度による公開範囲では，開示レベル 0～5 を定義したが，クローズド SNS である Path [79]では，サービス開始当初，より深いつながりを持つ少ない人数として，家族を含めた友だちリストの上限を 50 人に設定し，閉ざされた SNS エゴネットワークを表現した．これにより，Facebook 等のオープン SNS からプライバシを保ちたいユーザを多く獲得してきたことを考慮し，本分類表では，開示レベル 0 の“家族”をすべての開示レベルに含むものとし，50 人を開示範囲の上限として，開示レベル 1～3 を分類の重要度に応じて設定した．

表 3-6　SNS におけるプライバシ侵害情報分類表

非公開とすべき情報の重要度による分類及び公開範囲 / 非公開とすべき情報の内容による分類		開示レベル 1 親友 （1-5 人） 週 1 回以上のコミュニケーション 個人の特に重大な秘密であって，当該情報を公にすることにより，当該個人の生存中の権利・利益を不当に害する恐れがある	開示レベル 2 友だち （6-15 人） 月 1 回以上のコミュニケーション 個人の重大な秘密であって，当該情報を公にすることにより，当該個人の社会生活上の権利・利益を不当に害する恐れがある	開示レベル 3 知人以上 友だち未満（16-50 人） 半年 1 回以上のコミュニケーション 個人の秘密であって，当該情報を公にすることにより，当該個人の権利・利益を不当に害する恐れがある
個人の内心に関する情報	思想，信条	個人の信条主張		一般的な社会信条
	宗教	個人の宗教観		
個人の心身の状態に関する情報	病歴	個人の病気・疾患（重度）		個人の病気・疾患（軽度）
	心身の記録		個人の精神的な症状	個人の身体情報 (身長・体重含む)
個人の基本情報，生活の状況に関する情報	写真		個人特定可能な写真	
	家庭状況			家族構成・家庭状況等の家族情報
	行動傾向	非常識な行動傾向		日常的な行動傾向
個人の社会的活動等に関する情報	犯罪及び不法行為	犯罪行為（重度）	犯罪行為（軽度）	

3.3 評価

本節では, 前節で提案したSNSにおけるプライバシ侵害情報分類表の妥当性の評価として, SNS 投稿時にプライバシの観点から，本分類表の各分類がどの程度重要であるのか，情報の重要性を一対比較法[62, 63]を用いて評価調査を行う．ただし，本評価は社会調査ではなく，4 章以降のシステム提案・実装への適用を目的として実施する.

3.3.1 評価方法

一対比較法の評価方法に従って，下記の条件で評価を実施した.

(1) 調査対象者

調査対象者は，大学生，大学院生，教員，会社員，主婦から構成される 20 代～60 代の 39 人（男 : 21 人，女 : 18 人）とした．調査対象者の年代は，30 代（40%）が最も多く，続いて 20 代（33%），40 代（18%）・60 代（7%）・50 代（2%）である.

(2) 評価対象の分類

前節で提案した SNS におけるプライバシ侵害情報分類表の 13 種類の分類を用い，各分類および，分類ごとにあらかじめ分類したサンプル Tweet メッセージに対して，SNS 投稿時のプライバシ観点での情報の重要性を，一対ずつ比較し，順位付けおよび，順位の信憑性評価を行った. その際, 内心と行動傾向等, 異なる性質の比較を極力避けるため, 評価対象となる 13 種類の分類を 2 つの比較グループ（(A) 個人の内心・心身の状態に関する情報，(B) 個人の基本情報・生活の状況，および社会的活動等に関する情報）として，前者のグループ(A)を分類の(1)～(7) ，後者のグループ(B)を分類の(8)～(13)に分けて実施した．表 3-7 に本評価で比較対象とする分類を示す.

3.3.2 評価手順

同一グループ内から異なる分類を 2 つ取り上げ，一方の分類を基準分類，他方の分類を比較分類とし，SNS 投稿時にプライバシの観点からどのくらい重要であるかを主観的に評価した．これを比較グループ A，B のそれぞれにおいて総当りに実施する．調査対象者は，基準となる分類を 0 ととらえた上で，0 を含む－2（重要でない）から＋2（重要である）までの 5 段階評価を行う．また，すべての分類を基準分類として評価する.

評価手順を以下に示す.

(1) 調査対象者に 13 種類のプライバシ侵害情報の分類を教示し，分類の説明を行う．（図 3-2 Step 2）

(2) 調査対象者に，基準分類，比較分類の 2 分類を提示する.

(3) 調査対象者は，提示された基準分類を基準に，比較分類が SNS 投稿時にプライバシの観点からどの程度重要であるかを－2～2 の 5 段階の評価尺度により評価する.（図 3-2　Step 3-1）

(4) すべての分類を基準分類として提示し，グループ内の全分類と総当りに比較を行う.

(5) 実際に過去に投稿され，あらかじめ各分類に分類したサンプル Tweet メッセージおよび，メッセージに紐付いた画像を提示し，(3), (4)を繰り返す.（図 3-2　Step 3-2）

(6) 調査対象者の回答のうち，SNS を週 1 回以上利用する調査対象者 26 人（男：13 人，女：13 人）の回答に対して，一対比較法の評価方法*2 に従い，グループごとに分類間の順位の算出と，分類間の有意差の有無を確認する.

表 3-7　比較対象とする分類

比較グループ	番号	分類名
(A) 個人の内心・心身に関する情報	1	個人の信条主張
	2	一般的な社会信条
	3	個人の宗教観
	4	個人の病気・疾患（重度）
	5	個人の病気・疾患（軽度）
	6	個人の精神的な症状
	7	個人の身体情報 （身長・体重含む）
(B) 個人の基本情報・生活の状況，及び社会的活動等に関する情報	8	個人特定可能な写真
	9	家族構成・家庭状況等の家族情報
	10	非常識な行動傾向
	11	日常的な行動傾向
	12	犯罪行為（重度）
	13	犯罪行為（軽度）

*2 一対比較法の最低評価対象者数は 20 人程度[80]とされるため，評価としては成立すると考えているが，社会調査としては評価対象者の代表性を満たせていない.

図 3-2　評価手順

3.3.3 評価結果

比較グループ A, B におけるプライバシ観点での情報の重要性の評価結果を図 3-3 に示す. 図 3-3 の横軸は，分類ごとの尺度値を比較するために，数直線上に尺度表現している．プラス方向である左に行くほど，SNS 投稿時におけるプライバシ観点での情報の重要性が高いことを示しており，マイナス方向の右に行くほど，重要性が低いことを示している．(6) 個人の精神的な症状を除き，評価グループ A，B ともに，表 3-6 の SNS におけるプライバシ侵害情報分類表の開示レベルに沿った順位付けとなることが分かった.（図 3-4 および図 3-5 を参照）たとえば，評価グループ B：個人の基本情報・生活の状況，および社会的活動等に関する情報において，最も重要性が高いとされた(12) 犯罪行為（重度），次点の(10) 非常識な行動傾向は，開示レベル 1 に分類されている．また，最も重要性が低いとされた(11) 日常的な行動傾向は，開示レベル 3 に分類されている．分類表上で開示レベル 2 に分類される(6) 個人の精神的な症状は，より重要性が高くとらえられる結果となったが，評価手順 5 で実施したあらかじめ分類したサンプル Tweet メッセージの評価結果においては, (3) 個人の宗教観よりも下位の順位であり，かつ，両分類の間には，有意差があった．以上の結果より，表 3-6 で提案した SNS におけるプライバシ侵害情報分類表が，SNS 投稿時のセンシティブデータであるかの判断基準として有効であることが分かった．ただし，本評価の調査対象者に偏りの可能性とその人数が十分とは言えず，社会評価としては不成立であるため，さらなる評価の充実を今後の課題とする.

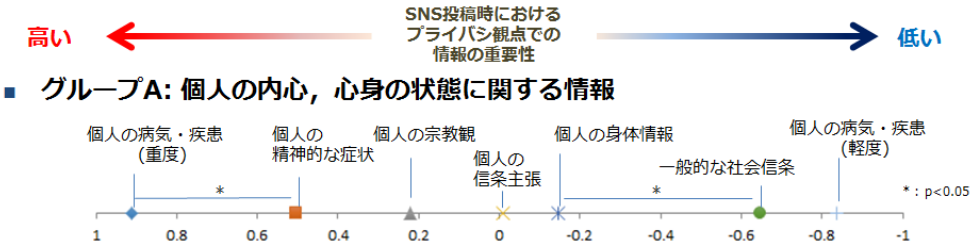


図 3-3　プライバシ観点での情報の重要性の評価結果

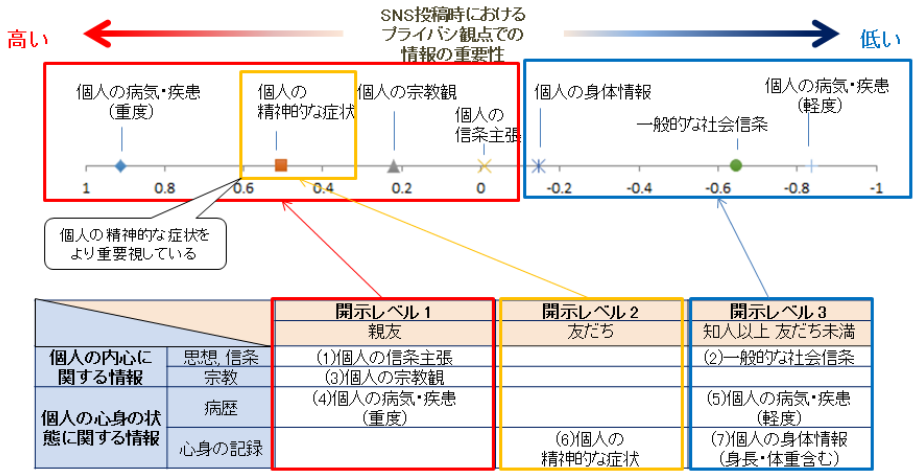


図 3-4　グループ A：評価結果と提案分類表の対応

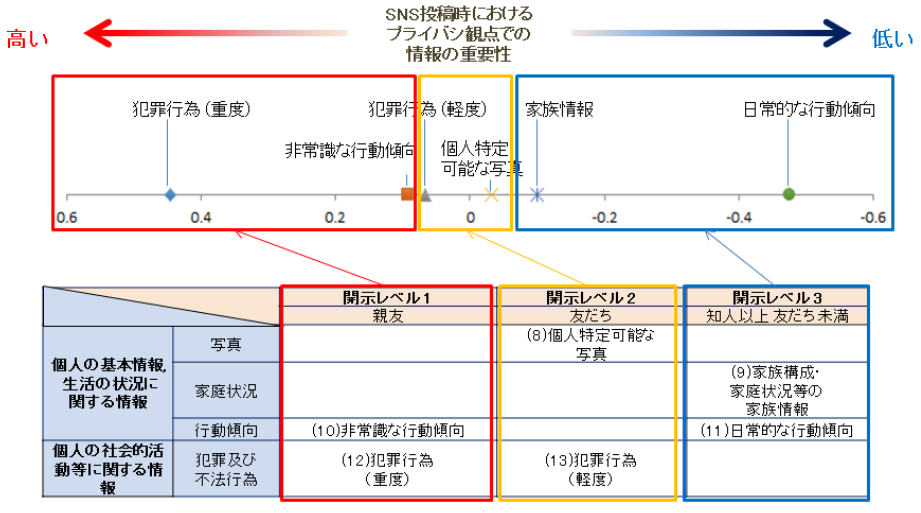


図 3-5　グループ B：評価結果と提案分類表の対応

3.4 まとめ

最初に，SNS ユーザの不用意な SNS 投稿によるセンシティブデータの漏洩が，実際に発生していることを確認するために，SNS 投稿後のユーザの後悔事例をもとに，137,877,745 件の Twitter アーカイブにおけるセンシティブデータの存在率の試算評価を行った．その結果，計 0.008%のセンシティブデータの漏洩を検知し，非常に少数であるが，センシティブデータを含んだ SNS 投稿が実際に行われていることを示した．本試算評価を受け，SNS 投稿時のセンシティブデータの漏洩を事前に自動検知するための第一歩として，公文書館におけるプライバシへの取り組みをベースに，非公開とすべき情報の内容分類と開示レベルを対応付けた SNS におけるプライバシ侵害情報分類表を提案した．本分類表の妥当性評価として，一対比較法を用い，SNS 投稿時のプライバシ観点での情報の重要性の評価を実施し，社会調査としては未成立であるが，システム適用に向けてその妥当性を示した．4 章では，本分類表を用いて，センシティブデータが含まれた SNS 投稿メッセージの内容から漏洩の有無とその情報に応じた開示レベルを判断し，センシティブデータの漏洩有無の通知と，投稿の公開範囲・公開対象者の自動設定を提供する設定方式を提案する．

第4章　センシティブデータの漏洩検知に基づく公開範囲の設定方式

投稿ユーザが投稿メッセージ内にセンシティブデータを含んでいることに気付かず投稿してしまい，その結果，センシティブデータの漏洩と意図していないユーザへ投稿を共有してしまう課題に対して，前章で提案した SNS におけるプライバシ侵害情報分類表を適用することで，SNS 投稿メッセージの内容からセンシティブデータ漏洩の有無と，そのメッセージの公開範囲として開示レベルを客観的に判断できるようになる．本章では，提案した分類表が持つこの特性を活かし，センシティブデータの漏洩有無の検知・通知と，その検知結果に基づいた公開範囲・公開対象者の自動設定を提供する設定方式を提案する．さらに，対象 SNS を Facebook として，提案手法のシステム実装を示す．

4.1 センシティブデータの漏洩検知に基づく公開範囲の設定方式

2016 年 6 月現在，Facebook では投稿時の公開範囲の設定として，“公開”，“友だち”，“自分のみ”，“SNS 側で作成される特定グループ”，“ユーザによりカスタマイズ可能なグループ”を提供している．Twitter においては，“公開”・“個別ユーザに対する”ダイレクトメッセージ機能を提供している．多くの SNS ユーザは通常，公開範囲 :“友だち”を設定している[17]．しかし，公開範囲 :“友だち”の中には，関係性の弱いユーザが含まれており，公開範囲を狭めたとしても，依然としてプライバシ漏洩の危険性がある[12]．また，ユーザによっては，投稿内容によって，公開対象として望ましいユーザの選択や望ましくないユーザの除外を行っているが[31]，投稿のつど行っていくことは難しい．このことから，SNS ユーザは，現在 SNS 側より提供されている投稿の公開範囲設定以外に，投稿内容によって特定の個人や興味を持つグループに対して，安全に効果的に公開したいと考えている[29]．

そこで，ユーザが SNS へ投稿する前に，投稿予定メッセージに対して，表 3-6 の SNS におけるプライバシ侵害情報分類表を適用し，センシティブデータの検知を行う．センシティブデータを含まない場合，ユーザが事前に設定する公開範囲設定に従い，SNS へ投稿する．センシティブデータを含む場合，センシティブデータ漏洩の通知と，投稿ユーザを中心とした SNS ネットワークにおける投稿メッセージの公開範囲を自動設定するシステムを提案する．

4.2 システム設計

4.2.1 概要

SNS ユーザの不用意な投稿によるセンシティブデータの漏洩を防止するために，センシティブデータの漏洩検知に基づく公開範囲の設定方式を提案する．本方式では，ユーザが SNS へ投稿を行う前に，事前利用することで，投稿予定メッセージ内にセンシティブデータが含まれるか否かを判別する．センシティブデータを含む場合，センシティブデータの漏洩通知と，誰に対して参照可能とすべきかの公開範囲を導出し，ユーザに提案を行う．ユーザはこの公開範囲の提案に対して，公開対象者の追加・除外といった改訂をすることができ，最終的な SNS への投稿判断は投稿ユーザに委ねることとする．また，その改訂情報は次回投稿時の公開範囲提案に考慮される．これにより，ユーザはセンシティブデータを，意図していない他ユーザに対して公開することなく，公開範囲の管理が容易に可能となる．

4.2.2 処理フロー

本システムのプロセスフローを図 4-1 に示す．本システムは大きく分けて，センシティブデータの漏洩検知，公開範囲の導出，結果通知の 3 プロセスから構成される．以下では，そのプロセスを説明する．

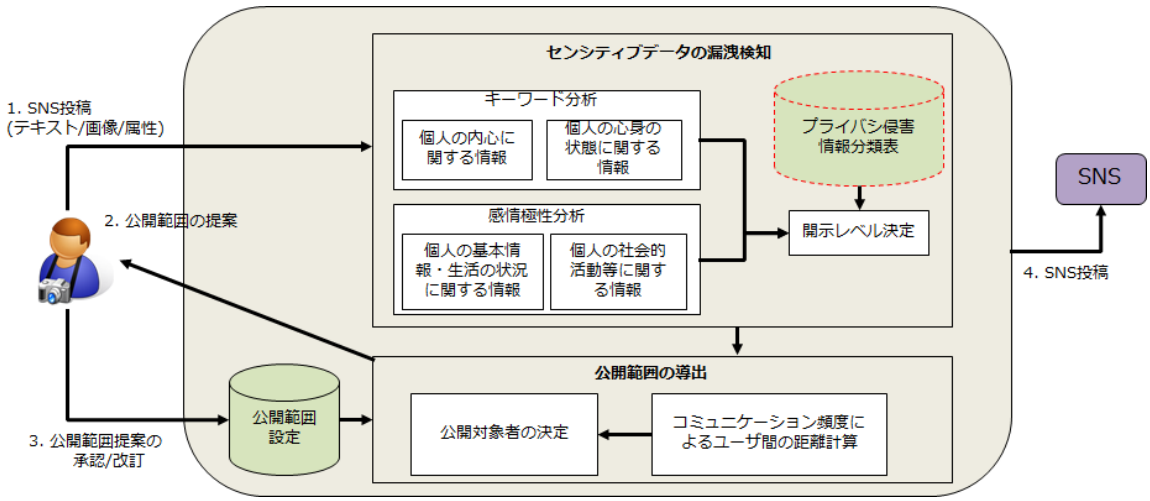


図 4-1　提案システムのプロセスフロー

Step 1. センシティブデータの漏洩検知

テキスト・画像・属性情報から構成される SNS 投稿メッセージを分析し，センシティブデータの含有を検知する．その判断基準として，3 章で定義した SNS におけるプライバシ侵

害情報分類表を適用するが，分類表の縦軸の非公開とすべき情報の内容による分類により，分析方針が異なる．個人の内心に関する情報，個人の心身の状態に関する情報は，関連キーワードをベースとした検知を行う．それ以外の個人の基本情報，生活状況に関する情報，社会的活動等に関する情報等は，感情の高ぶりを含めた感情起伏が出ているときに判断が緩くなり投稿した場合，センシティブデータの漏洩傾向があるため，投稿テキストにおける感情極性の分析によりポジティブ（P）/ ネガティブ（N）を検知する．また，センシティブデータの漏洩が検知された場合，プライバシ侵害情報分類表に従い，投稿の開示レベルを決定する．

Step 2. 公開範囲の導出

Step 1 で決定した開示レベルに応じて，自身を中心とした SNS エゴネットワーク内における友だちとのコミュニケーション頻度による距離を測り，最終的な公開対象者を決定する．その際，以前の投稿に対する改訂情報を考慮し，公開対象者の追加・除外を行う．これにより，コミュニケーション頻度が少なくても，友人関係性が強いケース，またはその逆のケースに対する考慮となる．

Step 3. 結果通知と公開範囲の提案

分析結果であるセンシティブデータの漏洩通知と，投稿に対する公開範囲として公開対象者を SNS ユーザへ通知する．SNS ユーザは本システムからの公開範囲の提案に対する投稿可否判断として，承認/改訂を行った後，SNS へ本投稿する．承認/改訂情報はシステム側に蓄積され，次回の投稿時の公開範囲提案に考慮される．

4.2.3 要求機能

前項までの本システムの概要および，分析プロセスをふまえ，システムの実現時に要求される基本機能および，データ処理手順を説明する．

■　基本機能

本システムでは，SNS ユーザのプライバシを保護する手法として公開範囲にフォーカスし，(1) SNS 投稿メッセージに含まれるセンシティブデータの漏洩検知，(2) 投稿メッセージの内容による適応的な開示範囲の提案手法の確立を目的としている．本手法の目的のために，システムの基本機能として以下の 6 機能が必要とされる．

- ユーザ視点
  1. 公開範囲として開示レベルを自由に設定できる．
  2. 設定した開示レベルに沿うようにユーザ単位で公開対象者を自由に選択できる．
- システム視点
  1. 投稿予定メッセージにおけるセンシティブデータの含有を検知する．

2. 検知したセンシティブデータに応じた公開範囲として開示レベルを導出し，ユーザへ提案する．
3. 開示レベルの範囲内に指定された公開対象者のみに，投稿メッセージを公開する．
4. 最終的に公開対象者としたユーザ情報を保持し，次回の開示レベル提案時に考慮する．

■　**データ処理手順**

データ処理手順として，基本機能のシステム視点 1〜4 を説明する．

Step 1. センシティブデータの漏洩検知

あらかじめ，プライバシ侵害情報分類表の分類ごとに，開示レベル 1〜3 および，その他として判別する多値分類器を作成する．分類器の作成方法としては，3.1 節の Twitter アーカイブにおけるセンシティブデータの存在率の試算評価と同様に，個人の内心に関する情報と個人の心身の状態に関する情報は，キーワード分析後の教師データから分類器を作成する．個人の基本情報，生活の状況に関する情報，個人の社会的活動等に関する情報については，感情分析後の教師データから同様に分類器を作成する．ユーザが SNS へ投稿予定メッセージを本システムに投稿した場合，あらかじめ作成されたすべての分類器を通し，最も狭い開示レベルを適用する．

Step 2. 公開範囲の導出

Step 1 より得られた開示レベルから，コミュニケーション頻度情報を得る．次に，友人ごとにユーザと友人のメッセージのやりとり等のコミュニケーション頻度から，自身と友人における紐帯の強度を算出し，友人の順位付けを行う．その際に，コミュニケーション頻度が少ないが公開対象者となりうるユーザが存在する可能性があるため，これまでの投稿履歴を考慮して，順位を決定する．最後に，開示レベルごとの人数上限に沿うように最終的な公開対象者を提案する．提案を受けたユーザは，基本機能のユーザ視点 1, 2 にあるように，提案された開示レベルおよび公開対象者を自由に変更が可能である．

Step 3.　SNS 投稿

ユーザから SNS 投稿依頼が行われた時点で開示レベルの範囲内に指定されたユーザのみに，投稿メッセージを限定公開する．また，公開の対象・対象外としたユーザ情報を保持し，次回以降に本システムを使用する際の初期値提案，および，Step 2 の公開範囲の導出時にも適用する．

4.3 システム実装

SNS ユーザの不用意な投稿によるセンシティブデータの漏洩を防ぐために，前節で提案した方式案をベースに，Facebook を対象としたアプリケーション : Adaptive Disclosure Controller for Facebook（以下，ADCF）を実装した．本アプリケーションは，SNS 投稿時の公開範囲に注目し，2 つの主要機能 : (1) 投稿内容に含まれるセンシティブデータの漏洩検知，(2) 投稿内容に応じた適応的な公開範囲の提案から構成される．

4.3.1 概要

ADCF のユーザインターフェースを図 4-2 に示す. 本アプリケーションは 2 つのボタン（Post ボタン，Detect ボタン），投稿メッセージを入力するためのテキストフィールド，開示レベルを選択するためのプルダウンリストを持つ．Post ボタンは，テキストフィールドに入力された投稿メッセージを，本アプリケーションで設定された公開範囲に基づき Facebook に投稿する．Detect ボタンは，投稿メッセージにセンシティブデータを含むか否かをアプリケーションに判断させるために使用し，SNS 投稿前に使用されることを期待している．画面右上のプルダウンリストは，開示レベル 1 : 親友 〜 5 : 他人の 5 レベルを持ち，選択した開示レベルにより，その投稿の公開範囲を制御する．また，画面中央に Dunbar’s circle として 5 つの円を表示する．最も内側の小さい円は自分自身を表現し，その他 4 つの円は，開示レベル 1 : 親友 〜 4 : 知人に応じて，段階的に異なる円のサイズで上位の開示レベルを包含するように表現する．最も外側のエリアは開示レベル 5 : 他人となる．また，各円の上に自身の Facebook 友人リストから抽出したユーザと，そのプロフィール写真を友人関係の紐付き強度に合わせて配置する．

4.3.2 コミュニケーション頻度による友人関係性の表現

図 4-2 は，ADCF の初期画面を表現している．Dunbar’s Circle の中央をユーザ自身とし，Facebook の友達リストから取得した各ユーザを，親友，友人，知人以上友人未満などの開示レベルごとの円内に配置する. Dunbar’s Circle 内におけるコミュニケーション頻度による友人分類は，実世界における友人分類と類似性があるとされる[81]．そこで，自身と各ユーザとの友人関係の紐付き強度を表現するために，Facebook におけるコメントやいいね！ボタン押下などのコミュニケーション頻度をもとに，各ユーザの配置位置を決定する．紐付き強度の算出には，自分自身によるコメント・いいね！ボタン押下（Outgoing コミュニケーション），他ユーザによる自身の投稿へのコメント・いいね！ボタン押下（Incoming コミュニケーション）の履歴とその頻度を利用する. さらに, 自身から積極的にコミュニケーションを取る Outgoing コミュニケーションは，他ユーザから受ける Incoming コミュニケーションよりも友人関係の

紐付き強度に与える重要性が高いと仮定し，より重みを与える．これにより，一方的にコミュニケーションを取ってくる他ユーザとの紐付き強度を抑える効果がある．同様に，気軽にコミュニケーションを行いやすい，いいね！ボタン押下よりもコメントの入力は，ユーザ間の紐付きが高いと仮定し，より重要性が高い行為と見なす．また，各円内に配置可能なユーザ数は，開示レベルの基準人数を上限とし，基準人数を超える場合は，算出した紐付き強度を基準に下位の開示レベルに含めるものとする．例えば，紐付き強度を算出し，開示レベル 1：親友（基準人数 1〜5 人）に 6 ユーザが含まれた場合，これらユーザの中で最も強度の低いユーザを開示レベル 2：友人に属するものとする．これらの処理手順を最も外側の開示レベル 5 まで繰り返して実行し，最終的な各ユーザの配置を決定する．

図 4-2　Adaptive Disclosure Controller for Facebook

4.3.3 任意の公開範囲・公開対象者の選択

画面右上に配置される開示レベルのリストボックスより，任意の開示レベルを選択することで，容易に投稿の公開範囲を拡大・縮小の設定をすることができる．この開示レベルの選択により，Dunbar's Circle の縁に沿い，赤ラインでその投稿範囲を視覚的に表現する．開示レベル 4：知人を選択後の画面を図 4-3-(a)に示す．

また，Dunbar's Circle 上に配置された各友人の画像をドラッグ＆ドロップすることにより，開示レベルの内外へ友人を容易に再配置することができる．この機能により，投稿内容によって，任意の公開対象者をアドホックに選択することができる．また，各友人の画像にマウスオーバー時に友人の名前を動的表示することで，公開対象者とするか否かの判断を容易にする．図 4.3-(b)に開示レベル 2：友だちから，開示レベル 3：知人以上友だち未満へ再配置中の画面を示す．

(a) 任意の開示レベルの選択

(b) 任意の公開対象者の選択

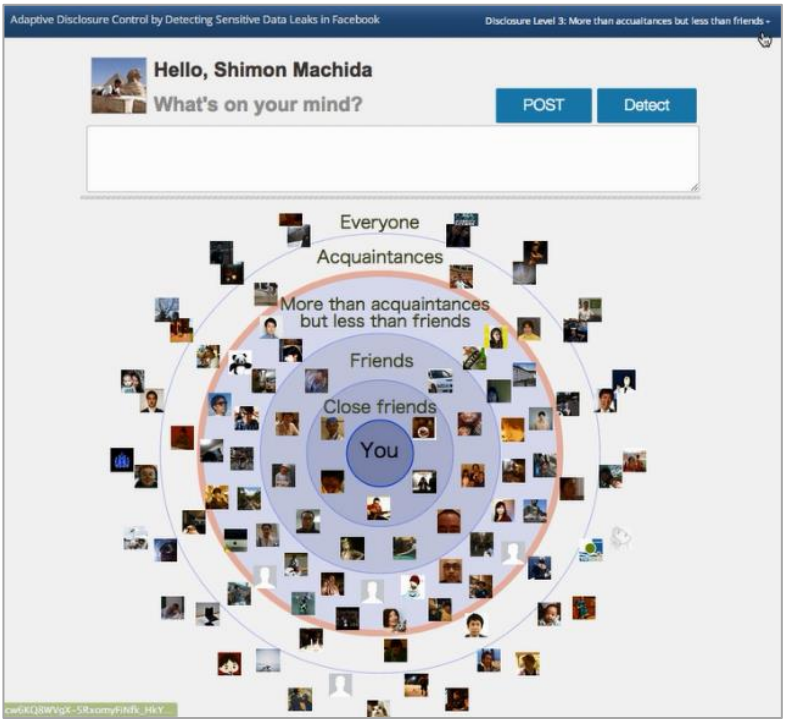
図 4-3　公開範囲・公開対象者の選択

4.3.4 センシティブデータ検知に応じた開示レベルの提案

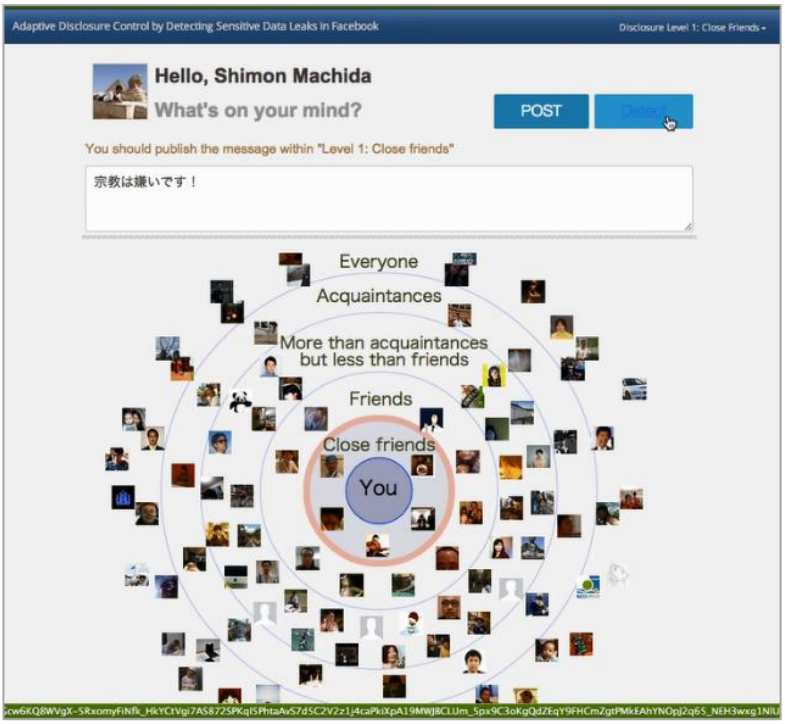
メッセージに含まれるセンシティブデータを検知し，その情報の重要性に応じた開示レベルを提案するために，ADCF では事前に生成済みの分類器を通じて，キーワード分析，感情分析を行う．これらの分類器は，3 章で示した過去 1 年間分の 1 億 4,000 万件の Twitter アーカイブをベースに実施した SNS におけるセンシティブデータ漏洩の試算評価時の結果データを教師データに用い, Support Vector Machine で実装した. ユーザが Detect ボタンを押下時に，分析ごとの分類器によりチェックを行い，学習データに沿った開示レベルが提案される．なお，キーワード分析，感情分析の結果，異なる開示レベルとなった場合は，極力安全策を取るようにするため，上位の開示レベルを採用することとした.

ユーザが Detect ボタンを押下時，メッセージに含まれるセンシティブデータを前述の分類器を通して検知し，その開示レベルと警告メッセージをユーザに提示する．例えば，メッセージから個人の宗教観，病気・疾患（重度），非常識な行動傾向を検知した場合，最も狭い開示範囲となる開示レベル 1 を提案する．提案される開示レベルは，画面上の当該開示レベルに対応した Dumber’s Circle の縁に沿い, 赤ラインでその投稿の公開範囲を視覚的に表現する. その結果, ユーザは SNS 投稿をする前にメッセージにセンシティブデータを含んでいること，その公開範囲に含まれる友人を容易に認識することができる．ADCF からの提案後，ユーザはメッセージの修正，前項の機能により，開示レベルの再選択，開示対象ユーザとなる友人を選択できる．例えば，ユーザが投稿の開示レベルを開示レベル 3: : 知人以上友だち未満と選択した上で，メッセージに「宗教は嫌いです！」と入力・投稿しようとする場合，投稿前に Detect ボタンを押下してもらう．その結果，ADCF はメッセージにセンシティブデータが含まれていることを検知し，開示レベル 1 : 親友を提案する．本例を図 4-4 に示す.

また，先行研究[16]において，SACL 作成時のユーザの作業負荷を軽減するためには，過去に使用した設定履歴を用いることが最も有効な手法とした．そこで，ADCF では，最終的にユーザが投稿した開示レベルと Dunbar’s Circle 上の各友人の配置位置を記録し，次回アプリケーション利用時の初期提案値として利用する.



(a) ユーザ：開示レベル 3 を手動選択



(b) システム：開示レベル 1 を自動提案

図 4-4　センシティブデータ漏洩の検知と開示レベルの提案

4.4 まとめ

本章では，3 章で提案・評価した SNS におけるプライバシ侵害情報分類表を基に，SNS 投稿前のメッセージに含まれるセンシティブデータ有無の検知と，検知した情報の重要度に応じた開示レベルを公開範囲として通知・設定する方式を提案した．そして，本提案方式の実現化として，Facebook を対象としたアプリケーション：Adaptive Discosure Controller for Facebook を実装した．本アプリケーションでは，センシティブデータ有無の検知機能のみでなく，コミュニケーション頻度を基に算出した友人関係性を Dunbar's Circle 上に可視化し，投稿ユーザに当該投稿の公開対象者を視認できるようにし，注意を促すようにした．さらに，SNS ユーザが要望する投稿内容に応じた特定の個人や興味を持つコミュニティへのメッセージ公開を実現するために，各友人の画像をドラッグ＆ドロップすることで，開示レベルの内外へ友人の再配置を可能とした．これにより，コミュニティやグループにとらわれることなく，投稿内容に応じた特定の個人へ投稿を容易に共有することができる．

第5章　被写体のコミュニティベース・プライバシポリシの設定方式

SNS 投稿ユーザには，写真の被写体のプライバシポリシが分からないため，自身の主観的な判断基準のもとに投稿を行い，その結果，被写体のセンシティブデータ漏洩が発生している現状がある．そこで本章では，写真の被写体のプライバシを保護するために，当該人物が属するコミュニティ内外におけるプライバシの振る舞いをポリシとして埋め込んだタグを用いて，コミュニティ内外で当該人物の顔領域を適応的に保護する手法を提案する．本手法を適用することにより，SNS 投稿ユーザ，または写真の撮影者の主観的な判断基準のみに依存せずに，被写体のプライバシポリシに基づいたプライバシ保護が可能となる．

5.1 被写体のプライバシ保護

5.1.1 SNS 投稿写真に写る人物のプライバシ保護

SNS 提供者側では，ユーザのセンシティブデータの漏洩を防ぐために，様々なプライバシ設定機能を提供している．しかし，プライバシ管理は複雑性があるだけでなく，その維持に多くの労力が必要となる[16]．そのため，投稿の公開範囲提案などの情報の開示境界に関する研究[16, 39, 82]を中心に活発に研究が行われている．しかし，多くの従来研究は，投稿ユーザ自身のプライバシ保護を主目的としており，投稿内容に含まれる他ユーザのプライバシの注目まで至っていない．実際に，SNS ユーザの投稿後の後悔事例[24, 29, 32]において，個人特定可能な写真を無断投稿した結果のトラブルが報告されているように，SNS へ投稿する写真に写る/写り込む人物を含む被写体のプライバシを考慮する必要がある．そこで，投稿ユーザの判断基準のみに依存したプライバシ保護ではなく，被写体のプライバシポリシを反映させる手法を検討する．

5.1.2 被写体のプライバシポリシ適用

SNS 投稿時に被写体のプライバシポリシを反映させるためには，被写体のポリシを投稿ユーザに参照させる必要がある．その際，考え得る手法としては，(1) 顔認識を用いる手法，(2) 電波認識を用いる手法，(3) タグ認識を用いる手法が挙げられる．顔認識を用いる手法では，予め顔の特徴量とプライバシポリシを紐付け登録しておき，他ユーザによる写真投稿時に顔認識を行い，被写体のポリシを適用・通知を行う[28, 53]．しかし，これらの手法の場合，

タグを常に身に付ける煩雑さはないものの，システム上に顔の特徴量等の人体特性を登録する必要があるため，プライバシ懸念を理由にユーザから利用を拒否される可能性がある[54, 55]．次に，電波認識を用いる手法では，プライバシポリシが埋め込まれた RFID タグを身につけ，電波を介してポリシ情報を読み取り，そのポリシに応じて非特定化処理を行う手法[56, 57]が提案されている．しかし，RFID タグを認識するために，デジタルカメラなどのデバイスに，統一的な仕様のもとに組み込む必要があることが課題となる．最後に，タグ認識を用いる手法では，自身のポリシを埋め込んだタグを身に付け，投稿時にタグ解析を行い，被写体のポリシを適用する．従来研究では，機械のみでなく，人間も容易に判断可能な単純化されたタグを身に付け，周囲にアピールする手法[58]や，QR コードに多くのポリシを埋め込み，プライバシをコントロールする手法[25, 60]が提案されている．しかし，単純化されたタグは多くの情報を保持できないため，コミュニティに応じたプライバシの振る舞いを表現できない．また，QR コードのように多くの情報を含む複雑なタグは，撮影者と被写体までの距離によっては検知・解析精度に課題が残る．

5.1.3 プライバシアピール

タグ認識を用いる手法で，ポリシを含むタグを身に付ける場合，自身のポリシを周囲にアピールすることとなる．自身のポリシもセンシティブデータの一つと捉え，隠すべきとする考え方があるが，撮影者にアピールをすることで，被写体のポリシ尊重を促す効果がある[53]．そこで，本提案では，システムに人体特性の登録が不要であるタグ認識を用いる手法を採用し，投稿時に被写体が身に付けたタグから取得したポリシを適用させることとした．

5.2 被写体のコミュニティベース・プライバシポリシの設定方式

写真に写る/写り込む人物のプライバシを保護するために，当該人物が属するコミュニティ内外におけるプライバシの振る舞いをポリシとして埋め込んだタグを用いて，コミュニティ内外で当該人物の顔領域を適応的に保護する手法を提案する．本手法を適用すると，SNS へ写真投稿時に，写真に写る人物が身に付けたタグを検知・解析し，その結果得られるポリシに基づき，当該人物の顔領域をぼかして非特定化する．また，タグに含まれるコミュニティ情報を用いて，予め，ユーザが属するコミュニティの定義と，そのコミュニティに属するその他ユーザを紐付ける．これにより，ユーザはコミュニティごとに異なるポリシを持つタグを身に付けることで，状況に応じたプライバシの振る舞いが可能となる．さらに，コミュニティの所属ユーザを投稿の公開対象者として利用することで，投稿メッセージの公開範囲を限定し，センシティブデータの漏洩を防ぐ．

本提案方式は，(1) PrivacyTag，(2) Photo Privacy Realizer，(3) Privacy Wall の 3 つの主要機能か

ら構成される．図 5-1 に本方式の概略フローを示す．次節では，これら機能の説明，提案方式の流れを説明する．

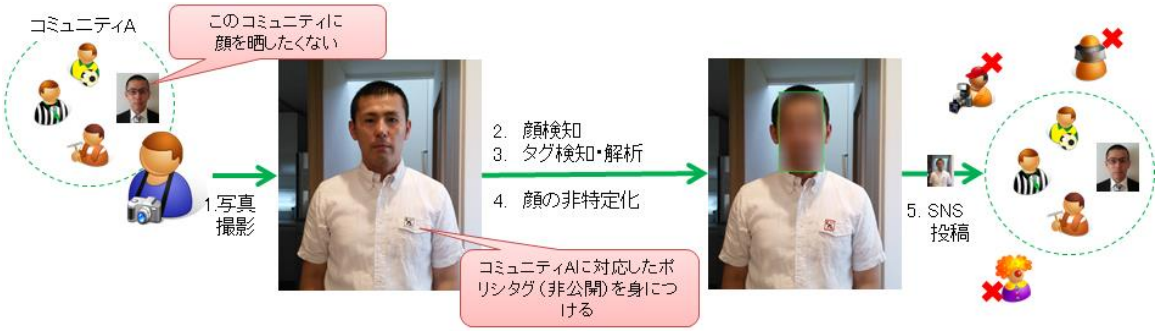


図 5-1　概略フロー：被写体のコミュニティベース・プライバシポリシの設定方式

5.2.1 PrivacyTag

PrivacyTag とは，インターネット上で自身の情報を公開/非公開といったプライバシの振る舞いと，その振る舞いを適用するコミュニティ情報をプライバシポリシとしてシンボル化したタグである．ユーザは，衣服やアクセサリ等，ファッションの一部として本タグを身に付け，周囲に自身のプライバシポリシをアピールする．プライバシタグのデザイン，解析アルゴリズム，および評価については，5.3 節で説明する．

5.2.2 Photo Privacy Realizer

Photo Privacy Realizer (PPR) は，SNS 投稿時に，写真から人物の検知，当該人物が身に付けるプライバシタグの検知および，解析を行い，取得したポリシに基づき，その人物の顔領域の保護を実現する Web アプリケーションである．ユーザがスマートフォン等から利用することを想定しており，(1) コミュニティ管理，(2) 写真撮影および非特定化 の 2 機能から構成される．PPR は，取得したポリシに従い，被写体の顔領域にぼかしを入れて非特定化する．そして，非特定化後の写真とメッセージをコミュニティメンバのみに限定公開する．5.4 節では，これら機能について，対象 SNS を Facebook として実装した PPR を説明する．

5.2.3 Privacy Wall

Privacy Wall とは，日常で使用されるデジタルカメラ等の PPR 以外のデバイスで撮影・投稿された際に，プライバシタグを身に付けた被写体のプライバシ保護を行う機能である．本機能は，SNS 提供者の機能の一部として構築され，投稿時にタグ検知のみを行い，タグを身に付けた全被写体の顔領域を非特定化する．なお，本論文では機能の提案のみとし，今後検討を行う．

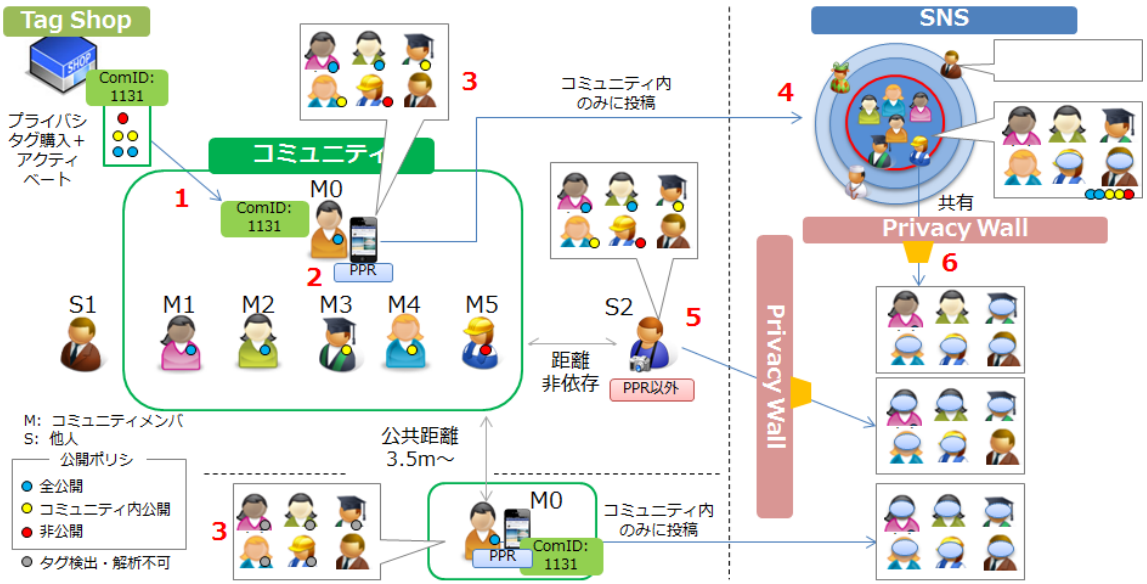


図 5-2　提案方式のプロセスフロー

5.2.4 提案方式の流れ

提案方式のプロセスフローを図 5-2 に示す．図中の番号は，下記の流れと対応している．

1.　タグの入手

タグは，タグを取り扱う店から，共通のコミュニティ ID を含む全公開/コミュニティ内公開/非公開を示す 3 種のタグを束ねたタグパックとして購入，また，ユーザ自身によるタグのプリントが可能とする．

2.　タグのアクティベートとコミュニティ登録

タグを入手後，PPR を利用して，タグのアクティベートと新たなコミュニティを登録する．また，コミュニティの所属メンバを SNS の友人リストから選択して紐付ける．各コミュニティメンバは，自身のポリシに合致するタグを身に付ける．

3.　写真撮影と非特定化

PPR を利用して写真を撮影し，被写体が身に付けたタグを解析する．この時，被写体のポリシが“非公開”の場合は，顔領域にぼかしを入れる．また，当該写真にコミュニティメンバ外の他人が写り込んでいる場合，タグを身に付けていないため，ポリシを取得できないが，プライバシを尊重して非特定化する．同様に，被写体からの距離等が原因により，タグを検知したが，解析ができない場合も当該被写体を非特定化する．

4.　SNS 投稿

保護処理後の写真にメッセージを添え，SNS へ投稿する．その際，当該コミュニティのメンバを対象に限定公開する．

5.　PPR 以外のデバイスによる撮影・投稿

コミュニティメンバ外の他人が PPR 以外のデバイス・アプリケーションで撮影・投稿したとする．この場合，SNS 提供者側の Privacy Wall により，タグ検知のみを行い，タグを身に付けた全被写体の顔領域を非特定化する．

6.　コミュニティ外へ再投稿

既に投稿されたメッセージの共有など，コミュニティ外に再投稿される場合は，写真のメタ情報からポリシを取得し，そのポリシに基づき非特定化する．

5.2.5 撮影者と距離による保護の対応

被写体の顔領域の保護は，(1) 写真の撮影者と被写体の距離，(2) 撮影および SNS への投稿が PPR，またはデジタルカメラなどのその他デバイス・アプリケーションにより行われたかに依存する．

写真の撮影者と被写体の距離について検討する．Instagram を対象とした投稿写真の分類[83]では，ユーザによって投稿された写真は，自撮り・食べ物，ペットなどの 8 カテゴリに分類される．このうち，投稿者自身の自撮り，友人など 2 人以上と写った写真が 46.6%を占める結果であった．このように撮影・SNS へ投稿された写真における撮影者と被写体の距離を計るにあたり，自身と他者との関係性を示すパーソナルスペースの分類[84]を参照した．本分類では，人が相手や状況により取るスペースを 4 分類で示している．(表 5-1 を参照) 先に挙げた投稿写真の分類では，自撮り，友人など 2 人以上と写った写真が多い傾向である事から，SNS に投稿される写真の多くは密接距離～社会距離であると仮定し，被写体のプライバシ保護範囲も同様の距離を対象にすることとする．よって，被写体が身に付けたプライバシタグの検知・解析は，撮影者から 350 cm まで正確に行える必要がある．

表 5-1　Hall によるパーソナルスペースの分類

分類	説明	距離
密接距離	非常に親しい人に許可される距離	0 – 45 cm
個体距離	友人などとの会話で取る距離	45 – 120 cm
社会距離	他人との会話で取る距離	120 – 350 cm
公共距離	相手と公的な関係性である時に取る距離	350 – 750 cm

図 5-3 に，撮影者（コミュニティメンバ/他人）と距離（密接距離～社会距離，公共距離～）による顔領域の保護処理の対応を示す．コミュニティのメンバが M0～M5 の 6 名存在する．また，コミュニティメンバ外の他人として S1，S2 がいる．このうち，M0 と S2 が撮影者となる．M0 による撮影は，社会距離以内であれば，被写体のポリシに応じて顔領域を保護する．しかし，S2 によるその他デバイス・アプリケーションを用いた撮影・投稿は，SNS 提供者側の Privacy Wall により，撮影距離に依存することなく，タグを身に付けたユーザの顔領域を保護する．このように本手法では，ユーザのポリシが反映できないケースに対して，極力安全策を取るようにする．

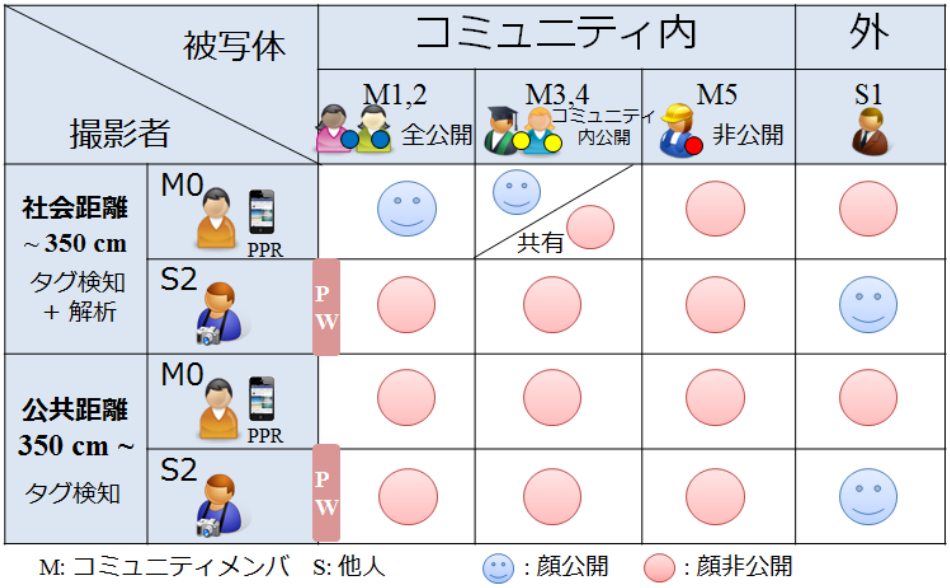


図 5-3　撮影者と距離による保護処理の対応

5.3 プライバシタグ

我々が提案するプライバシタグのデザインと検知・解析アルゴリズムを説明する．また，従来手法で用いられる QR コードベースのタグとの比較，評価を行い，課題であった被写体までの距離による検知・解析精度の改善を示す．

5.3.1 予備評価

タグデザインするにあたり，タグの検知・分析手法を検討した．最初に従来手法で用いられる QR コードにおいて，距離課題があることを確認した．（図 5-4 参照）タグの検知手法と

して，タグ全体を囲うフレームを設け，最初にこのフレームを検知し，次にフレーム内のビットパターンを読み込み，プライバシタグであるかの判断を行うと仮定した．この仮定のもと，(1) フレーム線幅と距離による検知精度，(2) 距離によるビットパターンの読み取り精度の予備評価を行った．両評価とも 5 cm 四方のタグを用い，屋外で約 2090 万画素のデジタルカメラを利用して撮影し，フレームまたはビットの輪郭を捉えた場合に成功と見なした．評価の結果として，図 5-5 にフレーム線幅（1 mm〜5 mm）と距離（1.5 m ごとに 12 m まで）による検知結果を示す．フレーム線幅 2 mm 以下では，撮影者との距離が遠くなるにつれ，輪郭を検知できないが，3 mm 以上の場合，12 m まで検知可能であることが分かった．次に，1 辺 7 ビットを配置したビットパターンの読み込み結果を図 5-6 に示す．撮影者からの距離が 6 m 以下の場合，概ねビット読み込みができることが分かった．本結果をもとに提案タグのデザインを行った．

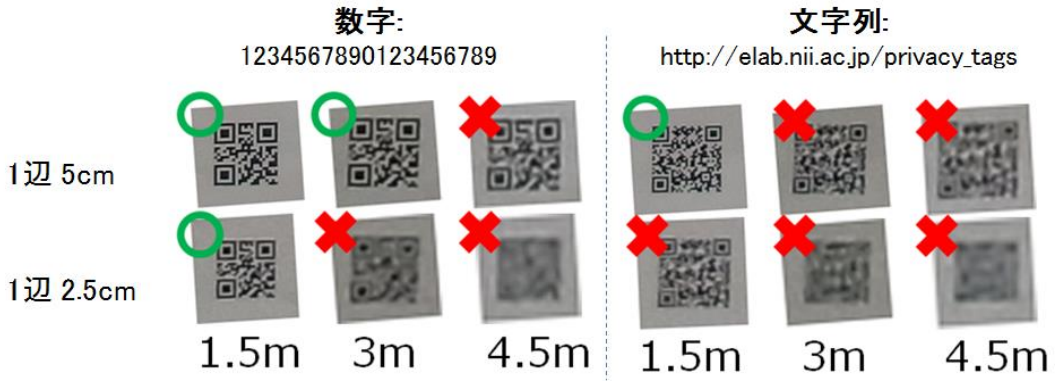


図 5-4　従来手法におけるタグサイズと距離による検知評価

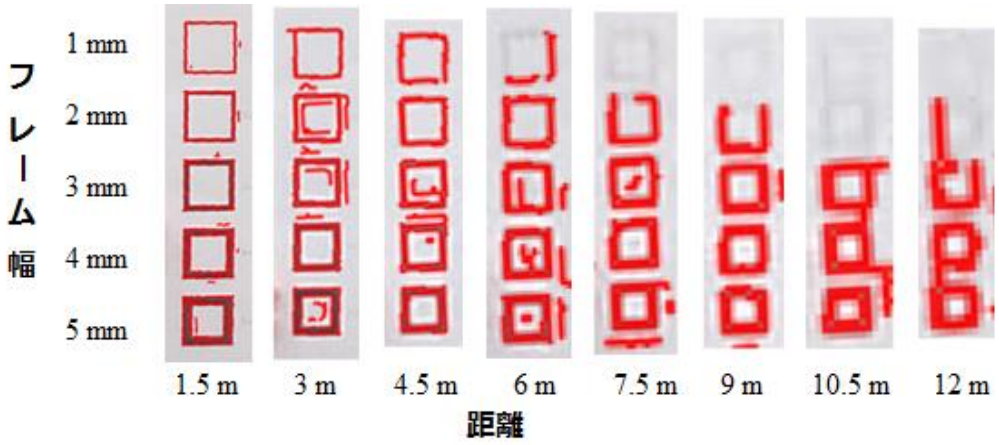


図 5-5　フレーム線幅と距離による輪郭の検知（サイズ：5 cm 四方）

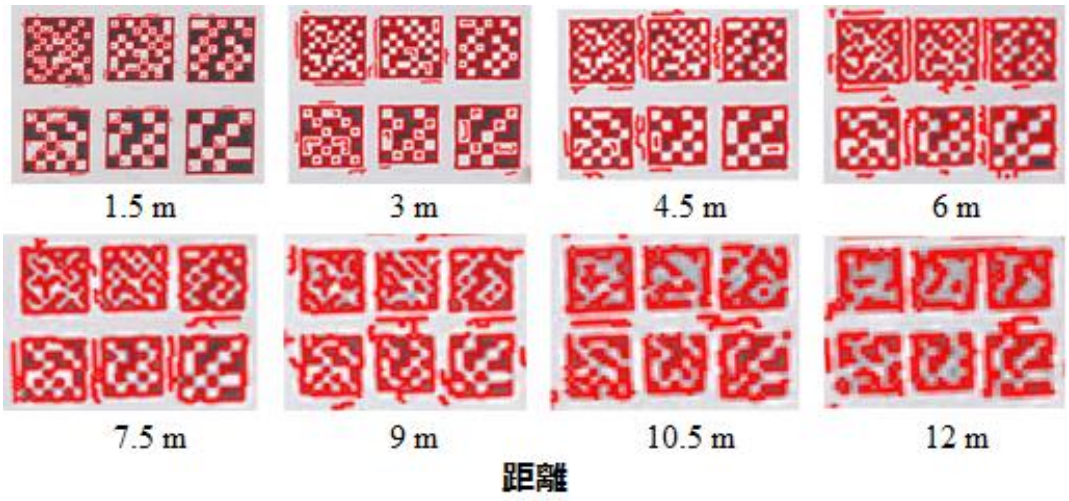


図 5-6　距離ごとのビットパターン読み込み（サイズ：5 cm 四方）

5.3.2 デザイン：公開ポリシ

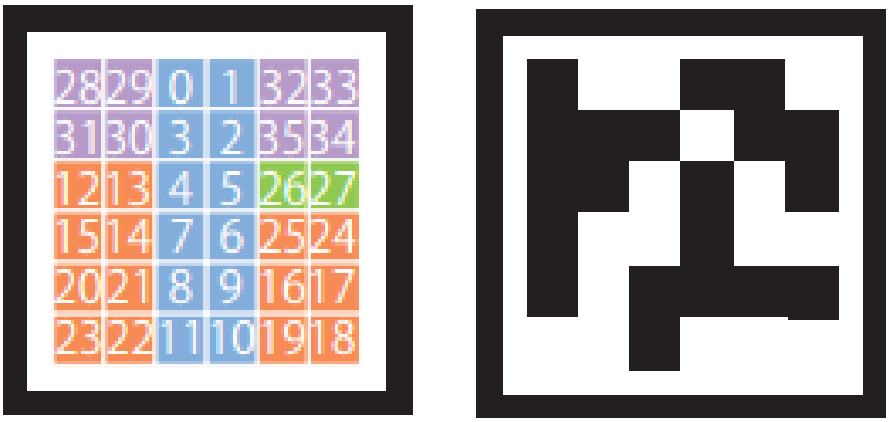
写真のセンシティブデータ漏洩を防ぐために, 複雑な設定は必要なく, 顔情報やタグ付け, 位置情報等の公開・非公開を表す数ビットで対応可能である[85]. 我々の提案方式は，コミュニティを考慮した被写体の顔領域保護を行うため，公開ポリシとして，(1) 非公開，(2) コミュニティ内公開, (3) 全公開の 3 種を定義した. 表 5-2 に提案タグで用いる公開ポリシを示す.

表 5-2　提案タグで用いる公開ポリシ

公開ポリシ	説明
非公開	コミュニティ内外問わず，自身の顔領域を非公開
コミュニティ内公開	特定コミュニティ内であれば，自身の顔領域を公開可能
全公開	コミュニティに関わらず，自身の顔領域を公開可能

5.3.3 デザイン：ビットパターン

定義した公開ポリシや適用対象となるコミュニティ情報を，タグ内にビットとして配置するため，タグ内のビットパターンを検討した．本タグ内のビットパターンとタグサンプルを図 5-7 に示す.



(a) ビットパターン　　(b) PrivacyTagの例

図 5-7　プライバシタグ

タグに求められる要件として，タグの向きに依存しないビット可読性，バースト誤りを考慮したエラー訂正がある．これらに対応するため，タグ上でバースト誤りが発生しやすい領域を，左右・上部→左右・下部→中央の順であると仮定した．この仮定のもと，タグ検知および，タグの向きを判定するために，位置決めビットパターンをヘッダ部として中央に配置した（図 5-7-(a)，bit: 0-11）．次に，被写体のプライバシポリシとなるコミュニティ ID と公開ポリシを左右・下部へ配置した（図 5-7-(a)，bit: 12-25, 26-27）．誤り訂正符号には，リード・ソロモン符号を採用し，1 シンボル=4 ビットごとにエラー訂正が可能である．ただし，ヘッダ部はエラー訂正の対象外とした．このエラー訂正符号を左右・上部に配置した（図 5-7-(a)，bit: 28-34）．

5.3.4 タグ検知と解析

写真から被写体が身に付けたタグの検知・解析フローを図 5-8 に示す．検知・解析フローは，顔検知とタグ検知・解析を並列に実行する．顔検知は，顔検出アルゴリズム: Viola-Jones 法[86]をベースとした手法を適用した．次に，タグ検知・解析は，最初にタグのフレーム境界を検出する．次に，タグの傾き等への対応として画像補正を行い，タグ内のビットパターンを読み取る．最後に，位置決めパターンであるヘッダ部とのマッチングにより，プライバシタグであるかを判断する．なお，読み取り不可ビットがある場合は，エラー訂正符号を用い，ビットを補完する．

タグ検知・解析後，タグの持ち主となる被写体を判別する．被写体は当該タグを上半身に身に付けている，かつ，検知した顔の幅(W)，高さ(H)を用い，顔直下の 3W×4H の領域にタグが存在すると仮定して，マッチング処理を行う．最後に，取得したポリシに従い，顔領域にぼかし処理を行う．

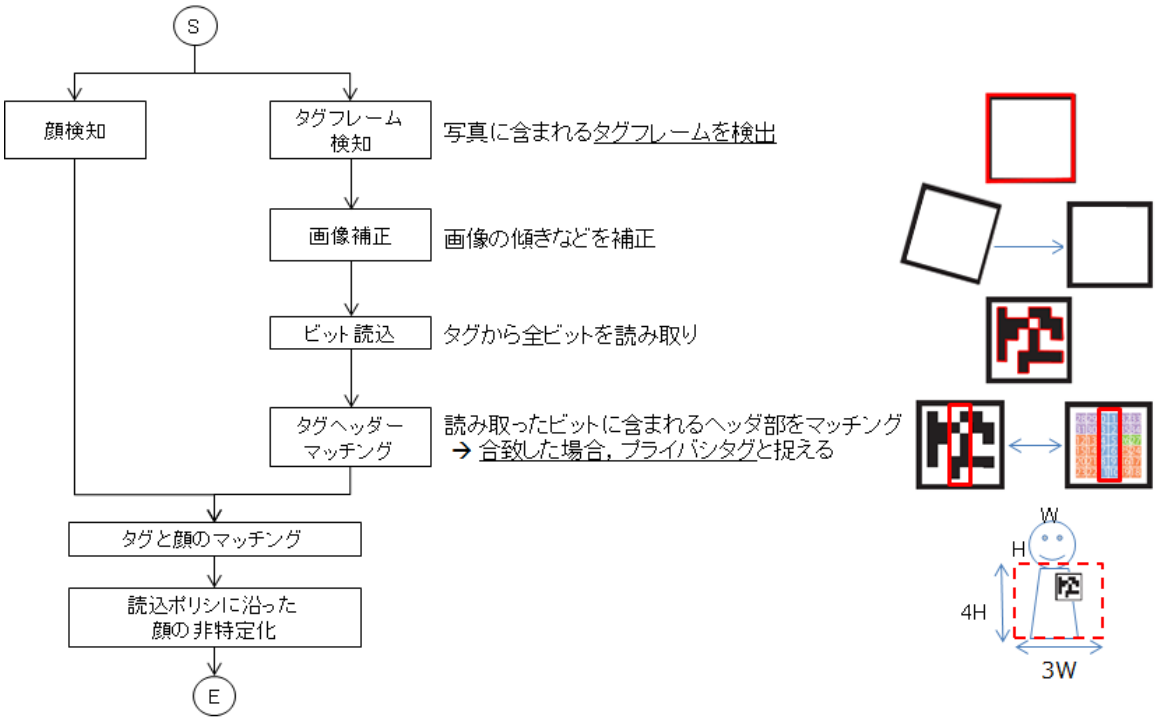


図 5-8　タグの検知・解析フロー

提案タグ

サイズ/距離	45 cm	82.5 cm	120 cm	235 cm	350 cm	450 cm	550 cm	650 cm	750 cm	850 cm	950 cm	1050 cm
2 cm	✓	✓	✓	✓	✓	NA	NA	NA	NA	NA	NA	NA
3 cm	✓	✓	✓	✓	✓	✓	NA	NA	NA	NA	NA	NA
5 cm	✓	✓	✓	✓	✓	✓	NA	NA	NA	NA	NA	NA
7.5 cm	✓	✓	✓	✓	✓	✓	✠	NA	NA	NA	NA	NA
10 cm	✓	✓	✓	✓	✓	✓	✓	NA	NA	NA	NA	NA

密接距離　個体距離　社会距離　公共距離

✔ - タグ検知 & 解析
✠ - タグ検知
NA - 検知不可

QRコードタグ

サイズ/距離	45 cm	82.5 cm	120 cm	235 cm	350 cm	450 cm	550 cm	650 cm	750 cm	850 cm	950 cm	1050 cm
2 cm	✓	✓	✓	NA	NA	NA	NA	NA	NA	NA	NA	NA
3 cm	✓	✓	✓	✓	NA	NA	NA	NA	NA	NA	NA	NA
5 cm	✓	✓	✓	✓	✓	NA	NA	NA	NA	NA	NA	NA
7.5 cm	✓	✓	✓	✓	✓	✓	✓	NA	NA	NA	NA	NA
10 cm	✓	✓	✓	✓	✓	✓	✓	✓	NA	NA	NA	NA

図 5-9　提案タグと QR コードベースタグの評価結果

5.3.5 評価

従来手法で用いられる QR コードのタグと提案タグの比較評価の第一歩として，各タグにおいて，タグサイズ：2 cm〜10 cm かつ，被写体までの距離：45 cm〜1050 cm の検知・解析精度を評価した．なお，QR コードはバージョン 1，評価実験は屋外で約 2090 万画素のデジタルカメラを利用して撮影した. 図 5-10 に提案タグと QR コードベースタグの評価実験の例，図 5-9 に実験結果を示す．評価の結果，実生活において許容されると考えられる 5 cm 以内のサイズにおいて，検知・解析可能な距離が向上した．タグサイズ：2〜5cm において，QR コードは距離：120 cm〜350 cm までが読み取り上限であったが，提案タグは，450 cm まで読み取り可能であった．図 5-10-(b)は，サイズ：3 cm，距離：350 cm において，QR コードは各ビットが暈けており，読み取ることができなかったが，提案タグはタグの検知とビット解析が可能であった．被写体のプライバシ保護範囲として，密接距離〜社会距離を対象とすると仮定したが，評価の結果，社会距離（350 cm）までは，実用的なタグサイズにおいて検知・解析が行えることが分かった． なお，タグサイズ：10 cm の着用例から，実社会において実用的なサイズとして 5 cm 以下であることが判断できる．（図 5-10-(a)を参照）

(a) サイズ: 10 cm, 距離: 350 cm

(b) サイズ: 3 cm, 距離: 350 cm

図 5-10　提案タグと QR コードベースタグの着用例

5.4 システム設計

5.4.1 概要

プライバシタグを身に付けた人物の顔領域をポリシに応じて保護するために，5.3 節で提案したタグデザイン，検知・解析フローをベースに，アプリケーション：Photo Privacy Realizer を設計した．PPR は，(1) コミュニティ管理，(2) 被写体の非特定化の 2 機能から構成される．本アプリケーションでは，最初に投稿ユーザはプライバシタグに含まれるコミュニティコードをアクティベートし，SNS の友人リストをベースに当該コミュニティに属するコミュニティメンバを管理する．次に，写真撮影を行い，SNS へ投稿する前に本アプリケーションを用いて，写真に対して被写体のプライバシポリシを反映させる．そして最後に事前に定義したコミュニティメンバのみを投稿の公開範囲として，SNS へ公開する．これにより，投稿ユーザは自身の主観的な投稿判断のみに依存することなく，写真に写る被写体のプライバシポリシを考慮した結果の写真投稿が可能となる．

5.4.2 処理フロー

本アプリケーションのプロセスフローを図 5-11 に示す．本アプリケーションは大きく分けて，コミュニティ管理，被写体の非特定化の 2 プロセスから構成される．以下では，そのプロセスを説明する．

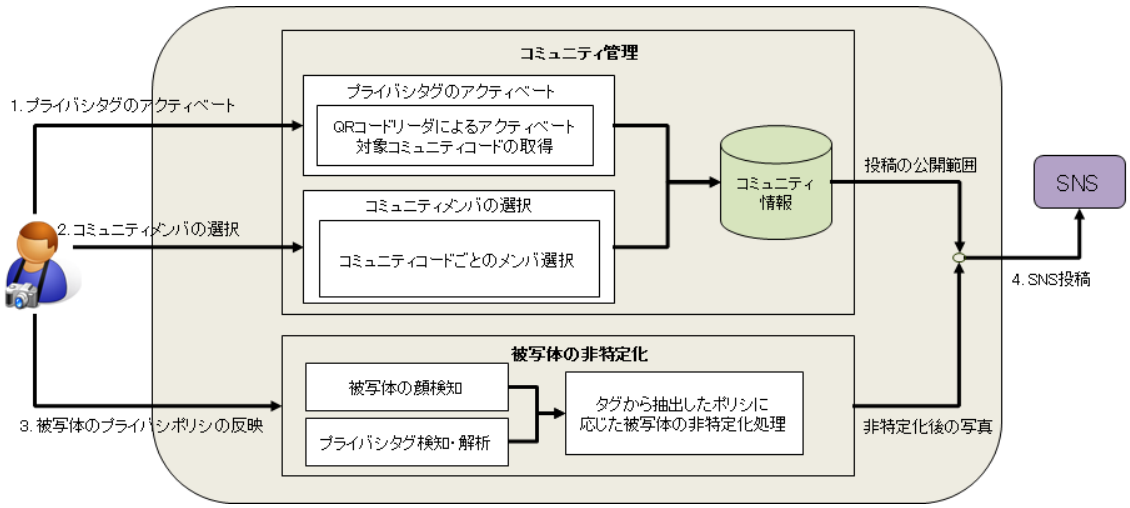


図 5-11　Photo Privacy Realizer のプロセスフロー

Step 1. プライバシタグのアクティベート

プライバシタグをパック化したタグパックに付帯する QR コードから，QR コードリーダな

どを用いてアクティベート対象のコミュニティコードを取得する．その際，当該プライバシタグの有効期限が設定される．

Step 2. コミュニティメンバの選択

SNS で定義済みの友人リスト情報を取得し，Step 1 で取得したコミュニティコードに属するコミュニティメンバを決定し，コミュニティとの紐付けをコミュニティ情報として保持する．

Step 3. 被写体のプライバシポリシの反映

投稿ユーザは，本アプリケーションを利用して撮影した写真，または事前に撮影済みの写真に対して，当該写真に写る被写体のプライバシポリシを反映する．その際，本アプリケーションでは，5.3.4 項で説明した検知・解析フローを用いて，被写体の顔検知および，プライバシタグの検知・解析を行う．次に，取得したプライバシポリシに応じて，被写体の顔領域に非特定化処理を行う．なお，被写体の顔検知のみ，プライバシタグの検知のみできる場合は, 5.2.5 項の撮影者と距離による保護に従って, 非特定化処理を行うこととする．

Step 4. SNS 投稿

Step 3 で非特定化処理後の写真を SNS に投稿するが，その際，Step 2 で管理したコミュニティ情報から当該コミュニティに属するコミュニティメンバを抽出し，投稿の公開範囲として利用する．

5.5 システム実装

前節のシステム設計をもとに，対象 SNS に Facebook を対象として，アプリケーション : Photo Privacy Realizer for Facebook を実装した．本アプリケーションのユーザインターフェースを図 5-12 に示す．

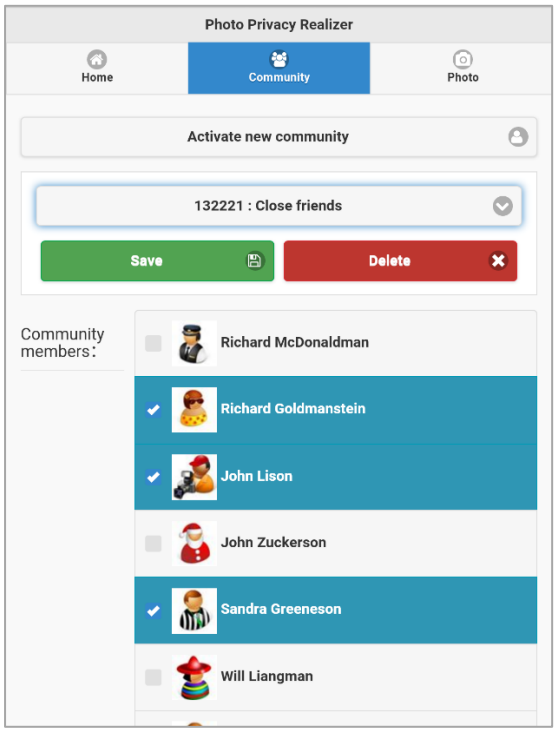
5.5.1 コミュニティ管理

コミュニティ管理機能を図 5-12-(a)に示す．本機能では，コミュニティのアクティベート，コミュニティに属するユーザの決定を行う．最初に Facebook へログインが必要となる．ログイン後，自身の Facebook 上の友人が一覧で表示される．次に，コミュニティのアクティベートを行うために，Activate new community ボタンを押下し，予めインストール済みの QR コードリーダを用いて，購入したタグパックに付帯する QR コードからコミュニティ ID を登録する．（図 5-12-(b)を参照. ）そして，登録したコミュニティ ID に対して，属するユーザを画面

タップで選択/解除を行い，最後に，コミュニティ情報を保存する．また，コミュニティのプルダウンリストより，登録済みのコミュニティを切り替えることで，コミュニティに属する友人のメンテナンスが可能となる．

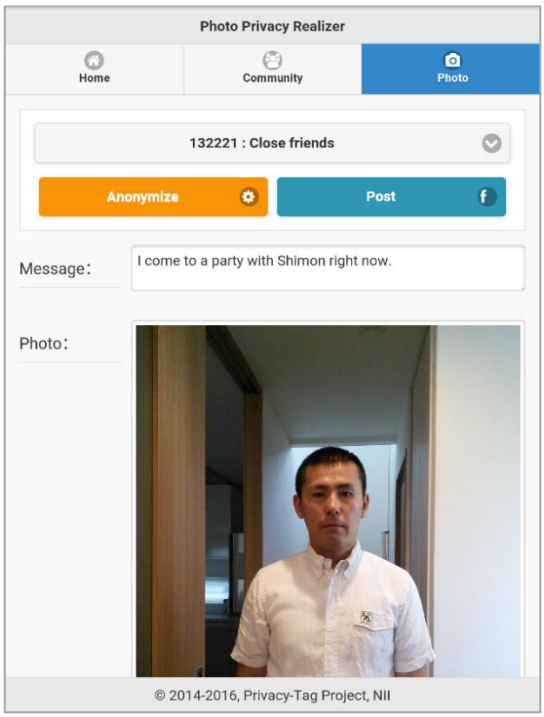
5.5.2 写真撮影・非特定化

本機能では新たに写真撮影，または既に撮影済み写真に対して，顔検知および，タグ検知・解析を行い，取得したポリシと PPR で指定されたコミュニティに応じて顔領域にぼかし処理を行う．最初に投稿メッセージの入力，写真撮影を行う．（図 5-12-(c)を参照．）この時点では，まだタグ検知・解析は行われていない．次に，Anonymize ボタンを押下し，撮影された写真に含まれるタグの解析を行い，その結果に応じた非特定化を行う．（図 5-12-(d)を参照．）また，タグのコミュニティが合致しない場合も，非特定化を行う．最後に Post ボタンを押下し，SNS へメッセージと保護処理後の写真を投稿し，コミュニティメンバのみに限定公開する．（図 5-12-(e)を参照．）なお，顔の非特定化後の緑枠，プライバシタグ検知後の赤枠は，処理結果の分かりやすさのために表示しており，実運用時には非表示となる．

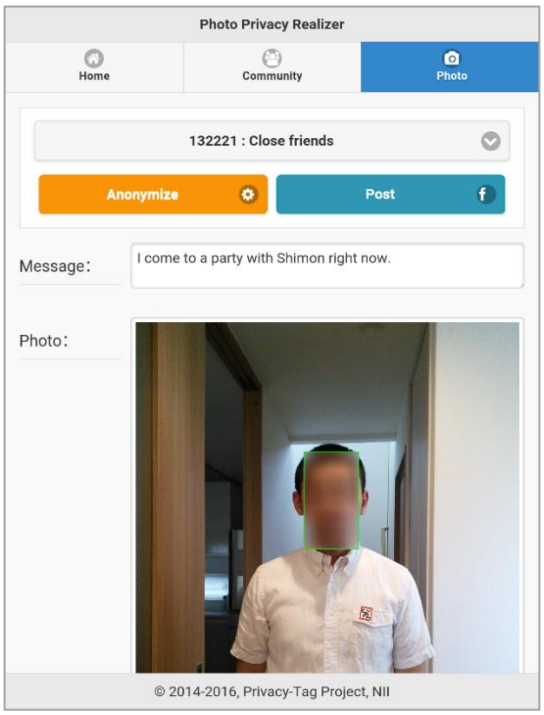


(a) コミュニティ管理

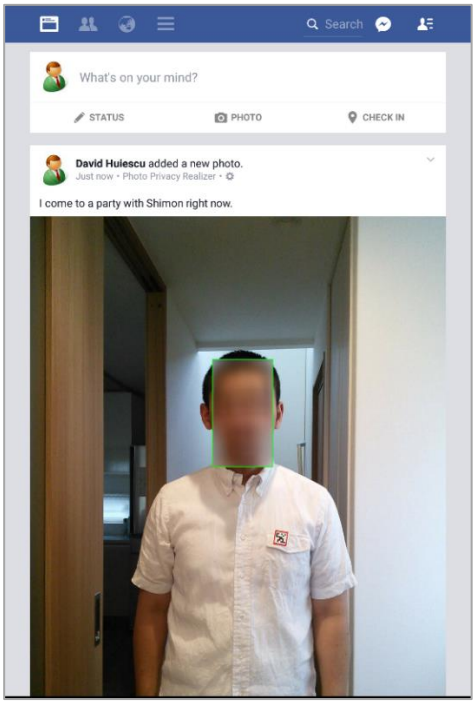
(b) プライバシタグのアクティベート



(c) 写真撮影



(d) 顔領域の保護



(e) コミュニティメンバのみへ投稿公開

図 5-12　操作フロー

5.6 まとめ

本章では，SNS 投稿ユーザ，または写真の撮影者の主観的な判断基準のみに依存せずに，写真に写る/写り込む人物のプライバシポリシに基づいたプライバシを保護するために，当該人物が属するコミュニティ内外におけるプライバシの振舞いをポリシとして埋め込んだタグを用いて，コミュニティ内外で当該人物の顔領域を適応的に保護する手法を提案した．また，本手法で提案したプライバシタグのデザイン検討をするにあたり，タグフレーム，ビット読み込みの予備評価を実施し，検知・解析手順を得た．そして，従来手法で用いられる QR コードベースのタグと提案タグをタグサイズ，撮影者から被写体までの距離の観点で評価を行い，これまでの課題であった被写体までの距離が理由による検知・解析精度の改善を示した．さらに，本タグをベースに Facebook を対象としたアプリケーション: Photo Privacy Realizer for Facebook を実装した．本方式の適用シーンとしては，被写体となった際に撮影者に自身のプライバシポリシを直接伝えることが難しい病院や警察などの関係者、または，子供を通じた弱い友人関係の紐付きであるママ友サークル，趣味を通じたサークルなどの小さなコミュニティが合っていると考えている．なお，本提案では非特定化の対象を顔領域に限定しているが，著作物など顔領域以外への展開も可能である．

第6章　結論

6.1 本研究の成果

本研究では，SNS の特徴である他ユーザとのオンラインコミュニケーションの楽しさを極力損なうことなく，SNS 投稿ユーザの主観的な判断基準のみに依存せずに，客観的なセンシティブデータの判断基準，および投稿内容に含まれる他ユーザのプライバシポリシに基づいて，投稿メッセージを適応的に保護する手法の確立を目的とした．本目的に向けて，1 章で挙げた 3 つの課題に沿い，研究を進めた．以下に，課題に対するその成果を示す．

課題 1：SNS におけるセンシティブデータの客観的な判断基準の定義

SNS ユーザの主観的な投稿判断を起因とする不用意な投稿と，それによるセンシティブデータ漏洩を防ぐために，最初に SNS におけるセンシティブデータ漏洩の発生状況を試算評価した．過去 1 年間分の 1 億 4,000 万件の Twitter アーカイブデータを用い，SNS 投稿後の後悔事例から抽出した後悔理由をもとに，これらセンシティブデータの漏洩有無とその検知手法を試算評価した．その結果，ごく少数であったものの，実際に漏洩が発生していることが確認できた．この評価結果を受けて，公文書におけるプライバシへの取り組みを参照し，SNS におけるプライバシ侵害情報分類を提案・評価し，SNS において客観的な判断基準として有効であることを示した．

従来研究との大きな違いは，過去に投稿された SNS メッセージアーカイブに含まれたセンシティブデータの漏洩有無の検知に留まらず，その漏洩の発生原因の 1 つがユーザの主観的な判断基準にあるとし，客観的な判断基準を定義するべく，公文書におけるプライバシの取り組みを参照して，非公開とすべき情報の内容分類と開示レベルを対応付けた，SNS におけるプライバシ侵害情報の分類を提案した．そして，ユーザの評価調査からその有効性を示したことである．本分類表をシステムに適用することにより，SNS ユーザは，投稿を行う前に客観的な判断基準のもと，投稿判断を可能とする第一歩を示した．

課題 2：センシティブデータの漏洩検知に基づき投稿ユーザへ指摘・通知

提案した SNS におけるプライバシ侵害情報分類表が持つ，SNS 投稿メッセージに含まれるセンシティブデータの有無とその開示レベルを客観的に判断可能とする特性を活かし，センシティブデータの漏洩有無の検知・通知と，その情報の重要度に沿った公開範囲の自動設定を提供する設定方式を提案した．また，本方式を基に，Facebook を対象とした Adaptive

Disclosure Controller for Facebook を実装して，検知したセンシティブデータの重要度に応じた開示レベルとセンシティブデータの含有を示すメッセージにより投稿ユーザへ視覚的に示し，センシティブデータが含まれていることを認識できるよう通知した．また，SNS の友達リストに含まれる友達ユーザをコミュニケーション頻度による友人関係性を Dunbar’s Circle 上に表現し，当該投稿がどのユーザへ公開されるかを視覚的に示した．さらに，開示レベルの選択による任意の公開範囲の選択，ドラッグ＆ドロップによる公開対象者の容易な選択ができる機能を提供した.

従来研究において，投稿前に投稿予定メッセージからセンシティブデータを検知し，投稿ユーザに再確認を促す，または，投稿に対する公開対象者を視覚的に表現し，投稿ユーザに認識させる手法が提案されていることを説明した．本研究では，これらに加え，SNS ユーザが持つ，投稿内容に応じた特定の個人や興味を持つコミュニティへのメッセージ公開要望に対して，容易に任意の公開範囲・公開対象者を定義可能とすることで，特定コミュニティ・グループにとらわれることなく，投稿メッセージを公開できること示した.

課題 3：コミュニティに応じた被写体のプライバシポリシの反映

写真に写る・写り込む被写体のプライバシ保護のために，被写体のコミュニティに応じたプライバシの振る舞いを埋め込んだプライバシタグを用いる方式を提案した. また, Instagram に投稿される写真の 46.6%は，自撮り，友人など 2 人以上と写った写真であることから，SNS 画像投稿時に予期される撮影者と被写体の距離を密接距離～社会距離とあると仮定し，提案したプライバシタグと従来手法で用いられる QR コードベースタグとの比較評価を行い，検知精度の改善を示した．さらに，本プライバシタグの検知・解析を可能としたアプリケーション：Photo Privacy Realizer for Facebook の実装を行った.

多くの従来研究は，投稿ユーザのプライバシ保護を目的としているが，本手法では，写真に写り込む被写体のプライバシ保護に注目した．また，被写体のプライバシ保護を目的とした従来手法との大きな違いは，人々が属するコミュニティや状況により，異なるプライバシの振る舞いを持っていることに注目し，コミュニティに応じたプライバシポリシの概念を取り込み，コミュニティごとに異なるポリシの適用・コミュニティメンバを公開対象者とすることで，当該投稿を限定的にしたことである．これらにより，SNS 投稿ユーザの主観的な判断基準のみに依存せずに，被写体のコミュニティや状況に応じたプライバシポリシを反映させることを可能とした.

6.2 今後の課題

前節では本研究の成果を示した一方，いくつかの制約や課題が残っている．以下に，本研究で提案した 3 つの提案手法に対するその課題および制約を示す．

提案 1：SNS におけるセンシティブデータの分類

提案した SNS におけるプライバシ侵害情報分類表の評価調査は，一対比較法を用いて実施し，評価条件としては満たしているが，社会調査として評価対象者の代表性を満たした評価とは言えず，結果に偏りが出ている可能性がある．

提案 2：センシティブデータの漏洩検知に基づく公開範囲の設定方式

アプリケーション：Adaptive Disclosure Controller for Facebook では，自身と友人の関係性を表現するために，先行研究をもとに，コミュニケーション頻度を用いた計算を行った．しかし，自身と友人の紐付き度合いを計る指標として，コミュニケーション頻度以外に，状況に応じたコンテキストの考慮など，別の指標が存在する可能性がある．また，SNS へ投稿後，時間経過とともに公開範囲や開示対象者が変化していくことが予想されるため，この時間経過の検討が必要である．

提案 3：被写体のコミュニティベース・プライバシポリシの設定方式

被写体が身に付けたプライバシタグを検知後，ポリシに従って自動的に非特定化を行っているが，投稿回数や投稿の経緯などの状況から非特定化の有無を判断するよう，投稿自体のコンテキストの考慮が必要である．また，提案タグの評価における課題として，従来手法の QR コードベースタグと提案タグの比較評価を行ったが，タグサイズと距離による基本評価のみとなっていることやその撮影条件を含め，評価を進める必要がある．

謝辞

本論文を執筆するにあたり，多くの方々のご指導とご協力を賜りました．ここにお世話になった方々への感謝の意を表します．

最初に，主任指導教員を務めてくださり，総合研究大学院大学における 3 年半のあいだ研究を支えて頂きました国立情報学研究所・総合研究大学院大学　越前功 教授に心より感謝致します．特に入学直後に訪れた研究の方向性が定まらなかった時期，仕事が忙しく十分な研究時間が確保出来ず，成果・進捗が上がらない時期に，状況に応じた厚い御指導と的確な方向性を示して頂き，誠にありがとうございました．また，学会発表以外にも，多くの場で発表，執筆の機会を与えてくださり，研究成果を最大限活かせるようにご配慮頂きました．ここに厚く御礼申し上げます．

指導教員を務めてくださいました国立情報学研究所・総合研究大学院大学 曽根原登 教授，に心より感謝申し上げます．本学入学前からプライバシ研究の本質を常に意識させるご指導をしてくださりました．ここに厚く御礼申し上げます．

本論文の審査をして頂いた国立情報学研究所・総合研究大学院大学 相原健郎 准教授，岡田仁志 准教授，津田塾大学 小舘亮之 教授に心より感謝申し上げます．相原准教授は，ヒューマンコンピュータインタラクションの観点から，本研究におけるプライバシ保護手法の実現化をご指導くださりました．岡田准教授は，法情報の観点から本研究におけるプライバシとセンシティブデータの位置付けについて，ご指導くださりました．小舘教授には，情報セキュリティの観点から様々なご意見を頂戴しました．ここに厚く御礼申し上げます．

そして，青山学院大学 梶山朋子 助教には，研究内容の議論のみでなく，評価手法や論文執筆手法など基礎的な研究手法と多岐に渡り，多くの御指導をして頂きました．梶山助教から頂戴した多くの手法は私の研究の礎となっております．厚く御礼申し上げます．

さらに，修士課程において主任指導教員を務めてくださり，私に本論文の研究テーマであるプライバシへの関心を植え付けて頂いた元首都大学東京産業技術大学院大学 嶋田茂 教授に深く感謝致します．博士課程進学後も気に掛けてくださり，多くのご意見を頂戴しましたことに厚く御礼申し上げます．

本研究を遂行するにあたり，総合研究大学院大学　越前功研究室 技術スタッフ　大金建夫氏，同研究室 技術補佐員　清野由美子氏には，多くの御指導，御助言を頂きましたことに深く感謝致します．また，孤独になりがちな研究・学生生活を共に過ごし，励まして頂いた同研究室 Nguyen Son Hoang Quoc さん，Bui Van Thach さんに感謝致します．

プライバシ侵害情報分類表の評価実験，実装システムの評価において，貴重な時間を割いて頂いた首都大学東京産業技術大学院大学の皆様，青山学院大学の皆様，SAP ジャパン株式会社の同僚の皆様に厚く御礼申し上げます．

最後に，私の学位挑戦に伴い，修士課程から含め 5 年半という長い時間を使うことを快諾してくれた最愛の妻・敬子に感謝いたします．時には，私の研究について熱く議論し，評価実験にも積極的に参加して頂きました．そして，医療従事者の視点からも多くの助言を頂きました．どんなに感謝してもしきれません．また，日に日に成長していく姿を見せ，私に安らぎを与えてくれた娘・史帆に感謝いたします．そして，いつも遠くから温かい目で見守ってくれた町田家 父・征之，母・加代子，大澤家 父・勝，母・久美子に感謝いたします．誠にありがとうございました．

2016 年 9 月

町田 史門

参考文献

[1] Joanna Brenner and Aaron Smith. 72% of Online Adults are Social Networking Site Users. Pew Internet Project, 2013.

[2] Cara Pring. 100 social media statistics for 2012, available from <http://thesocialskinny.com/100-social-media-statistics-for-2012/>, (accessed 2016-09-30).

[3] Dan Noyes. The Top 20 Valuable Facebook Statistics – Updated April 2016, available from <https://zephoria.com/top-15-valuable-facebook-statistics/>, (accessed 2016-09-30).

[4] Kevin Systrom. 300 Million: Sharing Real Moments, available from <http://blog.instagram.com/post/104847837897/141210-300million>, (accessed 2016-09-30).

[5] Bianca Bosker. The Twitter Typo That Exposed Anthony Weiner, available from <http://www.huffingtonpost.com/2011/06/07/anthony-weiner-twitter-dm_n_872590.html>, (accessed 2016-09-30).

[6] Lee Humphreys, Phillipa Gill and Balachander Krishnamurthy. Twitter: a content analysis of personal information. Information, Communication & Society, Volume 17, Issue 7, pages 843-857, 2014.

[7] Tehila Minkus, Kelvin Liu and Keith W. Ross. Children Seen But Not Heard: When Parents Compromise Children's Online Privacy. In Proceedings of the 24th International Conference on World Wide Web, pages 776-786, 2015.

[8] Priya Kumar and Sarita Schoenebeck. The Modern Day Baby Book: Enacting Good Mothering and Stewarding Privacy on Facebook. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, pages 1302-1312, 2015.

[9] Caroline Lang and Hannah Barton. Just untag it: Exploring the management of undesirable Facebook photos. Computers in Human Behavior 43, pages 147-155, 2015.

[10] 独立行政法人情報処理推進機構. 2014 年度 情報セキュリティの倫理に対する意識調査－調査報告書－, 入手先 <https://www.ipa.go.jp/files/000044094.pdf>, (参照 2016-09-30).

[11] 独立行政法人情報処理推進機構. 「ゴールデンウィーク（GW）の行楽写真を投稿する際

はご注意を」～ブログや SNS に投稿した写真からプライバシー漏洩の可能性～，入手先 <http://www.ipa.go.jp/security/txt/2015/05outline.html>, (参照 2016-09-30).

[12] Fred Stutzman and Jacob Kramer-Duffield. Friends only: examining a privacy-enhancing behavior in facebook. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 1553-1562, 2010.

[13] Arunesh Sinha, Yan Li and Lujo Bauer. What you want is not what you get: predicting sharing policies for text-based content on facebook. In Proceedings of the 2013 ACM workshop on Artificial intelligence and security, pages 13-24, 2013.

[14] Yabing Liu, Krishna P. Gummadi, Balachander Krishnamurthy and Alan Mislove. Analyzing facebook privacy settings: user expectations vs. reality. In Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, pages 61-70, 2011.

[15] Michelle Madejski, Maritza Lupe Johnson and Steven Michael Bellovin. The failure of online social network privacy settings, Columbia University Computer Science Technical Reports, 2011.

[16] Mainack Mondal, Yabing Liu, Bimal Viswanath, Krishna P. Gummadi and Alan Mislove. Understanding and Specifying Social Access Control Lists. In Proceedings of the Tenth Symposium On Usable Privacy and Security, pages 271-283, 2014.

[17] Fred Stutzman, Ralph Gross and Alessandro Acquisti. Silent Listeners: The Evolution of Privacy and Disclosure on Facebook. Journal of Privacy and Confidentiality: Vol. 4: Iss. 2, Article 2, 2012.

[18] 総務省パーソナルデータの利用・流通に関する研究会，入手先 <http://www.soumu.go.jp/main_sosiki/kenkyu/parsonaldata/>, (参照 2016-09-30).

[19] Lee Humphreys, Phillipa Gill, and Balachander Krishnamurthy. How much is too much? Privacy issues on Twitter. In Proceedings of the Conference of International Communication Association, 2010.

[20] Serge Egelman, Adrienne Porter Felt, and David Wagner. Choice architecture and smartphone privacy: There'sa price for that. The Economics of Information Security and Privacy, pages 211-236, 2013.

[21] パーソナルデータに関する検討会．個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律，入手先

<http://www.kantei.go.jp/jp/singi/it2/pd/info_h270909.html>, (参照 2016-09-30).

[22] パーソナルデータに関する検討会. 個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律 新旧対照条文, 入手先 <http://www.kantei.go.jp/jp/singi/it2/pd/pdf/taihihyo.pdf>, (参照 2016-09-30).

[23] 総務省統計局: 統計調査と個人情報保護, 入手先 <http://www.stat.go.jp/info/today/007.htm>, (参照 2016-09-30).

[24] Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro Giovanni Leon and Lorrie Faith Cranor. "I regretted the minute I pressed share": a qualitative study of regrets on Facebook. In Proceedings of the Seventh Symposium on Usable Privacy and Security, pages 10, 2011.

[25] Cheng Bo, Guobin Shen, Jie Liu, Xiang-Yang Li, YongGuang Zhang and Feng Zhao. Privacy.tag: privacy concern expressed and respected. In Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems, pages 163-176, 2014.

[26] Shane Ahern, Nathan Good, Simon King, Mor Naaman and Rahul Nair. Privacy Decisions for Location-Tagged Media. In Proceedings of the 8th international conference on Ubiquitous Computing, 2006.

[27] Benjamin Henne and Matthew Smith. Awareness about Photos on the Web and How Privacy-Privacy-Tradeoffs Could Help. In Proceedings of the Seventeenth International Conference on Financial Cryptography and Data Security, pages 131-148, 2013.

[28] Benjamin Henne, Christian Szongott and Matthew Smith. SnapMe if you can: privacy threats of other peoples' geo-tagged media and what we can do about it. In Proceedings of the sixth ACM conference on Security and privacy in wireless and mobile networks, pages 95-106, 2013.

[29] Manya Sleeper, Rebecca Balebako, Sauvik Das, Amber Lynn McConahy, Jason Wiese and Lorrie Faith Cranor. The post that wasn't: exploring self-censorship on facebook. In Proceedings of the 2013 conference on Computer supported cooperative work, pages 793-802, 2013.

[30] Andrew Besmer and Heather Richter Lipford. Privacy Perceptions of Photo Sharing in Facebook. In Proceedings of the Fourth Symposium on Usable Privacy and Security, 2008.

[31] Sanjay Kairam, Mike Brzozowski, David Huffaker and Ed Chi. Talking in circles: selective sharing in google+. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 1065-1074, 2012.

[32] Manya Sleeper, Justin Cranshaw, Patrick Gage Kelley, Blase Ur, Alessandro Acquisti, Lorrie Faith Cranor and Norman Sadeh. "i read my Twitter the next morning and was astonished": a conversational perspective on Twitter regrets. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 3277-3286, 2013.

[33] Lu Zhou, Wenbo Wang and Keke Chen. Tweet Properly: Analyzing Deleted Tweets to Understand and Identify Regrettable Ones. In Proceedings of the 25th International Conference on World Wide Web, pages 603-612, 2016.

[34] Sasa Petrovic, Miles Osborne and Victor Lavrenko. I Wish I Didn't Say That! Analyzing and Predicting Deleted Messages in Twitter. 2013.

[35] Parantapa Bhattacharya and Niloy Ganguly. Characterizing Deleted Tweets and Their Authors. In Proceedings of the Tenth International AAAI Conference on Web and Social Media, 2016.

[36] Huina Mao, Xin Shuai and Apu Kapadia. Loose tweets: an analysis of privacy leaks on twitter. In Proceedings of the 10th annual ACM workshop on Privacy in the electronic society, pages 1-12, 2011.

[37] Leysia Palen and Paul Dourish. Unpacking "privacy" for a networked world. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 129-136, 2003.

[38] Altman, Irwin. The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding. Brooks/Cole Publishing Company, 1975.

[39] Serge Egelman, Andrew Oates and Shriram Krishnamurthi. Oops, I did it again: mitigating repeated access control errors on facebook. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 2295-2304, 2011.

[40] Alessandra Mazzia, Kristen LeFevre and Eytan Adar. The PViz comprehension tool for social network privacy settings. In Proceedings of the Eighth Symposium on Usable Privacy and Security, Article No. 13, 2012.

[41] Anna Squicciarini, Dan Lin, Sushama Karumanchi and Nicole DeSisto. Automatic social group organization and privacy management. In Proceedings of the 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing, pages 89-96, 2012.

[42] Yang Wang, Pedro Giovanni Leon, Kevin Scott, Xiaoxuan Chen, Alessandro Acquisti and Lorrie Faith Cranor. Privacy nudges for social media: an exploratory Facebook

study. In Proceedings of the 22nd International Conference on World Wide Web, pages 763-770, 2013.

[43] Massimiliano La Gala, Valerio Arnaboldi, Marco Conti and Andrea Passarella. Ego-net digger: a new way to study ego networks in online social networks. In Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research, pages 9-16, 2012.

[44] ReThink, available from <http://www.rethinkwords.com/>, (accessed 2016-09-30).

[45] Anna Cinzia Squicciarini, Smitha Sundareswaran, Dan Lin and Josh Wede. A3P: adaptive policy prediction for shared images over popular content sharing sites. In Proceedings of the 22nd ACM conference on Hypertext and hypermedia, pages 261-270, 2011.

[46] Hoang-Quoc Nguyen-Son, Minh-Triet Tran, Tien-Dung Tran, Hiroshi Yoshiura, Noboru Sonehara, and Isao Echizen. Automatic Anonymous Fingerprinting of Text Posted on Social Networking Services. IEICE Transaction Information and System, pages 78-88, 2015.

[47] Jannik Strötgen and Michael Gertz. HeidelTime: High quality rule-based extraction and normalization of temporal expressions. In Proceedings of the 5th International Workshop on Semantic Evaluation, pages 321-324, 2010.

[48] Lujo Bauer, Lorrie Faith Cranor, Saranga Komanduri, Michelle L. Mazurek, Michael K. Reiter, Manya Sleeper and Blase Ur. The post anachronism: the temporal dimension of facebook privacy. In Proceedings of the 12th ACM workshop on Workshop on privacy in the electronic society, pages 1-12, 2013.

[49] Snapchat, available from <https://www.snapchat.com/>, (accessed 2016-09-30).

[50] Lukasz Piwek, Adam Joinson. "What do they snapchat about?" Patterns of use in time-limited instant messaging service. Computers in Human Behavior 54, pages 358-367, 2016.

[51] Oshrat Ayalon and Eran Toch. Retrospective privacy: managing longitudinal privacy in online social networks. In Proceedings of the Ninth Symposium on Usable Privacy and Security, Article No. 4, 2013.

[52] Roxana Geambasu, Tadayoshi Kohno, Amit A. Levy and Henry M. Levy. Vanish: Increasing Data Privacy with Self-Destructing Data. In Proceedings of the 18th International Conference on USENIX Security Symposium, pages 299-316, 2009.

[53] Primal Pappachan, Roberto Yus, Prajit Kumar Das, Tim Finin, Eduardo Mena and Anupam Joshi. A semantic context-aware privacy model for faceblock. In Proceedings of the 2nd International Conference on Society, Privacy and the Semantic Web - Policy and Technology, pages 64-72, 2014.

[54] Yan Li, Ke Xu, Qiang Yan, Yingjiu Li and Robert H. Deng. Understanding OSN-based facial disclosure against face authentication systems. In Proceedings of the 9th ACM symposium on Information, computer and communications security, pages 413-424, 2014.

[55] JD Woodward. Biometrics: Privacy's Foe or Privacy's Friend?. IEEE, pages 1480-1492, 1997.

[56] S.-C.S. Cheung, M.V. Venkatesh, J.K. Paruchuri, J. Zhao and T. Nguyen. Protecting and Managing Privacy Information in Video Surveillance Systems. Protecting Privacy in Video Surveillance, pages 11-33, 2009.

[57] Jehan Wickramasuriya, Mahesh Datt, Sharad Mehrotra and Nalini Venkatasubramanian. Privacy Protecting Data Collection in Media Spaces. In Proceedings of the 12th annual ACM international conference on Multimedia pages 48-55, 2004.

[58] Frank Pallas, Max-Robert Ulbricht, Lorena Jaume-Palasí and Ulrike Höppner. Offlinetags: a novel privacy approach to online photo sharing. In Proceedings of CHI '14 Extended Abstracts on Human Factors in Computing Systems, pages 2179-2184, 2014.

[59] Adrian Dabrowski, Edgar R. Weippl and Isao Echizen. Framework based on Privacy Policy Hiding for Preventing Unauthorized Face Image Processing. In Proceedings of 2013 IEEE International Conference on Systems, Man, and Cybernetics, pages 455-461, 2013.

[60] TagMeNot, available from <http://tagmenot.info/>, (accessed 2016-09-30).

[61] Ponnurangam Kumaraguru and Lorrie Faith. Cranor. Privacy indexes: a survey of Westin's studies. Technical report Carnegie Mellon University, 2005.

[62] 中前光弘，田畑洋二，大賀泰文，角田充弘，宇都文昭，奥西孝弘，越智　保，前田　要. Scheeffe の一対比較法による主観的評価法. 日本放射線技術学会雑誌, pages 1561-1565, 1996.

[63] 沈　潔如，稲葉由之，伊藤　一. 一対比較法を用いた観光客の期待度に関する調査と分

析: 台湾人観光客の事例. 日本経営工学会論文誌, pages 273-282, 2004.

[64] Twitter Streaming APIs, available from <https://dev.twitter.com/streaming/overview>, (accessed 2016-09-30).

[65] Wikipedia: 宗教一覧, 入手先 <http://ja.wikipedia.org/wiki/%E5%AE%97%E6%95%99%E4%B8%80%E8%A6%A7>, (参照 2016-09-30).

[66] Yahoo! 家庭の医学, 入手先 <https://medical.yahoo.co.jp/katei/>, (参照 2016-09-30).

[67] 岡本　裕. 9 割の病気は自分で治せる. page 1-46, 中経出版, 2009.

[68] 東山昌彦, 乾健太郎, 松本裕治. 述語の選択選好性に着目した名詞評価極性の獲得. 言語処理学会第 14 回年 次大会論文集, page 584-587, 2008.

[69] 工藤拓. MeCab: Yet Another Part-of-Speech and Morphological Analyzer, 入手先 <http://taku910.github.io/mecab/>, (参照 2016-09-30).

[70] 国立公文書館, 入手先 <http://www.archives.go.jp/>, (参照 2016-09-30).

[71] 公文書館における記録の公開と審査, アーカイブズ no.23, 2006.

[72] 戸嶋明. 地方公文書館における公開をめぐる問題と対応について. アーカイブズ, No.35, pages 40-44, 2009.

[73] 古賀崇. 政府・自治体における個人データの「時効」とアーカイブ. 第 13 回情報科学技術フォーラム, 2014.

[74] International Council on Archives, available from <http://www.ica.org/>, (accessed 2016-09-30).

[75] 独立行政法人国立公文書館利用規則, 入手先 <http://www.mext.go.jp/b_menu/hakusho/nc/k19720425001/k19720425001.html>, (参照 2016-09-30).

[76] RIM Dunbar. The Social Brain Hypothesis. Evol Anthropol 6, pages 178–190, 1998.

[77] SAM G. B. Roverts and ROBIN I.M. Dunbar. Communication in social networks: Effects of kinship, network size, and emotional closeness. Personal Relationships Volume 18, Issue 3, pages 439–452, 2011.

[78] Bruno Goncalves, Nicola Perra and Alessandro Vespignani. Validation of Dunbar's number in Twitter conversations. PLoS ONE 6(8), 2011.

[79] Path, available from <https://path.com/>, (accessed 2016-09-30).

[80] 福田忠彦，福田亮子．増補版　人間工学ガイド　感性を科学する方法．pages 73-123，サイエンティスト社，2011.

[81] Valerio Arnaboldi, Marco Conti, Andrea Passarella and Fabio Pezzoni. Analysis of Ego Network Structure in Online Social Networks. In Proceedings of 2012 International Conference on and 2012 International Confernece on Social Computing, pages 31-40, 2012.

[82] Shimon Machida, Tomoko Kajiyama, Shimada Shigeru and Isao Echizen. Analysis of Facebook Friends Using Disclosure Level. In Proceedings of the Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pages 471-474, 2014.

[83] Yuheng Hu, Lydia Manikonda and Subbarao Kambhampati. What We Instagram: A First Analysis of Instagram Photo Content and User Types. In Proceedings of the 8th International AAAI Conference on Weblogs and Social Media, 2014.

[84] ET. Hall. The hidden dimension. New York: Anchor Books/Doubleday, 1966.

[85] Ashwin Ashok, Viet Nguyen, Marco Gruteser, Narayan Mandayam, Wenjia Yuan and Kristin Dana. Do not share!: invisible light beacons for signaling preferences to privacy-respecting cameras. In Proceedings of the 1st ACM MobiCom workshop on Visible light communication systems, pages 39-44, 2014.

[86] Paul Viola and Michael J. Jones. Robust Real-Time Face Detection. International Journal of Computer Vision, vol. 57, issue 2, pages 137-154. Springer, 2004.

[87] 君塚正臣．憲法の私人間効力論．悠々社, 2008.

研究業績

1. 学術論文

[1] 町田史門, 梶山朋子, 嶋田茂, 越前功, “SNS におけるセンシティブデータの漏洩検知に基づく公開範囲の設定方式”, 情報処理学会論文誌　特集号 “新しいリスクに対応するコンピュータセキュリティ技術”, vol. 55, no. 9, pp. 2092-2103, 2014 年 9 月.

2. 国際会議論文

[1] Erwan Chaussy, Shimon Machida and Isao Echizen, “Definition of Private Information for Image Sharing in Social Networking Services”, in proceedings of the 13th International Workshop on Digital-forensics and Watermarking (*IWDW 2014*), pp. 544-556. Springer, October 2014.

[2] Shimon Machida, Tomoko Kajiyama, Shigeru Shimada and Isao Echizen, “Analysis of Facebook Friends using Disclosure Level”, in proceedings of the Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (*IIH-MSP 2014*), pp. 471-474. IEEE, August 2014.

[3] Shimon Machida, Tomoko Kajiyama, Shigeru Shimada and Isao Echizen, “Adaptive Disclosure Control System Using Detection of Sensitive Information in SNSs”, Symposium On Usable Privacy and Security (*SOUPS 2014*), 2 pages, https://cups.cs.cmu.edu/soups/2014/posters/soups2014_posters-paper14.pdf, July 2014.

[4] Shimon Machida, Shigeru Shimada and Isao Echizen, “Settings of Access Control by Detecting Privacy Leaks in SNS”, in proceedings of Signal-Image Technology & Internet-Based Systems (*SITIS 2013*), pp. 660-666 IEEE, December 2013.

3. 学会発表

[1] 町田史門, 越前功, “プライバシータグによるコミュニティを考慮した被写体のプライバシー保護手法”, コンピュータセキュリティシンポジウム 2015 (*CSS 2015*)予稿集, pp. 1119-1126, 2015 年 10 月.

[2] 町田史門, 梶山朋子, 嶋田茂, 越前功, “正規化順位法を用いた SNS におけるプライバシー侵害情報分類表の評価”, IEICE マルチメディア情報ハイディング・エンリッチメント研究会 (*EMM*), vol. 114, no. 118, pp. 145-148, 2014 年 7 月.

[3] 町田史門, 梶山朋子, 嶋田茂, 越前功, “センシティブデータの漏洩検知による適応的な公開範囲設定システムのプロトタイプ実装”, IEICE マルチメディア情報ハイディング・エンリッチメント研究会 (*EMM*), vol. 113, no. 480, pp. 51-56, 2014 年 3 月.

[4] 町田史門, 嶋田茂, 越前功, “SNS 上のプライバシーセンシティブ情報の漏洩検知に基づ

く公開範囲の設定方式”, コンピュータセキュリティシンポジウム 2013 (*CSS 2013*)予稿集, vol. 2013, no. 4, pp. 566-573, 2013 年 10 月. <**情報処理学会 2014 年度年度山下記念研究賞, コンピュータセキュリティシンポジウム 2013 優秀論文賞, コンセプト論文賞　受賞**>

[5] 町田史門, 嶋田茂, 越前功, “デジタル私文書におけるプライバシーセンシティブ情報の漏洩検知に基づく公開範囲の設定方式の提案”, IEICE マルチメディア情報ハイディング・エンリッチメント研究会 (*EMM*), vol. 113, no. 212, pp. 31-36, 2013 年 9 月.

4. 書籍等出版物

[1] 町田史門, 越前功, “SNS 写真データにおけるプライバシー保護 －コミュニティに基づく被写体情報の保護と活用－”, 統計と情報の専門誌「エストレーラ」5 月号, 2016 年 5 月.

5. 招待講演

[1] Shimon Machida, Isao Echizen, “Settings of Access Control by Detecting Privacy Leaks in SNS”, International Workshop on Security (*IWSEC 2014*), 2014 年 8 月.

6. その他発表等

[1] 町田史門, 越前功, “～写真共有によるプライバシー侵害を防止するには～ PrivacyTag : SNS 投稿ユーザの主観的な判断基準のみに依存しない被写体のプライバシー保護”, 国立情報学研究所 オープンハウス 2016, 2016 年 5 月. [ポスター発表]

[2] 町田史門, 越前功, “PrivacyTag : コミュニティを考慮した被写体のプライバシー保護手法”, コンピュータセキュリティシンポジウム 2015 (*CSS 2015*), 2015 年 10 月. [デモンストレーション・ポスター発表]

[3] 町田史門, 越前功, “～写真共有によるプライバシー漏洩を防止するには～ プライバシータグによるコミュニティを考慮した画像プライバシーの保護手法”, 国立情報学研究所 オープンハウス 2015, 2015 年 6 月. [ポスター発表]

[4] 町田史門, 越前功 “Disclosure Control with Reasonable Expectation of Privacy in Online Social Networks”, 社会のイノベーションを誘発する情報システム (*ISSI 2014*), 2015 年 2 月. [ポスター発表]

[5] 町田史門, 越前功, “‘Understanding and Specifying Social Access Control Lists’の紹介“, IPSJ 情報セキュリティ心理学とトラスト研究会 (*SPT*), 2014 年 10 月. [口頭発表]

[6] 町田史門, 梶山朋子, 嶋田茂, 越前功, “～ソーシャルメディアを安心・安全に楽しむため～ SNS でプライバシー漏洩を防ぐための公開範囲の設定方法”, 国立情報学研究所 オープンハウス 2014, 2014 年 5 月. [ポスター発表]

[7] 町田史門, 梶山朋子, 嶋田茂, 越前功, “センシティブデータの漏洩検知による適応的な公開範囲設定システムのプロトタイプ実装”, IEICE マルチメディア情報ハイディング・エンリッチメント研究会 (*EMM*), 2014 年 3 月. [ポスター発表]

[8] 町田史門, 越前功 “Settings of Access Control by Detecting Privacy Leaks in SNS”, 社会のイノベーションを誘発する情報システム (*ISSI 2013*), 2014 年 2 月. [口頭発表]

7. 受賞

[1] 国立情報学研究所優秀学生賞(受賞者：町田史門), 2016 年 3 月.

[2] 情報処理学会 2014 年度山下記念研究賞(2013 年 10 月・SNS 上のプライバシーセンシティブ情報の漏洩検知に基づく公開範囲の設定方式，学会発表[4]，受賞者：町田史門)

[3] 情報処理学会コンピュータセキュリティシンポジウム 2013(*CSS 2013*) 優秀論文賞（2013 年 10 月・SNS 上のプライバシーセンシティブ情報の漏洩検知に基づく公開範囲の設定方式，学会発表[4]，受賞者：町田史門，嶋田茂，越前功）

[4] 情報処理学会コンピュータセキュリティシンポジウム 2013(*CSS 2013*) コンセプト論文賞（2013 年 10 月・SNS 上のプライバシーセンシティブ情報の漏洩検知に基づく公開範囲の設定方式，学会発表[4]，受賞者：町田史門，嶋田茂，越前功）

付録

A.1 公文書の公開に関する運用基準： プライバシー等侵害情報分類表（戸嶋 私案）

別表1

公文書の公開に関する運用基準　プライバシー等侵害情報分類表（私案）

非公開とすべき情報の内容による分類例 ＼ 非公開とすべき情報の重要度による分類及び非公開期間		a　非公開期間120年	b　非公開期間50年	c【非公開期間30年】	（参考）秋田県公文書館の点頭結果内訳（27,079件のうち一部非及び非公開簿冊件数及び割合） 件数	割合(%)
I 個人の内心に関する情報	ア　思想、信条	一般人の思想、信条		一般的な生活信条、人生・社会・政治観	52	0.9
	イ　宗教	一般人の宗教信徒情報			26	0.4
II 個人の心身の状態に関する情報	ア　病歴	遺伝性疾患、伝染性疾患、精神性疾患及び重度の疾患		一般的な病気の既往症、軽度の疾患	652	11.3
	イ　心身の記録		特殊な身体記録、精神症状	一般的な身体記録	3	0.1
III 個人の基本情報、生活の状況に関する情報	ア　戸籍、外国人登録、写真	門地、戸籍、外国人登録指紋	写真（個人が特定されるもの）		1,520	26.3
	イ　家庭状況	特殊な生育歴、悲惨な家庭状況		一般的な家庭状況	136	2.4
	ウ　行動傾向	反社会的、非常識的な行動傾向（性格、趣味、嗜好）		一般的な行動傾向（性格、趣味、嗜好）	119	2.1
IV 個人の経歴、社会的活動等に関する情報	ア　学歴、試験結果		最終学歴（卒業、中退）懲戒（退学・停学）試験結果		129	2.2
	イ　職歴、団体・社会活動		職歴、団体・社会活動懲戒、公務員の分限処分、叙勲表彰関係（非受賞者）	公務員の一般的な職歴叙勲表彰関係（受賞者）	2,784	48.2
	ウ　犯罪及び不法行為	犯罪歴（罰金刑以下を除く）	犯罪歴（罰金刑以下）民事事件の不法行為		96	1.7
	エ　犯罪被害及び不法行為の被害	犯罪被害で不名誉なもの民事事件の不法行為の被害で不名誉なもの	犯罪被害（罰金刑以下）民事事件の不法行為の被害		5	0.1
	オ　訴訟及びその他事件	その他の事件、事故、災害で不名誉なもの破産、個人事業倒産	民事訴訟の敗訴、強制執行行政事件の不利益処分	民事訴訟の勝訴その他の民事・行政事件、その他事件、事故、災害	58	1.0
	カ　歴史的事実	同和問題関係戦犯	軍歴、戦争関係小作調停関係		15	0.3
V 個人の財産状況に関する情報	ア　財産状況		財産全体	財産の一部	45	0.8
VI　その他（原課協議によるもの。法人営業情報、住民の安全確保等）					141	2.4
					5,781	100.0

備考

1　この表は、表頭に非公開とすべき情報の重要度（「個人の秘密」、「個人に重大な秘密」、「個人の特に重大な秘密」）、表側にそれぞれに該当する可能性が考えられる一般的な情報の類型を情報公開条例の非開示情報を参考に分類表示したものであり、歴史公文書等の非公開情報の点検に当たっては、当該情報の具体的性質、当該情報が記録された当時の状況等を総合的に勘案して個別に判断するものとする。

2　本人の死後遺族にも影響を及ぼすような特に重大な秘密（遺伝性疾患、不名誉な犯罪被害等）、現在も継続中であり公開が困難な門地、同和関係、法的判断の定まっていない戸籍等については、将来の公開時期の社会状況等を勘案して延長も可能とするものとする。

※複数理由の資料があるため合計は件数5，575件とは一致しない。