卒業論文　　2012年度（平成24年度）

ソーシャルメディアにおける匿名度評価手法の提案と実装

慶應義塾大学 環境情報学部

氏名：有馬 怜文

担当教員

慶應義塾大学 環境情報学部

村井 純

徳田 英幸

楠本 博之

中村 修

高汐 一紀

Rodney D. Van Meter III

植原 啓介

三次 仁

中澤 仁

武田 圭史

平成25年1月22日

卒業論文要旨 - 2012 年度 (平成 24 年度)

ソーシャルメディアにおける匿名度評価手法の提案と実装

近年，一般ユーザにおけるソーシャルメディアの利用が活発になり，ユーザが自由に情報を発信することのできる Web ページ（以下，パーソナルページ）の数も増加している．同時に，パーソナルページにおいて公開されている個人情報を元にし，なんらかの不祥事をきっかけに爆発的な注目を集めてしまう事件（以下，炎上事件）が発生している．そのため，多くの識者がパーソナルページの扱い方へ注意を喚起する一方，ソーシャルメディアの利便性も伴い，炎上事件の発生が終息する傾向は見て取られない．これらの問題を解決するため，パーソナルページにおいて，個人情報を公開することに対する危険性を明らかにすることが必要とされている．

本論文では，まず個人情報に焦点が当てられている事例に対し，公的な機関や民官の企業から発信されている情報を元に事例分析を行った．そしてパーソナルページにおける匿名度を評価することによって，潜在する危険性を示した．

本論文が提案する手法は大きく 3 つある．まず，パーソナルページのプロフィール欄において，それぞれのユーザにおけるキーワードとなる単語に対し，同じソーシャルメディアを使うユーザ内においての一意性を調査した．次に，パーソナルページの自由記述部分において，それぞれのユーザが発信している内容へ形態素解析を行い，出現回数の多い単語からユーザを特徴付けるものを抽出した．そして，抽出したユーザの特徴から，他ソーシャルメディアにおける同ユーザのパーソナルページを検索し，紐付けを行った．

これら 3 つの手法を用いて多数のユーザへ実験を行った結果，ソーシャルメディアにおけるパーソナルページから，ユーザの特定が容易であることが示された．これによって，潜在する危険性を顕在化させ，ソーシャルメディアを扱うにあたってのガイドライン作りに成功した．

キーワード:
1. ソーシャルメディア, 2. プライバシ, 3，個人情報, 4. 匿名性, 5. インターネット

慶應義塾大学 環境情報学部

有馬 怜文

Abstract of Bachelor's Thesis - Academic Year 2012

Evaluation method for anonymity on Social Media: Proposal and Implementation

In the recent years, the usage of the social media by general users are on the move, and the number of personal pages where the users can send information are increasing. At the same time, flaming incidents that collect explosive attention is happening, which are triggered by misconduct events. From this background, while many leaders spark up caution for the treatment of personal pages, the user-friendliness of the social media prevents the flaming incidents from comming to an end. To solve these problems, the high risks of exposing personal information on the personal page needs to be revealed.

In this thesis, the example analysis is preformed on the examples focusing on personal information from public institution and private companies. By evaluating the anonymity on the personal page, the hidden risk is clarified.

There are mainly 3 methods that this thesis proposes. First, the uniqueness of the keywords used in the profile section of the personal page was investigated, focusing on the users using the same social media. Secondly, the information that the user is sending from the free space of the personal page is put through the morphological analysis. Out of the words that are frequently used, the words which characterises the user is extracted. Thirdly, using by the extracted words about the user, a retrieval is tested on other social medias to find the personal pages of the same user. The page becomes linked and checked if there is a true relation.

The results from the experiments using the 3 methods showed that the identification of a specific user can be easily done from the personal page of the social media. From this result, by actualizing the hidden risks, the guideline of how to deal with social media is created, and was a success.

Keywords :
1. Social Media, 2. Privacy, 3. Personal Data, 4. Anonymity, 5. Internet

Keio University, Faculty of Environment and Information Studies

Remon Arima

目 次

図 目 次

表 目 次

第1章　序論

本章では，一般ユーザによるソーシャルメディア上でのパーソナルページの増加及び多様化により，情報発信の敷居が低くなったため，Web 上での個人情報の公開が容易になった現状を示す．また，そのため悪意を持ったユーザから意図しない状況での被害を受けてしまう可能性を明示し，パーソナルページの危険性を顕在化させる必要性を述べる．次に，本論文中にて扱う用語の定義を行う．そして本論文の構成を記す．

1.1　一般ユーザによる Web 上での情報公開の現状

近年，一般ユーザによるソーシャルメディアの利用が活発になり，同時にユーザが自由に情報を発信することのできる Web ページ（以下，パーソナルページ）の数も増加している．ブログサービスの利用に始まり，その後のソーシャルネットワーキングサイト（以下，SNS）の台頭を経て現在に至っている．SNS の黎明期は，招待制や承認制といった不特定多数のユーザへ非公開とする形式が主な形式であったが，近年では誰でも登録することができ，かつ不特定多数のユーザへ自身を紹介することのできる仕組みとなっているサービスが大手となっている．現在世界で最も多い登録者数を誇っている SNS は，アメリカのフェイスブック社が提供している Facebook[1] であり，その利用者数は 8 億人を越えるとされている．SNS には実名登録が原則となっているものもあり，そのメインコンテンツとなるパーソナルページへ内容として記述できる項目は，性別や年齢といったものから勤務先や結婚歴まで様々である．パーソナルページのトップにはプロフィール画像として自身の写真を設定することもできるため，現実世界におけるコミュニティを Web 上へ直接持ち込むことができるものとなっている場合が多い．しかしその利用の目的は様々であり，Facebook のように他人とのつながりを強調する目的のもの，また Mobage[2] や GREE[3] といった数々のソーシャルゲームを活用するための場としての SNS もある．ほか，SNS ではないものの，パーソナルページにて 140 字以内のつぶやきを投稿する，ミニブログ形式である Twitter[4] のようなサービスも多くのユーザを獲得している．

このようなパーソナルページ増加の要因には，アクセス手段の多様化も関係している．パーソナルコンピュータ（以下，PC）が一家に一台だった時代から一人に一台へと変遷し，近年ではスマートフォンの普及によって常時に渡るインターネットへの接続が可能となっている．またスマートフォンの機能としてのカメラ機能を用いることにより，自身に起きたことや外出先での出来事を気軽に画像として発信することができるようにもなっている．更に GPS 機能によって位置情報を付加した情報発信も可能になるなど，パーソナルページの使い方は多様化を極めている．

これらから，現状における SNS を始めとする，ソーシャルメディア上でのパーソナルページにおいては匿名性が非常に低く，対象ユーザの詳細を既知とする検索においてはほぼ確実に見つけることができ，またそうでない場合においても幾つかのキーワードを設定することにより見つけることができる可能性がある．Web 上でのつながりに特化することのできる反面，それにより起こりうる問題もある．

1.2 問題点

パーソナルページにおいて公開される情報において，公開者の意図は問わずその内容が誰かに著しく不利益を与えてしまう場合がある．そのような場合に，不利益を与えられた者，あるいは状況に対し正義感を働かせた者たちが，ユーザの情報を元に攻撃的な手段を用い，報復行為を行う事例が発生している．また，不利益を与えられる対象が公開者自身であった場合としては，ユーザに対しつきまとい行為を行う，いわゆるネットストーキングに遭遇し，直接的な被害を受ける事例もある．

このように，一般ユーザによるパーソナルページの増加及び多様化により，Web 上での交流が活発になった反面，副作用としての公開情報を用いた問題が発生しているのである．

1.3 本研究の目的

本論文の目的は，パーソナルページにおいて公開されている個人情報を元に，各ユーザにおけるパーソナルページの匿名度を評価する手法を提案し，潜んでいる危険性を明らかにすることである．これにより，一般ユーザがパーソナルページにおいて，自身の意図しない事件に巻き込まれないための Web ガイドライン作りが可能になることも期待される．

1.4 本論文中における用語の定義

本論文では，パーソナルページ及び炎上事件という用語を用いる．ここではこれらの用語について定義を行う．本論文におけるパーソナルページとは，SNS を始めとするソーシャルメディアに特有である，ユーザが自身の情報を公開する場所として存在する Web ページのことであると定義する．各ソーシャルメディアによって名称が違うため，本論文では統一した名称を用いるものとする．また，本論文における炎上事件とは，佐伯胖の著書 [5] に基づき，サイト管理者の想定を大幅に超え，非難・批判・誹謗・中傷などのコメントやトラックバックが殺到することであると定義する．

1.5 本論文の構成

本論文は全 7 章から構成される．第 2 章では，プライバシの扱いや個人情報の扱い，また SNS を始めとしたパーソナルページの現状について述べる．第 3 章では，第 2 章で述べた課題に取り組む関連研究を紹介する．第 4 章では，個人情報に焦点が宛てられている事例に対し，公的な機関や民官の企業から公開されている情報を元に事例分析を行う．第 5 章では，実際にパーソナルページの匿名度を評価するための手法を提案する．第 6 章では，第 5 章で提案した手法を元に実装したもの，それを使い行った実験とその結果，そして評価を述べる．最後に，第 7 章で本論文の結論と作成した Web ガイドライン，そして今後の展望を述べる．

第2章　背景

本章では，日本におけるプライバシと個人情報の扱いについて定義し，それらがどのように扱われているかを述べた．また SNS を始めとするソーシャルメディアを挙げ，それぞれの持つ特徴を紹介した．そしてプライバシが脅かされる危険性について調査し，主にソーシャルメディアのパーソナルページにおいて，ユーザ毎に危険性を顕在化させる必要性に着目した．

2.1　プライバシとは

プライバシとは，個人に関する全ての情報を扱う際にその本人の承諾が無い状態で利用されることを防ぐための権利であり，日本語で「私事権」と訳されることもある．プライバシの起源は Samuel D.Warren, Louis D. Brandeis らが 1980 年に発表した The Right to Privacy[6] にある the right to be let alone であると言われている．また日本においては，「宴のあと」裁判 [7] の東京地裁判決でプライバシという言葉が使われて以来，人格権として認められ言葉としても定着をしている．

プライバシ権は 2012 年現在，日本国憲法において明文化された規定はないが，主に憲法第 13 条によって保障されると解されるのが一般的であり，その内容は図 2.1 に示した内容となっている．そして，近年プライバシ権は請求権に基づいた積極的な権利とされており，他者が管理している自己の権利に対し訂正や削除を求めることができるものとなっている．また，1980 年に OECD 理事会は図 2.2 の内容となる「プライバシ 8 原則」を勧告し，世界各国の個人情報保護制度に影響を与えた．

> 第 13 条
>
> 全て国民は，個人として尊重される．生命，自由及び幸福追求に対する国民の権利については，公共の福祉に反しない限り，立法その他国政の上で，最大の尊重を必要とする．

図 2.1: 日本国憲法 第 13 条

プライバシによって守られる情報は主に個人情報とされ，これらが侵害されることによって事件へと発展した事例もある．

1 収集制限の原則
個人データは，適法・公正な手段により，かつ情報主体に通知または同意を得て収集されるべきである．

2 データ内容の原則
収集するデータは，利用目的に沿ったもので，かつ，正確・完全・最新であるべきである．

3 目的明確化の原則
収集目的を明確にし，データ利用は収集目的に合致するべきである．

4 利用制限の原則
データ主体の同意がある場合や法律の規定による場合を除いて，収集したデータを目的以外に利用してはならない．

5 安全保護の原則
合理的安全保護措置により，紛失・破壊・使用・修正・開示等から保護すべきである．

6 公開の原則
データ収集の実施方針等を公開し，データの存在，利用目的，管理者等を明記するべきである．

7 個人参加の原則
データ主体に対して，自己に関するデータの所在及び内容を確認させ，または異議申立を保証するべきである．

8 責任の原則
データの管理者は諸原則実施の責任を有する．

図 2.2: OECD8 原則

2.1.1 個人情報とは

個人情報とは，個人に関する情報として広い解釈を持たせられることが多いが，個人情報の保護に関する法律第 2 条には図 2.3 のように定められている．

また，日本工業規格「JIS Q 15001 個人情報保護マネジメントシステムー要求事項」には，図 2.4 のように定められている．法律とは違い「生存する」という言葉は入っておらず，事業者においては故人の情報も個人情報に含まれるとされる．

主にプライバシについて語られる際，個人情報との扱いに用語的な混同が生じる場合があるが，先述の OECD 理事会やアメリカ合衆国におけるプライバシ権利章典 [8]，ヨー

第 2 条
この法律において「個人情報」とは，生存する個人に関する情報であって，当該情報に含まれる氏名，生年月日その他記述等により特定の個人を識別することができるもの（他の情報と容易に照合することができ，それにより特定の個人を識別することができることとなるものを含む.）をいう.

図 2.3: 個人情報の保護に関する法律 第 2 条

JIS Q 15001
個人に関する情報であって，当該情報に含まれる氏名，生年月日その他の記述，又は個人別につけられた番号，記号その他の符号，画像若しくは音声によって当該個人を識別できるもの（当該情報だけでは識別できないが，他の情報と容易に照合することができ，それによって当該個人を識別できるものを含む）

図 2.4: JIS 規格

ロッパ連合（EU）におけるデータ保護規則提案 [9] においては明確に個人情報について定義がされている．それは個人を識別することのできる情報のことであり，また人によって管理のされている状態にある情報のことである.

2.1.2 個人情報の扱いにおけるユーザの立場

個人情報の扱いについて，データ保護規則提案では 2012 年 1 月 25 日，right to be forgotten という言葉を定義した．これは日本語訳で忘れられる権利と訳され，ユーザにおいて自身のデータが不要となった際，あるいはデータを消して欲しくなった際に，データの削除を要請することのできる権利である．これは EU においての情報管理に対し，ユーザから十分な信頼を得るために盛り込まれた内容であるといわれている．他，プライバシ権利章典においては，Do Not Track という概念が明確化された．これは消費者が事業者によるインターネット上での追跡をを拒否することのできる権利であり，アメリカ合衆国でのインターネット上における個人情報の扱いについて，消費者の権利を確立することが目的とされている.

2.2 プライバシを守るために行われている事例

プライバシが脅かされる事件の発生や，インターネットをはじめとしたネットワーク技術や情報処理技術の進展により，ユーザも自身のプライバシが利用する企業やサイトにお

いて守られているかを注意するようになった．ここでは，主にユーザのプライバシを守るために行われている事例として，制度的なものと技術的なものをそれぞれ述べる．

2.2.1 プライバシマーク制度

企業やサービスがプライバシを守っているかを判断する基準の一つとして，プライバシマーク制度 [10] が挙げられる．プライバシマーク制度とは，図 2.4 に示した JIS 規格に適合して，個人情報について適切な保護措置を講ずる体制を整備している事業者等を認定してマークを付与する制度である．図 2.5 にプライバシマークを示す．

図 2.5: プライバシマーク

このような認証マーク制度体制の整備においては，まず消費者の目に見えるプライバシマークで示すことによって，個人情報の保護に関する消費者の意識の向上を図ること，そして適切な個人情報の取扱いを推進することによって，消費者の個人情報の保護意識の高まりに応え，社会的な信用を得るためのインセンティブを事業者に与えることといった目的が与えられている．

2.2.2 Privacy by Design

そして，計画的なプライバシ対策が行われているかを評価する指標として，Privacy by Design[11] が挙げられる．これは，the Office of the Information and Privacy Commissioner の Ann Cavoukian 博士が 1990 年代に提唱したものである．図 2.6 に Privacy by Design のマークを示す．プライバシ侵害のリスクを低減するために，システムの開発においてプロアクティブにプライバシ対策を考慮し，企画から保守段階までのシステムライフサイクルで一貫した取り組みを行うことが掲げられている．

図 2.6: Privacy by Design のマーク

2.2.3 プライバシ保護データマイニング

プライバシマークと Privacy by Design が，制度としてプライバシを守るための取り組みであることに対し，プライバシ保護データマイニングは技術そのものである．ビッグデータを扱うようになってきた中，利用者がデータベースに質問しても真のデータベースの内容が利用者には取得できないようにし，かつ一方でできるだけ正確なデータマイニング結果も取得したい場合に行われる手法である．複数の組織がプライバシーに関わるクリティカルなデータを持ち，場合によっては公開をしている現状において，マイクロデータと呼ばれる詳細データが解析やデータマイニングに利用される．そのマイクロデータに雑音を加えることで実現される二種類の方法について述べる．

- **k-匿名性**
  k-匿名性とは，表 2.1（マイクロデータ）を元に，表 2.2 を作り上げる匿名化手法である．k-匿名性を用いたデータは，ある組み合わせについて少なくとも k 人以上の数が存在している．マイクロデータにおいて，一人目と二人目は同じデータである．しかし三人目以降はそれぞれが違う組み合わせを持ったデータとなっているため，二種類の特徴を抽出すれば個人が特定できてしまう．これを防ぐため，複数人のデータを同じ外見にまとめる手法を用いるものが k-匿名性である．表 2.2 では，三人目，五人目，七人目が同じ「学生」及び「1990 年代生まれ」とされ，また四人目と六人目が性別を隠されている．これによって最大三人の同じ組み合わせを持った人物が出現することになっている．この状態は 3-匿名性を持ったデータとされる．

- **l-多様性**
  l-多様性とは．同じく表 2.1 を元に，表 2.3，表 2.4 を作り上げる匿名化手法である．l-多様性を用いたデータは，あるグループについて，少なくとも l 種類以上の組み

表 2.1: マイクロデータ

Faculty	BirthDate	Gender	Points
環境	1989	F	100
環境	1989	F	100
総合	1990	M	150
総合	1990	F	80
環境	1991	M	150
総合	1990	M	80
環境	1992	M	150

表 2.2: k －匿名性を用いた匿名化データ

Faculty	BirthDate	Gender	Points
環境	1989	F	100
環境	1989	F	100
学生	199*	M	150
総合	1990	*	80
学生	199*	M	150
総合	1990	*	80
学生	199*	M	150

合わせが存在している．マイクロデータ内の組み合わせを，上 3 人及び下 4 人にグループ分けをし，また Points の項目に着目し頻度を抽出する．これによって，A グループにはそれぞれが 1 度ずつ出現，また B グループには 150 だけが 2 度出現することを表すことができる．この状態は 3-多様性を持ったデータとされる．

2.2.4 暗号化技術

情報漏洩に対し効果的とされる対策として，暗号化技術も挙げられる．情報を暗号化した技術で保存する手法であり，万が一漏洩した際にもその内容を復号化しない限り活用できない点が特徴である．ここでは，主に扱われる 2 つの手法について述べる．

- **共通鍵暗号方式**
  共通鍵暗号方式とは，暗号化と復号化で同じ鍵を用いる手法である．処理の高速性が特徴であり，鍵のアルゴリズムが公開されている場合が多い．また，通信を行う際に共通鍵を扱う利用者同士は，互いに秘密鍵を共有していることが前提とされて

表 2.3: l-多様性を用いた匿名化データ 1

Faculty	BirthDate	Gender	Group
環境	1989	F	A
環境	1989	F	A
総合	1990	M	A
総合	1990	F	B
環境	1991	M	B
総合	1990	M	B
環境	1992	M	B

表 2.4: l-多様性を用いた匿名化データ 2

Group	Points	Frequency
A	80	1
A	100	1
A	150	1
B	80	1
B	100	1
B	150	2

いる．利用者の対象が不特定多数となる場合は．鍵を共有するためのプロトコルを利用者が実行する必要がある．

- **公開鍵暗号方式**
  公開鍵暗号方式とは，暗号化と復号化で違う鍵を用いる手法である．暗号化の際に用いる鍵は公開鍵と呼ばれ，その名の通り公開されることに問題はない．公開鍵に対し復号化の際に用いる鍵は，暗号を受け取る利用者のみが所持するため，暗号化は誰でもできるものの復号化は一人しかできないこととなる．これもアルゴリズムが公開されている場合が多い．

2.3 SNS とは

SNS とは，ソーシャルネットワーキングサービス（Social Networking Service）の略である．社会的ネットワークを Web 上に構築することのできるサービス及びサイトであることが特徴であり，世界において多くの種類が存在している．それらが有する基本的な機能として，プロフィール機能，メッセージ機能，ユーザ間における検索や相互リンク機能，

ブログ機能，コミュニティ機能といったものがある．以下にそれぞれの SNS が持つ特徴を挙げる．

2.3.1 SNS の種類

- **Facebook**[1]
  フェイスブック株式会社の提供する SNS．2004 年にサービスが開始された，当初は学生のみに限定していたが，2006 年 9 月 26 日以降は一般にも開放された．13 歳以上が利用の条件で，日本では 2008 年にサービスが開始された．また実名登録制となっており，個人情報の登録も必要となっている．2011 年 9 月の段階でユーザが世界で 8 億人となり，事実上の世界最大手 SNS となっている．

- **Myspace**[12]
  News Corpration の運営する SNS．設立は 2003 年で，日本においても 2006 にサービスが開始された．当初から音楽マニアの間で発展し，現在も音楽やエンターテイメントを中心とした盛り上がりを見せている．会員の種類としてユーザとアーティストといった二つがあり，アーティスト登録の場合は音声ファイルの公開や登録内容に若干の違いがある．

- **LinkedIn**[13]
  2002 年に Reid Hoffman の自宅のリビングルームで開発されたソーシャルネットワーキングサービス．世界最大のプロフェッショナルネットワークとして，ビジネスに特化したサービスを提供する．ユーザは履歴書情報をサービス上で求人や商談を行うことや，専門家とコンタクトをとることができる．

- **mixi**[14]
  株式会社 mixi の運営する，日本においては最大規模のソーシャルネットワーキングサービス．当初は既登録ユーザによる完全招待制で始まり，ユーザにおいて健全で安全性の高いサービスを提供することが目的とされた．2010 年 3 月よりは誰でも参加することのできる登録制へと変更された．また，18 歳未満は禁止の制限があったが 2008 年 12 月 10 日より条件付きで 15 歳未満への制限と変更された．

- **Mobage**[2]
  株式会社 DeNA の運営するソーシャルネットワーキングサービス．2006 年に 2 月 7 日スマートフォン向けでサービスを開始し，2008 年 7 月以降は PC 版へ Yahoo!モバゲーを開始させた．2008 年からは携帯電話向けのゲームや SNS の提供サイトとしてのテレビ CM と同時に，携帯向けポータルサイトとしての宣伝を開始した．ゲームを中心としたアバター等による交流が特徴である．

2.3.2 その他のパーソナルページを持つサービスの例

- **Twitter**[4]
  Twitter Inc. の提供する情報サービス．ツイートと称される 140 文字以内の短文を投稿することのできるミニブログ形式である．相互フォロー機能がありユーザ間でのつながりが発生するため，広義でのソーシャルネットワーキングサービスとも言えるが，Twitter 自身がそれを否定している．アバターやゲーム等の機能はないものの，簡単な作りとその使いやすさから，2012 年 8 月にはユーザ数が 5 億人を突破したといわれている．

- **Timelog**[15]
  株式会社サンロフトの提供する Web サービス.「今何をしているのかをメモしていくだけのシンプルなサービス」と銘打ち，Twitter と同じミニブログ形式となっている．Twitter とは違い様々な機能を有し，例えば書いたメモは後から見返すことが容易となっている上，他ユーザからメモに対し評価を得ることもできる．2006 年にサービスが開始されたが，2011 年 8 月に個人の利用を停止，様々な内容が終了した．

2.4 プライバシが脅かされた例

人々の間でプライバシに焦点が当たる際，何かしらのプライバシ侵害に関する事件が起きている場合が多いが，そのような事件の件数，及び有名なものを挙げる．

2.4.1 インシデント件数

NPO 日本ネットワークセキュリティ協会（JNSA）[16] による「情報セキュリティインシデントに関する調査報告書 個人情報漏洩編」[17] によると，2011 年における漏洩事件の数は 1551 件，漏洩人数は 628 万 4363 人に登る．図 2.5 に表として示した．

表 2.5: 2011 年 個人情報漏洩インシデント 概要データ

漏洩人数	628 万 4363 人
インシデント件数	1551 件
想定損害賠償総額	1889 億 7329 万円
一件あたりの漏洩人数	4238 人
一件あたりの平均想定損害賠償額	1 億 2810 万円
一人あたりの平均想定損害賠償額	4 万 8533 円

また，漏洩媒体と経路に関して，一般財団法人日本情報経済社会推進協会（JIPDEC）による平成 19 年度「経済産業分野の事業者における個人情報の保護に関する取組み実態調査」報告書へ掲載されている個人情報漏洩の内訳を図 2.6 に示した．

表 2.6: 個人情報漏洩内訳

内容	割合
紙媒体	70.5%
USB 等可搬記録媒体	10.5%
電子メール	7.4%
インターネット	4.5%
PC 本体	4.0%
その他	2.4%
不明	0.7%

この報告書によると，紙媒体による情報漏洩が 70%を越える割合で最も多いものの，USB や電子メール，インターネットといった PC を用いた情報漏洩も多く発生していることがわかる．

2.4.2 企業から発生した事件

そして，PC を用いた情報漏洩のうち，主に企業から個人情報が漏洩及び流出した事件において，情報の件数が多く，また特徴が違う数種類の事件について以下に挙げる．

- **Yahoo!BB 顧客情報漏洩事件**
  2004 年 2 月 27 日に発生した，Yahoo!BB の登録者約 450 万人の個人情報が漏洩した事件である．この情報に対し，Yahoo!BB へ現金を要求していた犯人が 2 人逮捕されている．この事件においてソフトバンク BB の公表した被害額は 100 億円を超えた．この事件における漏洩経路は外部からの不正アクセスとされている．

- **大日本印刷個人情報流出事件**
  2007 年 3 月 12 日に発生した，大日本印刷の持っていた他企業における個人情報が流出した事件である．この事件は，業務委託先の元社員に顧客情報が持ちだされていており，43 社分 863 万 7405 件分であったとされる．そしてこれらは詐欺グループに売り渡されていたことが発覚し，その損害は 667 万 2989 円分であった．

- **PlayStation Network 個人情報流出事件**
  2011 年 4 月 21 日に発生した，PlayStation Network 利用者 7700 万人の個人情報が流出した事件である．外部からの不正侵入によりシステムがダウンすると同時に個人情報が流出した．1 週間遅れの公表というソニーの対応に対し様々な意見が飛び交った．また，この事件には米国ハッカー集団の Anonymous が関与している可能性があるともされている．

- **スクウェアエニックスオンラインショップ顧客情報流出事件**
  2012 年 9 月 13 日にスクウェアエニックスのオンラインショップへ不正アクセスがあり，顧客情報が流出した事件である．サーバが外部より攻撃を受けたものであり，スクウェアエニックスは同日付でオンラインショップを停止，以降閉鎖とした．

これらの事件において，主に重要とされている個人情報は氏名，年齢，住所，生年月日，クレジット番号等とされている．これらの情報において，まず氏名やクレジット番号は個人に直結する情報なのでとても危険度が高く，そして住所も世帯として絞られてしまうために危険であるとされている．しかし年齢や生年月日は日本において他人と被る可能性がとても多く，危険度は低いとされている．

2.4.3 パーソナルページから発生した事件

上記の事件が企業を中心に発生した事件であるが，逆にパーソナルページから事件へと発展した例もある．それらは主に炎上事件と呼ばれるが，以下に特徴的な事件を挙げる．

- **飲食店アルバイト動画炎上事件**
  2007 年 11 月 30 日，牛丼チェーン店に勤めるアルバイトの少年が，勤務時間外に店舗の食材を使って規定違反を犯していたことを撮影し，動画サイトにアップロードしたことでで炎上した事件である．また，弁明の方法が人々に理解されず，少年のパーソナルページから様々な情報が詮索され，最終的に少年は所属していた学校から退学をするまでに至った．

- **飲食店アルバイト虚偽書き込み炎上事件**
  2007 年 12 月 5 日，ファーストフードチェーン店に勤めるアルバイトの少年が，店舗の機材を使い不衛生な行為をしたと mixi に書き込み，その不適切さに反応した人々によって炎上した事件である．本人及び店舗は虚偽の書き込みとし，事実無根であるとした．しかし疑問の声は止まず，少年は保護者同伴で店舗へ謝罪した．

- **アイドルブログ炎上事件**
  2008 年 1 月 14 日に放映された TV 番組内で，出演していたあるアイドルが不適切な発言をしたことを発端に，そのアイドルのブログが炎上した事件である．炎上の内容は主にテレビ番組での発言に対してのものであったが，時間が経つとアイドルに対する誹謗中傷といった内容も現れ始め，結果的にアイドルが謝罪，所属していたグループを脱退するまでに至った．

- **芸人名誉毀損一斉検挙事件**
  2009 年 2 月 5 日，ある芸人が殺人事件に関与したという前提でブログのコメント欄に誹謗中傷する書き込みをした者を警視庁が検挙した事件である．芸人はその事件に対し無関与であると表明している．この事件はパーソナルページの所持者である芸人自体は事件の起爆剤になってはいないものの，炎上の場所がパーソナルページであったものである．

これらの事件は，主に個人の発した情報と場所を元に誹謗中傷を受けることが問題になったものである．余計な情報の発信がなければパーソナルページが荒れることはなく，また個人に繋がる情報の公開量が少なければ正しい対応で穏便に済んでいた可能性も高い．

2.5 本論文での着眼点

本章で述べたことから，まずは多くのデータから個人の特定は匿名化手法によって防ぐ努力がされているが，社会には悪意のある者が個人情報を求めており，またパーソナルページから個人情報を収集し情報主に対し影響を与える炎上事件が発生していることがわかる．このような炎上事件を防ぐために，パーソナルページにおける情報公開の方法及び内容には注意を払わなければならない．本研究では，個人の特定につながってしまう情報の種類を考察すると共に，現状のパーソナルページにおいて危険度がどの程度含まれているかを調査，そして匿名度として表現する手法を提案する．

2.6 まとめ

本章では，プライバシの定義と個人情報の定義，またデータ保護に関する匿名化手法やパーソナルページを持つソーシャルメディアの代表例を示し，そして近年発生している事件を挙げプライバシが脅かされる危険性を述べた．本研究では，パーソナルページにおける個人情報管理の重要性について注目し，事件の発生を防ぐため，匿名度評価手法の提案を目指す．

第3章　関連研究

本章では，プライバシを守るために行われている既存研究を述べる．また，個人情報を公開にするにあたっての既存研究についても述べる．

3.1　プライバシに関する研究

ここでは，プライバシを効率的に守るために行われている方法について述べている既存研究について分析し，記述する．

3.1.1　k-匿名性を利用した例

村本らの研究 [18] や，Latanya Sweeny[19] による研究において，k-匿名性を利用したデータの保護方法が述べられている．

マイクロデータの保護のため，氏名や電話番号，住所といった個人に対するユニークなデータは公開情報から削除されていることが多いが，これらの研究では，その作業によって本当に個人のプライバシー情報を推測することが不可能かという点に焦点をあてている．

また，これらの研究にて，k-匿名性を利用したデータ保護の長所と短所が述べられている．k-匿名性を利用することは，特異性のあるデータ項目は削除し，一般化されているデータとして匿名化処理を施すことができるので，長所としては個人を確実に他の k-1 人に紛れさせられることである．公開されたマイクロデータにおける，ある要素は少なくとも k-1 人存在する．よって，可能性としても個人特定は 1/k となる．そして，短所としては同種攻撃と背景知識攻撃が可能である点である．同種攻撃とは，一般化されたデータについて，もし一般化される前のデータを知っていた場合にどのグループに属しているのかが推測できてしまうために行われる攻撃である．そして背景知識攻撃とは，グループの特徴がとある団体の特徴と被っていた場合に推測できてしまうため行われる攻撃である．

3.1.2　l-多様性を利用した例

そして，Ashwin Machanavajjhala らの研究 [20] では，k-匿名性利用したデータ保護よりも更に安全度を高められるものとして，l-多様性を利用したデータ保護を挙げている．l-多様性を利用すると，k-匿名性を利用した際の欠点を補うことができる．あるグループに

おいて確実に l 種類以上の組み合わせが存在するということは，同種攻撃や背景知識攻撃における個人の推測特定を避ける事ができる．l-多様性を利用した際の欠点は，もしデータ全体において，例えば A グループの可能性が 99%，B グループの可能性が 1%であったとすると，グループ分けに極端に偏りが出てしまうことである．

k-匿名性及び l-多様性を利用したデータ保護から，マイクロデータ内にてある組み合わせを匿名化させる場合の特徴が示されている．本研究では，これらのデータ保護技術から，ある組み合わせをマイクロデータに紛れさせる手法を確立させる．

3.2 個人情報の組み合わせに関する研究

ここでは，個人情報の組み合わせにより実現されることについて述べている既存研究について分析し，記述する．

3.2.1 医療記録と投票者名簿の例

Latanya Sweeny の研究 [21] によると，当時の医療記録と投票者名簿から，マサチューセッツ州知事の医療記録を特定することができたとされる．また，1990 年におけるイギリスの国勢調査では，約 87%の人において，郵便番号や性別，誕生日等の情報から特定可能であったとされる．図 3.1 にイメージ図を示した．これらはプライバシ保護データマイニングの先駆けとなる研究とされ，一見それぞれに個人特定の意味を持たない情報を組み合わせることでの個人特定が可能であることを示している．この研究成果は現在においても生かすことができ，個人の持つ複数のパーソナルページがひもづけられる可能性を示唆している．

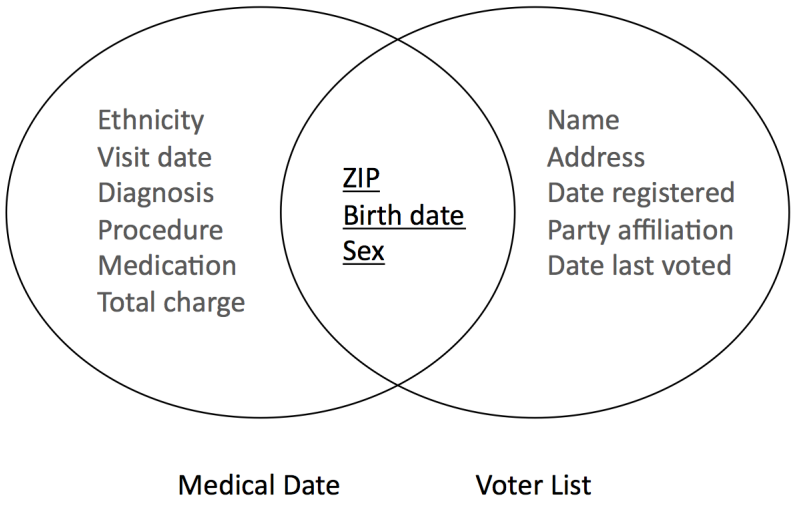

図 3.1: 医療記録と投票者名簿から特定可能

3.3 パーソナルページのセキュリティに対する意識に関する研究

ここでは，パーソナルページのセキュリティに対するユーザの意識を対象に行われている既存研究について分析し，記述する．

3.3.1 Facebook のプライバシ設定に関する例

Yabing Liu らの研究 [22] によると，彼らの調査内において Facebook を利用しているユーザの 36%がプライバシ設定についてデフォルトのままであったとされている．また，それぞれのユーザが公開している情報の範囲について，37%のユーザが想像の範囲を超えた人々にまで自分の個人情報を示してしまっていると回答したとされている．これらからパーソナルページを扱うユーザのプライバシに関する意識が高くないことが明らかになった．

3.4 ソーシャルメディアにおける事件に関する研究

ここでは，ソーシャルメディアでの事件を防ぐために行われている既存研究について分析し，記述する．

3.4.1 規範の概念に着目した例

平井の研究 [23] によると，ブログ炎上事件が発生する理由として主に注目すべき点は，利用するユーザ同士における規範であるとされる．不祥事に対し批判的なコメント等が殺到，そして対象となる人物が糾弾される様子は，規範の概念で補うことで包括的に捉えることができる．規範とは，日常での道徳やマナーといった社会的常識とは区別されるが，その規範に反してしまう意思は対象となる人物において無いことが特徴である．SNS における炎上事件でも同じことが言え，社会的な背景があるからこその事件発生であると筆者は結論づけている．

社会的な背景を元にしたこの研究とは違い，本研究では技術的な面からの解決を目指すが，炎上事件の中心となる人物において事件を起こしてしまう意思がないという点においては同じ視点であると言える．この無意識の事件発生をあらかじめ防ぐためにも，潜在する危険の顕在化が必要である．

3.4.2 マンガを教材とした例

折田らは自身の研究 [24] にて，Web サイト炎上事件を題材に，学生を対象にマンガを教材としてプライバシ教育を実施したとされる．主にマンガを教材とする教育の評価に対

する研究ではあるが，この中ではソーシャルメディアをサービスの理解だけではなく，作る側と利用する側のメタな理解を基にし，意思決定場面における問題発見及び問題解決能力の習得を目指している．この研究では炎上事件において必要とされる思考能力が養われると考えられるが，結果としては主にエンジニアへの共感が前提となり，経営者に対してネガティブな評価に偏る傾向が見られたとされる．

ユーザに対し危機管理能力の注意喚起をするにあたり，広い視野を持ち思考を巡らせることは重要と取れるが，本研究においてはパーソナルページにおいて必要とされる情報部分にだけ着目するため，敢えて潜在する危険性にのみ焦点を当てることとする．

3.5 まとめ

本章では，既存の匿名化技術や個人情報の持つ重要性について本研究において関連の深いものについて述べ，またパーソナルページを扱うユーザのプライバシに関する意識が高くないことを示した．そして炎上事件の注意喚起をする研究を取り上げたが，本研究においては違う視点から着手することを述べた．次章では，本章で述べた問題点を解決する手法を提案する．

第4章　事例分析

一般ユーザにおけるパーソナルページの増加から，なんらかの不祥事をきっかけに爆発的な注目を集めてしまう炎上事件が発生していることを示した．また，プライバシ保護に努める技術がある一方，パーソナルページごとにおけるプライバシ保護に対する意識が高くない現状も関連する研究より明らかになった．本章では，これらの問題を解決するにあたり，まずはインターネットを扱うことで到達することが可能な個人情報の範囲を検証した．

4.1　公的な機関から公開されている情報

本研究では，はじめに個人情報に焦点が当てられている事例に着目する．このような事例が発生した際，安全性に不安を持つ世論が大多数を占める．その不安に対し，二つの事案において数字としての割合を算出することでの可視化を目指すため，インターネット上にて取得することのできる情報を元に検証を行う．

本章では，まず公的な機関において公表されている資料を元に検証が行うことのできる事例を扱う．

4.1.1　武雄市立図書館とCCCの提携に関する事例

代表的な例として，ここでは2012年5月に行われた，佐賀県武雄市立図書館とカルチュア・コンビニエンス・クラブ株式会社（以下，CCC）[25] による，企画運営に対する提携基本合意の事例 [26] を取り上げる．この事例は，公的機関である武雄市立図書館の企画運営に対し，CCC という民間企業へ委託をしたことが全国初として注目を浴びた．民間企業が公的機関の個人情報を扱うことに対する不安の声が多かったが，武雄市個人情報審議会で議論された結果，情報収集に対し問題はないとされ，また保存する利用情報は年齢，性別，住んでいる町名の三種に限るとされた．

本研究では，この三種の個人情報からどの割合まで個人を絞ることができるのかを検証する．

4.1.2　事例分析

上記した通り，利用情報が年齢，性別，住んでいる町名の三種となっているため，ここでは佐賀県若木町に住む20-24歳の男性と仮定して算出を試みる．武雄市における統計情

報 [27] によると，武雄市には 51453 人の住民がおり，また 20 歳から 24 歳の男性は 399 人である．実に人口に対して 7.8%の割合となるが，これが一定と仮定し 947 人の住民がいる若木町に当てはめると，該当する人物は 15 人まで絞ることができる．

元浦安市立中央図書館長である常世田の著書 [28] によると，現在の日本における図書館の利用率は 10%から 20%とされており，武雄市においても同様の割合であるとすると，先述の 15 人のうち 2 人から 3 人程度まで特定が可能であることを示している．

4.2 民官の企業から公開されている情報

次に，民間の企業において公表されている資料を元に検証の行うことのできる事例を扱う．

4.2.1 救急車の誤出動に関する事例

次の例として，2013 年 1 月に横浜市港北区で発生した，救急車の誤出動に関する事例を取り上げる．この事例は，横浜市港北区に住む 60 歳の男性が 119 番通報にて救急隊の出動を要請したところ，通報者の発言が聞き取りづらかったこともあり苗字しか確認することができず，通報内容を元に救急車が出動したものの，誤って 2 軒隣の同姓別人宅へ到着してしまった内容である．救急車へ行き先を入力した管理センターの担当は，すぐ間違いに気づき行き先を入力しなおしたものの，救急車へは反映さず，結果的に救急車は行き先を間違えてしまったものである．具体的に技術的な内容をここでは取り上げないものの，元々同姓の別人宅が近くに存在しなければこの誤出動は発生しなかったと考えられる．

本研究では，この事例を参考にし，都道府県別の姓名における割合から，個人を特定することの割合を検証する．

4.2.2 事例分析

まず，神奈川県に住み姓名を持った人物と仮定する．2001 年発刊のハローページを元に作成された，同姓同名辞典 [29] によると，神奈川県にある姓名は 1208042 種，そのうち同姓同名は 188538 種であり，割合は 15.6%となる．救急車の出動に対し姓名だけの情報と仮定すると，通報の 15.6%は判断のつかないものとなり得ることがこの算出よりいえる，

4.3 まとめ

本章では，ソーシャルメディアを扱う以前に、インターネットを扱うことで取得することの出来る情報から，注目された事例を元に分析，その割合を算出した．武雄市が市民図書館の企画運営を CCC に委託した問題における，CCC が扱うことのできる個人情報からは，仮定上で 15 人まで絞ることができることが明らかになった．また，救急車が誤出動

した問題における，神奈川県での同姓同名についての割合は，15.6%の可能性があることが明らかになった．次章では，実際にソーシャルメディアにおける個人情報を対象に検証を行い，匿名度の評価手法を提案する．

第5章　提案手法

本章では，実際にパーソナルページにおける情報公開に関する危険性を顕在化させるため，匿名度の評価手法を提案する．

パーソナルページの代表例として，ここではまずTwitterを取り上げる．図5.1にTwitterの利用画面を示した．Twitterには登録名及び表示名，bioと呼ばれるユーザ自身の紹介部分となる自由記述欄，住んでいる場所，そしてユーザによるツイートの表示されるタイムラインが主に情報を公開する場所として存在している．Twitterは本名登録が必須ではないので，登録名および表示名には様々な名前が記述されており，本名を使っているユーザは少ない．またプロフィール欄には主に所属や趣味，参加している活動といった内容を書いている場合が多く，例えば学生はこの欄に所属学校やサークル団体といったものを書くユーザが多い．住んでいる場所には地域名を書くことが一般的ではあるが，それも決して強制ではないため，大喜利のような扱いをしているユーザもいる．そしてタイムラインにはツイートと呼ばれる各々の自由なつぶやき発言が記録されていく．今何をしているのか，どこにいるのか，誰と居るのか，といった5W1Hの情報からどのようなことを考えどのような話をするのかに至ってまで全て自由である．制限は一度のツイートに140文字までしか使えないということだけであり，もしそれ以上ツイートをしたい場合は二度に分けてつぶやくユーザも見かけられる．

そして，Twitterに続きパーソナルページの代表例としてFacebookを取り上げる．図5.2にFacebookの利用画面を示した．Facebookは実名登録が基本の言わばインターネット上での名刺のような役割を果たすパーソナルページを所持しており，名前や生年月日，出身や所属など，自身に関する情報を細かに発信しているユーザが基本となっている．実名制であることに起因し，Twitterに比べてプロフィールには真面目な内容を記入がされている場合が多い．出身学校を選択する機能も実装されているため，ユーザは過去の友人たちと再び触れ合うことの出来る場としてもFacebookを利用することができる．

5.1　手法概要

本研究では，個人情報がどのような量で公開されているかを調べるにあたり，このTwitterにおいて二つの手法を施し，そしてTwitterとFacebookを関連付けた手法を用いることでそれを実現する．

図 5.1: Twitter 表示画面

5.1.1 プロフィール欄の個人情報

まず，Twitter のプロフィール欄として登録名及び表示名，自由記述欄，住んでいる場所において記述されている個人情報について着目する．プロフィール欄の中で，特に自由記述欄には最大 160 文字の言葉を書くことができるため，名前や住んでいる場所の情報に加え自由な単語の組み合わせが発生することになり，それらの組み合わせを捉えるとそれぞれのユーザはおよそ確実に世界においてユニークなものとなる．本手法では，その単語の組み合わせはどの程度からユニークなものとなるのかを，それぞれのユーザにおいて類似度を調べることで検証する．呼称や地名，団体名といった，プロフィール欄に用いている言葉を自由に組み合わせ，キーワードによる検索を行う．それぞれのユーザの持つユニークとなりやすい単語，また単体ではユニークとなり得ないものの，組み合わせによってユニークとなってしまう単語を調べ，外部からの単純な閲覧によるパーソナルページの匿名度を評価する．

この手法は関連研究内にて示した，プライバシ保護データマイニング，特に k-匿名性を利用した考え方に準じているものである，k-匿名性は多くの組み合わせによってユニークとなってしまうことを防ぐための手法であるが，本手法では敢えてユニークとなってしま

図 5.2: Facebook 表示画面

う組み合わせを現していくことによって匿名性を評価することができる．

5.1.2 タイムラインの個人情報

次に，Twitter のタイムラインにおいて記述されている個人情報について着目する．それぞれのユーザにおけるタイムラインから個人情報を探し出し，ユーザの特徴を抽出することができるかを検証する．それぞれのツイートを言語解析し，様々な単語及びその単語の出現頻度を抽出する．先述の通りタイムラインには 140 字以内という制限があるのみで，他には自由な内容でツイートを投稿することができる．つまり内容には個人情報が含まれている可能性が大いにあり，この手法によりそれらを抽出することによって，個人の特徴として明らかにすることができる．

本手法によって，それぞれのユーザ像というものを機械的に割り出すことができるが，割り出しやすいユーザやその逆のユーザが存在すると考えられる．各々のユーザにおけるパーソナルページの特徴を明らかにすることで，一般的なパーソナルページにおける匿名度を評価するための指標が生まれると言える．

また本手法は，次に行う他ソーシャルメディアにおけるパーソナルページとの紐付けに関する検証と，同時に扱うことによっても価値を増すことができる．

この手法は関連研究内にて示した，ユーザのパーソナルページにおけるセキュリティ意識が高くないことを踏まえて行なっているものである．セキュリティ意識が高くないことから，タイムラインにおいても情報の操作を意図的に行うことはないと仮定され，多くの個人情報を抽出できると考えられる．

5.1.3 他サービス間のパーソナルページに対する紐付け

本手法では先述の Twitter に加え，同じパーソナルページを持つサービスとして Facebook を関連付ける．Twitter と Facebook といった異なるサービスにおけるユーザアカウントに対し，それぞれのパーソナルページにおける紐付けの可能性を検証する．Twitter におけるタイムラインの個人情報を扱った手法によって明らかになった，キーワードとなる単語を元に Facebook にて登録されている個人情報へのアプローチを試みる．

Twitter は実名の登録こそ強制ではないものの，様々な個人情報を顕在化することができると考えられる．また，Facebook は実名が基本となっているものの，ガイドライン通りに登録する居住地や所属以外の情報を詳しく表示していないユーザも多い．Twitter で抽出した個人情報を元に，Facebook における同一ユーザの発見を試み，ソーシャルメディア間におけるパーソナルページに対する紐付けの可能性を明らかにする，本手法によって，例えば Facebook においてあまり更新を行なっていないユーザの思想や活動といったものが，Twitter から明らかになってしまうことも考えられ，パーソナルページにおける匿名度を評価するにあたっては重要な手法であると言える．

この手法は関連研究内で示した，あるデータテーブルの個人情報から違うデータテーブルにおいて，秘匿されている個人情報を導き出せる可能性を示していることに基づいた手法である．Twitter から Facebook という他ソーシャルメディアのパーソナルページにアプローチを掛けることで，より詳細な個人情報を収集できる可能性があると考えられる．

5.2 要求事項

パーソナルページにおける匿名度を評価する手法を提案する上で，いくつかの要求事項がある．ここではその要求事項について詳しく述べる．

5.2.1 ユーザに規則性が無いこと

関連研究内で挙げた平井の研究 [23] にもあるように，人間はソーシャルメディアにおいても社会性を持った行動をとる，本提案手法にて扱うユーザも例外ではないため，特に Twitter におけるタイムラインには同コミュニティ内のユーザ同士に類似性が見られる．よって，タイムラインの個人情報を扱った手法を用いる際，扱うユーザに規則性が発生してしまうと，その内容にも同様に規則性が発生してしまう可能性がある．抽出した特徴

は，パーソナルページに対する紐付けを行う手法でも扱う．よって，内容は乱雑なものとなっていることに意味があるため，本要求事項を必要とする．

5.2.2 ユーザに対する前提知識が無いこと

パーソナルページに対する紐付けを行う手法を用いる際，ユーザに対して前提知識となるものが存在してしまうと，検証の可能性に影響を与えてしまうことが考えられる．タイムラインの個人情報を扱う手法によって明らかになった頻出単語を用いての紐付けを行うが，ここで扱う単語を選定し Facebook にてパーソナルページを探す際，ユーザに対しての前提知識が働くことで容易な紐付けが行えてしまう．飽くまで明らかになった単語を用いてのみの検証を行うことに意味があるため，本要求事項を必要とする．

5.3 まとめ

本章では，ソーシャルメディアでの情報公開の危険性を顕在化させるため，パーソナルページにおける匿名度を評価する手法を提案した．その内容は Twitter のプロフィール欄を扱った手法，及び Twitter のタイムラインを扱った手法，そして Twitter と Facebook のパーソナルページにおける紐付けを行う手法となっている．そしてそれらの手法を行う際に必要とされる要求事項についても述べた．

第6章　実験と評価

第5章で述べた手法を用いて，実際にソーシャルメディアのパーソナルページにおける匿名度を評価した．本章では，各手法ごとに行った実験とその結果，またそれに対する評価を述べる．

6.1　実験環境

本提案手法において検証実験を行う際に扱った，実装環境及び実験対象を表6.1及び表6.2に示した．

表 6.1: 実装環境

要素	属性	利用環境
OS	実機	MacOS X Snow Leopard
言語	実機	Ruby 1.9.3
		PHP5
ツール	言語解析	MeCab 0.994

言語解析に使用したMeCabは，京都大学情報学研究科と日本電信電話株式会社コミュニケーション科学基礎研究所の共同研究ユニットプロジェクトを通じて開発されたオープンソース形態素解析エンジンである．

表 6.2: 実験対象

性別	人数	属性	人数
男性	14名	社会人	5名
女性	6名	学生以下	15人

実験対象としたアカウントは全20アカウントとし，これらは筆者のアカウントを除き無作為抽出によって選出されたものであるが，条件として日本語での利用を行なっているユーザであること，Twitterに於いて1600件以上のツイートを行なっているユーザであることが挙げられる．

6.2 プロフィール欄の個人情報

ここでは，プロフィール欄の個人情報を扱った匿名度評価の手法について実験を行い，その結果から得られる評価の内容を述べる．

6.2.1 実験内容

Twitter Search API を用い，Twitter における登録名及び表示名，自由記述欄，住んでいる場所に対してのキーワードによる検索を，全世界のユーザに対して実行可能な状態とした．検索に用いる単語は 1 つから複数個まで対応させ，同じ単語をプロフィール欄に持ったユーザが存在した場合に，そのユーザのプロフィールと人数を表示する．実験対象としたそれぞれのユーザごとに，プロフィール欄に持つ単語を様々な組み合わせで検索を掛け，全世界でユニークとなるまでに使用した単語の数を記録する．この数が多いユーザは全世界においてユニークとなりづらい，つまり匿名度の高いプロフィール欄を持つユーザと言える．

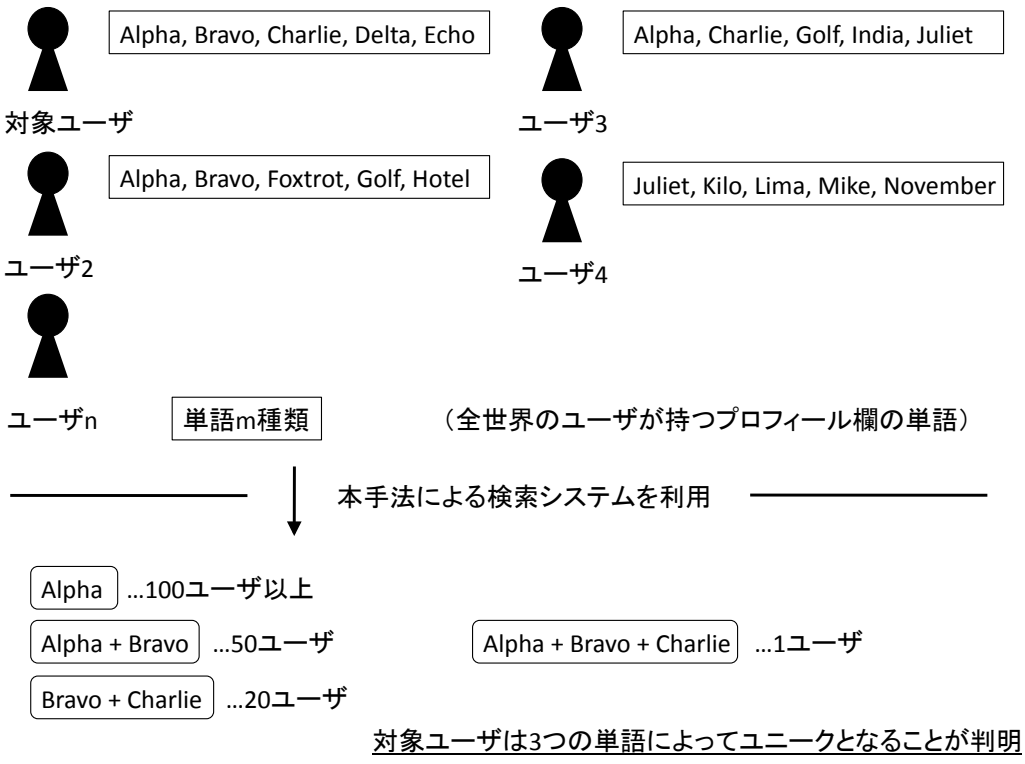

図 6.1: プロフィール欄の個人情報を利用した手法概要

6.2.2 実験結果

実験結果の一例として，筆者のアカウントにおいて行ったものの結果を表 6.3 に示す．なお，掲載している検索結果は人数が 100 人を下回った組み合わせのみである．

表 6.3: プロフィール欄に対する実験結果

検索キーワード	人数				
れもん+湘南+野球	6 名	れもん+湘南+音楽	32 名		
湘南+野球+音楽	27 名	れもん+湘南+野球+音楽	2 名		
村井研	15 名	武田研	6 名	村井研+武田研	2 名

これ以上の組み合わせを試行すると，検索結果はユニークなものとなる．つまり最短では村井研，武田研，何かという 3 つの単語によって，このプロフィール欄は全世界においてユニークなものとなった．

野球や音楽といった多くの人が持つ趣味に対し，れもんという呼称や湘南という地名を混ぜていくことで人数は大幅に絞られていく．また，村井研や武田研といったコミュニティにおける固有名詞は使用する人が少なく，この場合は研究会という単語を省略して扱っていることで更にその人数を少ないものとしていた．

6.2.3 評価

上記のアカウントにおいて，最短でユニークとなる組み合わせは 3 単語であった．同様に対象の全アカウントに対し実験を試みたところ，その平均の組み合わせ数は 3 単語となった．検索にかけた単語の平均文字数は 3.5 文字のため，160 字の自由記述欄だけを対象に見てもおよそ 50 単語弱書くことができるうちの 3 単語である．また，1 人のアカウントにおいて扱われている単語の平均数は 8 つであり，全アカウントにおいて目的を持った検索をかけることでユニークなものとして検出が可能であると言える．

以上のことから，k-匿名性の考え方より，4 単語以上を持ったアカウントが多数存在する場合において，あるアカウントはその集団の中でユニークとなり得る．また，それらの存在するアカウント集団においての匿名化処理を施す場合は，4 単語を越す数から k-匿名性の変数が 1 つずつ増えていく可能性を示している．

本手法によって，Twitter のプロフィール欄における検索機能を用いた危険性を数字を持って表すことに成功し，匿名度の評価手法として有意なものであることを明らかにした．

6.3 タイムラインの個人情報

ここでは，タイムラインの個人情報を扱った匿名度評価の手法について実験を行い，その結果から得られる評価の内容を述べる．

6.3.1 実験内容

Twitter REST API におけるユーザタイムライン取得のメソッドによって，任意のユーザから 1600 件のタイムラインを取得し，それらに対してリプライ用のアットマークやアンダーバー等の余計な文字を削除，後に MeCab を使用し形態素解析を行った．得られた結果から名詞のみを残し，同じ単語の出現回数を降順に並べた．タイムライン上には様々な文章の形でツイートが行われているが，これによって形式にとらわれず単語として重要なものを抽出し，また出現回数が多ければそれだけユーザにとって重要な単語であることを示している．抽出された単語からユーザ像が作られやすいものは匿名度が低く，逆は匿名度が高い状態と言える．

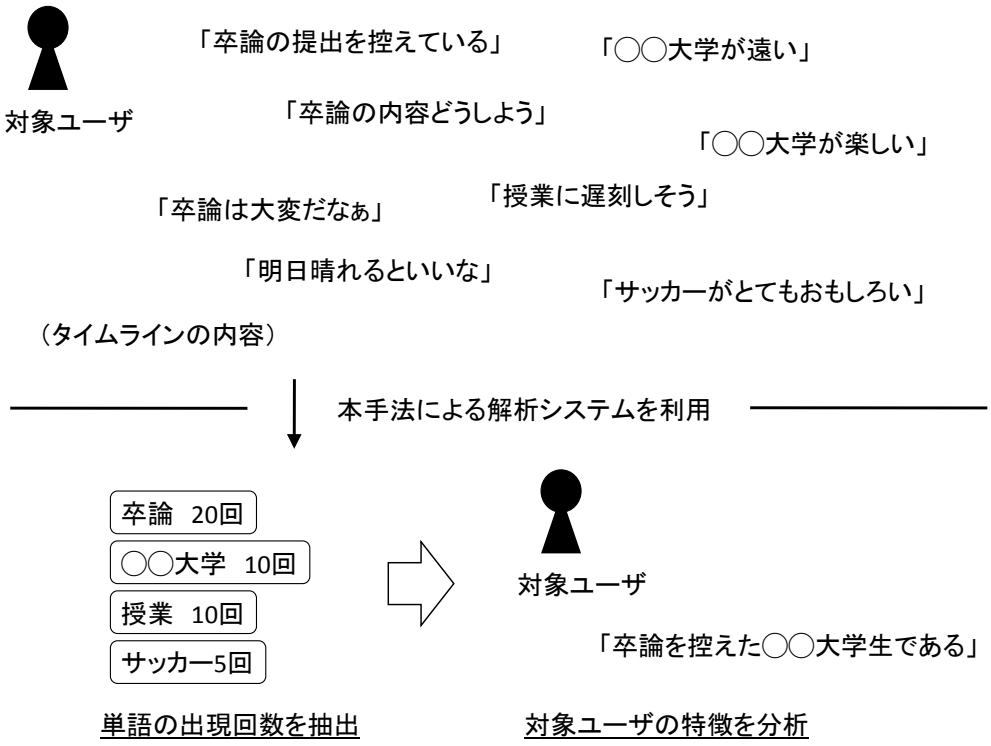

図 6.2: タイムラインの個人情報を利用した手法概要

6.3.2 実験結果

本手法においても実験結果の一例として，同様に筆者のアカウントにおいて行ったものの結果を表 6.4 に示す．なお，掲載している抽出結果は数千件を越えるもののうち，上位 30 件である．

表 6.4: タイムラインに対する実験結果

出現回数	単語				
65 回	人	19 回	時間	13 回	SR
49 回	私	18 回	SFC	12 回	提出
31 回	卒論	17 回	みんな		うち
30 回	今日		学校		車
	おれ		先生	11 回	有馬
28 回	明日	15 回	もの		最終
26 回	発表	14 回	レベル		話
24 回	家		先輩		章
19 回	大丈夫		ISC		夜
	自分		研究		PC

この実験結果より注目するべき重要な単語は，まず 31 回の出現回数を見せた卒論という単語である．この単語を多く扱う層は，日本において大学卒業を控えた学年にいる学生である．次に，18 回の出現回数を見せた SFC という単語，17 回の出現回数を見せた学校という単語である．先述の学生であることを踏まえると，このアルファベットは慶應義塾大学の湘南藤沢キャンパスの略称であることをほのめかしている．更に，14 回の出現回数を見せた ISC，研究という単語に注目し，慶應義塾大学湘南藤沢キャンパスに存在する研究グループの名前であると推測される．最後に，11 回の出現回数を見せた有馬という単語は，上記されていない部分に存在する他の人名単語らと包括し，このユーザの持つ名前である可能性を持つ．

6.3.3 評価

以上の実験結果より，このアカウントのユーザは慶應義塾大学の湘南藤沢キャンパスに所属する 4 年生で，ISC という研究グループに所属する，有馬という名を持つ可能性のある人物であるといえる．この人物像は，筆者に全く同じなものであり，ひとつとして間違っている情報はない．つまり，この実験より筆者のアカウントからは詳しい筆者の人物像を想定することができることが明らかになり，筆者のアカウントは非常に匿名度の低いタイムラインを所持しているということが言える．

筆者のタイムラインにおける実験結果には，出現回数の多い単語の中に重要な語句が大量に含まれていたため非常に匿名度の低いものであるとされた．同様に具体的な語句を多く出現させたアカウントが多くあり，これらはやはりユーザ像の想定が容易であった．関連研究で示されたとおり，ユーザのパーソナルページにおけるセキュリティ意識は高くないことが伺える．逆に，ユーザ像が全く想定できないタイムラインも存在した．それらの特徴として，基本的に趣味における会話しかしないこと，勉強内容や業務内容といった重

要な内容をツイートしていないこと，渾名やハンドルネームといった呼称しか用いず実名を一切出さないことなどが挙げられる．

本手法によって，Twitter のタイムラインから出現回数の多い単語をその回数とともに表示させ，同時にユーザ像の想定を行うことに成功し，匿名度の評価手法として有意なものであることを明らかにした．

6.4 他サービス間のパーソナルページに対する紐付け

ここでは，Twitter と Facebook におけるユーザアカウントに対し，それぞれのパーソナルページを扱った匿名度評価の手法について実験を行い，その結果から得られる評価の内容を述べる．

6.4.1 実験内容

Twitter のタイムラインの個人情報を扱った匿名度評価の手法から得られた，それぞれのユーザにおいて想定されたユーザ像を元に，同一ユーザの Facebook アカウントを探知する．例えば筆者のタイムラインからは，慶應義塾大学の湘南藤沢キャンパスに所属する 4 年生で，ISC という研究グループに所属する，有馬という名を持つ可能性のある人物であるということがわかっている．Facebook においてはまず，名前，学歴，勤務先，場所を用いて検索をかけることができる．この検索のほか，ヒントとなる単語からアカウントを導き出すことができた人数を記録，またその際にも容易に導き出せたアカウント，手順を踏んだものの導き出せたアカウント，近いところまでは導き出せたものの特定まで至らなかったアカウント，遠く及ばず導き出せなかったアカウントをそれぞれ明らかにする．紐付けができるということは，パーソナルページ内の情報がそのパーソナルページにとどまらず，他サービスにおいても活用できてしまうなど，ユーザにおいての様々情報を引き出すための手段となってしまうことが明らかにされることであり，パーソナルページにおいての匿名度が低いことを示している．

6.4.2 実験結果

本手法においての実験結果を表 6.5 にまとめた．タイムラインから想定した人物像で Facebook のアカウントに紐付けることができたアカウントは 13 名，逆に紐付けができなかったアカウントは 7 名となった．達成可能性は 65%となった．

容易に導き出せたアカウントは，Facebook の検索機能を 2 つ程度扱うことで表示されたものであり，手順を踏んだものの導き出せたアカウントは，検索機能を 3 種使う，もしくはヒントとなる単語から所属する団体のホームページを経由するなどして名前を探すなどの作業をしたものである．近いところまでは導き出せたものの特定まで至らなかったアカウントは，上記の手順を踏むことで数人までは対象を絞ることができたものの最終的

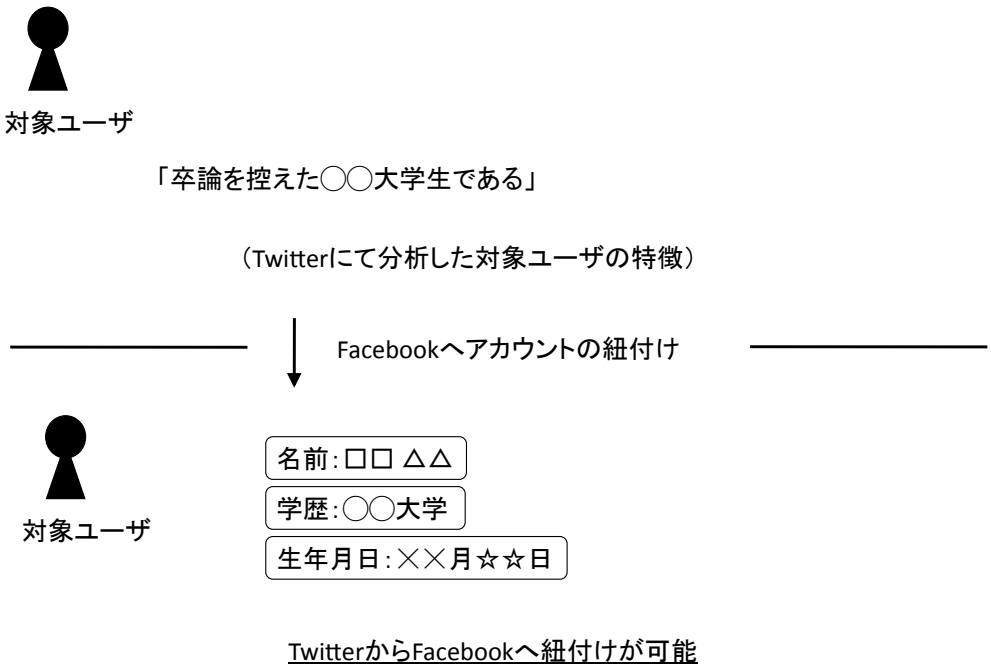

図 6.3: 紐付けの可能性を検証する手法

表 6.5: パーソナルページにおける紐付けの実験結果

可否	状態	人数
可	容易に導き出せた	8 名
	手順を踏んだものの導き出せた	5 名
否	近いところまでは導き出せた	4 名
	遠く及ばず導き出せなかった	3 名

な決め手に欠いたものであり，遠く及ばず導き出せなかったアカウントは，ヒントとなる人物像自体が詳細に想定されていなかったものである．

6.4.3 評価

以上の結果より，Twitter と Facebook のアカウントを紐付ける手法によって，およそ 65%の確率で目標を達成した．この確率は要求事項通りに前提知識のない状態においての

機械的な抽出による実験の結果であり，4 名の最終的な決め手に欠いたアカウントも，タイムラインの詳しい文章等から様々な知識を導入して探すことで目標を達成することができた．つまり実質的な可能性は 85%までのぼる．見つけるために重要であった単語は，実名，学歴が主であり，続いて地名，他ホームページを経由する際には所属する団体名が重要となった．

そして，Twitter には実名を書いていなかったユーザが，Facebook に書いている例を初め，Twitter において見つけることのできなかった詳しい情報が Facebook 載せられていた場合があった．また，Facebook は詳しい情報を載せていなかったユーザが，Twitter 上では色々なツイートとともに詳しい情報を書いていた場合もあり，どちらかに重点を置いて扱っているユーザが，重点を置いていないサービスのユーザから詳しい情報を探知できる可能性があることを示している．関連研究で示されたとおり，あるデータテーブルから他データテーブルの秘匿情報を収集することができることが明らかになった．

本手法によって，Twitter のタイムラインから想定されたユーザ像を元に，Facebook における同一ユーザのアカウントを紐付けることが 85%の確率で可能であることが示され，これも匿名度の評価手法として有意なものであることを明らかにした．

6.5 まとめ

本章では，ソーシャルメディア上のパーソナルページにおける情報公開に関する危険性を顕在化させるため，3 つの手法によって匿名度評価のための実験を行った．まず実装環境と実験対象を示し，そしてそれぞれの手法において実験内容と実験結果を述べた，同時に実験結果に対し評価を行い，それぞれの手法において匿名度評価手法として有意であることを示した．

第7章　結論と展望

本章では，本論文全体についてをまとめ，第1章にて述べた目的において達成状況を示す．そして，本研究の発展を実現するため，今後の展望を述べる．

7.1　まとめ

本論文の目的は，ソーシャルメディア上における個人情報の公開について，潜在する危険性を顕在化させるため，パーソナルページの匿名度を評価することである．これにより現代社会において問題になっている，パーソナルページにおける個人情報を元にした炎上事件等の発生を，予め防ぐためのWebガイドライン作りが可能になる．また，それぞれのユーザにおけるそれぞれのパーソナルページに対して評価を行うことで，ユーザにあった情報操作の必要性を認識することができると期待される．

そこで，本論文では上記のような目的を達成するための手段を3つ提示した．1つ目は，パーソナルページの代表例としてTwitterを用い，プロフィール欄の匿名度を評価する手法である．プロフィール欄に書かれている様々な個人情報から，他ユーザとどれだけ類似しているかを示すため，k-匿名性の考え方を利用し実験を行った．これによって，それぞれのユーザにおけるプロフィール欄の平均単語数が8つであることに対し，平均3単語によってパーソナルページは全世界においてユニークなものとなることが明らかになった．2つ目は，同じくTwitterを用い，タイムラインの匿名度を評価する手法である．タイムラインに書かれている多くの文章から，言語解析を行うことで重要な単語とその出現回数を抽出し，ユーザ像を想定する実験を行った．これによって，多くのユーザがパーソナルページに対するセキュリティ意識は低いことがわかり，同時にタイムラインでの頻出単語を用いユーザ像を想定することができると明らかになった．3つ目は，ここでは代表例としてFacebookも用い，同一ユーザにおけるTwitterのパーソナルページから，Facebookのパーソナルページへの紐付けに対する可能性を検証する実験を行った．これによって，65%のユーザが機械的な学習から紐付け可能であったことを示し，またユーザに対する前提知識を持てば85%までその可能性が上昇することが明らかになった．

これらから，本論文で示した3つの手法を用い，パーソナルページの匿名度評価が可能であること，またそれらが有意であることを示した．そして，本手法によって得ることのできたガイドラインを図7.1に記す．

本論文で提案，実現した匿名度評価手法を用いることによって，パーソナルページにおけるリスク管理の負担を軽減，容易なものとした．個人情報の公開を元に発生する事件はきっかけとなる出来事に注目されがちであるが，事件の発生を防ぐために予め対策をして

おくことで，不祥事が起きてしまった際にも最小限の被害で食い止めることができる．このように，本研究はソーシャルメディアのセキュリティ対策において意義がある．

検索によってユニークとなりづらいプロフィール欄を作る
他ユーザが扱うものと同じような単語を用いる
個性を出す場合は他ユーザに推測されづらい言葉を用いる

個人の識別に繋がりにくい内容を作る
名前，学歴，勤務先，所属，居住地といった具体的な情報を避ける
研究内容や業務内容といった重要な情報を書かない

他サービスにおけるパーソナルページの重要度で内容に差をつける
重要度の低いパーソナルページには特徴的な情報を載せない
他サービスのパーソナルページから紐付けられる可能性を懸念する

図 7.1: 匿名度評価手法によって提案される Web ガイドライン

7.2 今後の展望

本論文全体のまとめから，本研究のさらなる発展のため，今後の展望を述べる．今後の展望については，技術的な側面と手法的な側面の 2 つに分けて述べる．

7.2.1 技術的な側面

まず，技術的な側面における課題としては，それぞれの手法における動作を一貫した自動化を測ることである．現在，プロフィールにおける検索機能を用いる場合はそれぞれの単語を任意で選び組み合わせている状態であるが，形態素解析器を利用し，抽出された単語を自動で組み合わせ結果を表示するものが求められている．また，タイムラインにおける重要単語の抽出においても，一人称や代名詞が含まれている状態であり，より重要度の高い単語が見逃されてしまうことが考えられる．形態素解析の結果から除外する単語を選ぶ際，より条件の厳しいものを設定し，また固有名詞や頻出単語など重要なものを自動で抽出することが必要である．そして，Twitter と Facebook のパーソナルページに対する紐付けにおいても，Twitter のアカウント名を入力することによって自動で近い Facebook のアカウントを出力できるよう自動化するべきである．これらの課題は現在ある技術的に可能なことであるため，今後も実装を続ける必要がある．

7.2.2 手法的な側面

手法的な側面における課題としてはより精度の高い手法を確立することである．パーソナルページにおける危険度を顕在化させる際，現在は形態素解析の結果，名詞のみを扱っている状態である．しかし，人間の特徴は単語だけに現れず，口調や語尾，パーソナルページを扱う時間帯等様々な場所に現れるといえる．よって，今後はこれらの特徴も把握することのできる手法を確立させ，よりユーザに対し注意喚起が促せるようにしなければならない．

謝辞

本論文の作成にあたり，ご指導頂いた慶應義塾大学環境情報学部学部長 村井 純博士，同学部教授 徳田 英幸博士，同学部教授 中村 修博士，同学部准教授 楠本 博之博士，同学部准教授 高汐 一紀博士，同学部准教授 三次 仁博士，同学部准教授 植原 啓介博士，同学部専任講師 中澤 仁博士，同学部準教授 Rodney D.Van Meter III 博士，同学部教授 武田 圭史博士，同大学政策・メディア研究科 斉藤 賢爾博士，同研究科特別研究講師 佐藤 雅明博士に感謝致します．

特に武田 圭史博士に感謝致します．私が本研究会へ所属するきっかけを作って頂き，下級生であった頃から親身に指導をしてくださいました．知識も技術も持ち合わせていない私へ様々な道を示してくださり，おかげで無事本論文を書き終えることができたと確信しております．本当にありがとうございました．

そして，本研究を進めていく上で，様々な励ましと助言，お手伝いをいただきました，村井研究室卒業生である水谷 正慶氏，金井 瑛氏，上原 雄貴氏，重松 邦彦氏，梅田 昇翔氏，福岡 英哲氏，佐藤 文啓氏，相見 眞男氏，Doan Viet Tung 氏，Vu Xuan Duong 氏，Pham Van Hung 氏，吉原 洋樹氏に感謝致します．同時に，慶應義塾大学政策・メディア研究科修士課程，碓井 利宣氏，関根 冬輝氏，山本 知典氏に感謝致します．特に上原 雄貴氏は，右も左も分からない私へ対し真摯に向き合ってくださり，研究活動における基礎を作ってくださいました．また，碓井 利宣氏は研究を行なっていく上で分野の違う私に対して常に気を配ってくださいました．自身の研究やインターンシップによって多忙な中，それでも足繁く研究室へ来ていただき，私の障害を取り払うことに尽力してくださいました．関根 冬輝氏は研究生活におけるあらゆる活動を共に楽しんでくださり，研究室に通うことへの楽しみを作ってくださいました．山本 知典氏は研究会活動をまとめる大変な作業を行う上で尽力してくださいました．そして，実装作業で詰まっている私に何度も手を差し伸べてくださいました．彼らなしで充実した研究生活は無かったと常々感じておりました．

研究会で苦楽を共にした，石野 佑樹氏，大野 三津雄氏，大矢 崇央氏，恩田 優女史，Nguyen Anh Tien 氏，鴻野 弘明氏，小松 真氏，高岡 賢二氏，中島 明日香女史，藤原 龍氏，星出 直柔氏，三ツ木 あかね女史，由井 卓哉氏，吉原 大道氏，生方 悠介氏，小澤 麗女史，Tran Ngoc Anh 氏，露木 航平氏，中安 恒樹氏，深谷 哲史氏，岡田 英晃氏，川本 卓弥氏，芹澤 親氏，廣田 一貴氏，八木橋 優氏，山田 夏才氏に感謝致します．同時に，合同研究会にて支えてくださった，永山 翔太氏，三部 剛義氏，西山 勇毅氏，郭 智洋氏，倉田 彩子女史，小鷲 麻奈美女史，水谷 伊織氏，山岸 祐大氏に感謝致します．彼らと一緒に研究をすることでお互いを刺激しあい，より質の高い議論や研究をすることができまし

た．特にISCの同級生には，何より研究における良き相談役として，レクリエーションにおける良き仲間として，他にも様々な時間を質の高いものへとしていただきました．この場を借りてお礼を述べさせていただきます．

村井研究室秘書の渡辺 康恵女史，比嘉 宏美女史，宇佐美 由美子女史，青木 りか女史，渋谷 雪絵女史に感謝致します．研究会において，書類の提出や発表会の主ケジュール調整など，様々な場面でお世話になりました．ありがとうございました．

私の大学4年間の心の拠り所であった体育会準硬式野球部のメンバー全員に感謝致します．特に同期の池田 圭氏，浦 琢馬氏，郡司 拓也氏，清水 慈氏，武田 悠氏，田中 祐衣女史，中尾 圭佑氏，東 朋彦氏，増田 俊氏，松本 啓志氏は，とても未熟であった私に正面から向き合ってくださり，4年間の成長を支え続けてくださいました．彼らのおかげで余裕を持った学生生活及び研究活動をを行うことができたと確信しております．

学生生活を通じ良き理解者となってくださった，織戸 亜衣女史，笠原 優里女史，鴨見 ルリ女史，久慈 梨絵子女史，小畑 和博氏，斎田 太輝氏，斎藤 紗紀女史，佐藤 智麻女史，佐藤 龍ノ介氏，重田 瑞希女史，鈴木 悠平氏，瀬川 明日奈女史，豊島 有紀女史，三浦 綾子女史，柳生 奏一郎氏，吉田 光女史に感謝致します．私の言動へ真摯に付き合ってくださり，おかげで充実した学生生活を送らせていただきました．

慶應義塾大学湘南藤沢キャンパスの運営へ関わる方々に感謝致します．特に湘南コミュニティの方々にはデルタ館を清掃の面より支えていただき，おかげで快適な研究生活があったと感じております．そして，社長を始めとするデルタ館に住居する猫達が出会うたびに構ってくださったため，荒んだ心を癒すことができました．

最後に，大学入学からの4年間だけでなく22年間をあらゆる面で支えていただいた，父 有馬 公治，母 有馬 恵津子，兄 有馬 怜生，愛猫 めろんとラピス，愛犬 アロハに心から感謝致します．そして，祖父母を始めとした全ての親族の方々に感謝致します．これを以て，謝辞と致します．

参考文献

[1] Facebook Inc. Facebook. http://www.facebook.com, 12 2012.

[2] 株式会社 DeNA. Mobage. http://www.mbga.jp, 12 2012.

[3] グリー株式会社. Gree. http://gree.jp, 12 2012.

[4] Twitter Inc. Twitter. http://twitter.com, 12 2012.

[5] 佐伯胖. 学びとコンピュータハンドブック. 8 2008.

[6] Samuel D. vWarren and Louis D. Brandeis. The right to privacy. *Harward Law Review*, 4(5):193–220, December 1890.

[7] 東京地方裁判所 昭和 36 年 (ワ) 第 1882 号. 「宴のあと」事件, 9 1964.

[8] A consumer privacy bill of rights. http://www.whitehouse.gov/sites/default/files/privacy-final.pdf, 2 2012.

[9] Proposal for a regulation of the european parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (general data protection regulation. http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf, 1 2012.

[10] JIPDEC. プライバシマーク制度. http://privacymark.jp/index.html, 12 2012.

[11] Information and Privacy Commissioner of Ontario. Privacy by design. http://privacybydesign.ca/, 12 2012.

[12] News Corporation. Myspace. http://jp.myspace.com/, 12 2012.

[13] LinkedIn Corporation. Linkedin. http://jp.linkedin.com/, 12 2012.

[14] 株式会社 mixi. mixi. http://mixi.jp, 12 2012.

[15] 株式会社サンロフト. Timelog. http://timelog.jp, 12 2012.

[16] NPO 日本ネットワークセキュリティ協会. http://jnsa.org, 12 2012.

[17] NPO 日本ネットワークセキュリティ協会. 情報セキュリティインシデントに関する調査報告書 個人情報漏洩編. http://www.jnsa.org/result/incident/2011.html, 12 2012.

[18] 村本俊祐, 上土井陽子, and 若林真一. k-匿名性を利用したデータ一般化によるプライバシー保護. *DEWS2007*, 2007.

[19] Latanya Sweeny. Achieving k-anonymity privacy protection using generalization and ssppression. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 2002.

[20] Ashwin Machanavajjhala, Daniel Kifer, and Johannes Gehrke. l-diversity: Privacy beyond k-anonymity. *Journal ACM Transaction on Knowledge Discovery from Data*, 2006.

[21] Latany Sweeny. Guaranteeing anonymity when sharing medical data, the datafly system. 1997.

[22] Yabing Liu, Balachander Krishnamurthy, Krishna P. Gummadi, and Alan Mislove. Analyzing facebook privacy settings: User expectation vs. reality. *Internet Measure Conference*, 2011.

[23] 平井智尚. インターネットにおける「ブログ炎上」に関する一考察. **慶應義塾大学大学院社会学研究科紀要**, 2007.

[24] 折田明子, 吉川厚, and 山本秀男. マンガ教材によるソーシャルメディアのプライバシ教育の実施と評価. **情報処理学会研究報告 マルチメディア通信と分散処理研究会報告**, 2010.

[25] カルチュア・コンビニエンス・クラブ. Ccc. http://www.ccc.co.jp/, 12 2012.

[26] カルチュア・コンビニエンス・クラブ. 武雄市立図書館と ccc の企画運営における提携基本合意. http://www.ccc.co.jp/company/news/2012/20120504_003337.html, 5 2012.

[27] 武雄市統計情報. https://www.city.takeo.lg.jp/toukei/index.html.

[28] 浦安図書館にできること. **勁草書房**, 5 2003.

[29] レオ. 同姓同名辞典. http://www.douseidoumei.net/, 12 2012.