

「個人情報」に関する社会認識の変化に関する研究 —1990年から2016年における三大日刊紙の 社会面記事を手がかりに— (2017)

A study on the change of social awareness on Personal Information
- Based on the social section articles of three major Japanese
newspapers from 1990 to 2016 - (2017)

◎金 秀那¹
Suna Kim

¹ 中央大学大学院文学研究科博士後期課程社会情報学専攻
Graduate School of Literature, Chuo University

要旨…本研究の目的は、日本社会における「個人情報」に対する社会認識の変遷を『読売新聞』、『朝日新聞』、『毎日新聞』の社会面の記事の分析を通じて捉えることである。本研究では、大量の記事を言葉単位で分解し、数値化することによって、研究者の主観介入を抑えた分析を目指した。記事のトピック分類や共起語の関係の解釈を行った結果、「個人情報」が語られる文脈から社会認識の変遷を描写することができた。

キーワード 個人情報 新聞記事 テキスト分析 社会認識

1. はじめに

(1) 研究の背景

近年、注目されているビックデータ分析は、人々が日常で残すすべてのデータを活用し、組み合わせなどの処理を施すことで、新たな価値を持つ情報を創出すると期待されている。その反面、活用されるデータの中には、何らかの加工をすることで、データの主である個人が特定されてしまい、個人情報保護法で定められたデータ活用の範囲を超えているケースもある。日本政府は、このような問題を背景に、2005年に制定された「個人情報保護に関する法律」に関して、2008年と2009年の一部改正を経て、2016年に全面的な改正を行い、2017年5月30日から実施することを決めた。

このように個人情報をめぐる法律や政策は、その時代の社会的要請や影響を受け、さまざまに修正が加えられてきた。その反面、情報技術発展の凄まじいスピードに法律が対応することは難しくなっており、それを悪用した犯罪や対応不足による事故が増えていく今日の状況では、個人情報に対する人々の不安は広がりつつある。その不安を軽減させるための政策的配慮が求められるが、その際、人々がどのように個人情報を認識しているのかを捉えることが重要な課題となってくる。

そこで本研究は、新聞記事から社会的に「共有」される個人情報に対する認識の抽出を試みる。そのために、1990年から2016年までの『読売新聞』『朝日新聞』『毎日新聞』三紙における「個人情報」に関する社会面記事を対象とし、計量テキスト分析を行う。

本研究における新聞記事のテキスト分析は、膨大な情報を時系列的に分析することができる点、過去の新聞記事がその時代に多くの人々に共有される情報源を用いることで、その時代の個人情報に対する社会的認識を調べることができる点で有意義であると思われる。テキスト分析の手法に加えて、言葉のネットワーク分析を用い、各時代の言説空間における比較から、個人情報をめぐる語られ方から読み取れる社会的認識の変化を明らかにすることを目的とする。

(2) 先行研究

個人情報を人々のどのように理解しているかを捉えるために、これまでアンケート調査やインターネット調査が内閣府や自治体によって多数行われてきた。国民を対象とした意識調査には、内閣府や消費者庁が実施した「個人情報保護に関する世論調査 (2003.9, 2006.9)」「個人情報保護法改正に関する世論調査 (2015.10)」などがある。しかし、これらの調査は、その当時の人々の意見を問うものだけで、各時期における質問項目が統一されていないため、比較により変化を読み取ることが困

難である。その一方、たとえば、個人情報保護法施行以降に現れた過剰反応を解消するため、各自治体で行われた個人情報保護に関するヒヤリング調査など（個人情報保護委員会が実施した「個人情報保護に関するいわゆる『過剰反応』に関する実態調査報告書（2010.3）」など）からは、人々の個人情報に対する具体的な認識を読み取ることができるが、行われた地域と対象が限られているという限界がある。

他方で、制度の変遷から社会の認識の変化を捉えようとする研究も多い。たとえば、堀部(1988;2010)は日本社会における個人情報保護の歴史を振り返り、保護政策の変遷をたどることで、法律上におけるプライバシー保護を、特定の時代のニーズに応じる形で変遷してきたとしている。宮下(2010)は、プライバシーの歴史を整理し、近未来の超高度情報化社会においては、どのように政策を適用できるかまた変えなければいけないかを語ることで、新技術が訪れた世界で暮らすための認識転換の必要性を指摘している。本研究では、以上の先行研究と別のアプローチを試みる。具体的には、新聞記事データを用いて、日本国内全体に発信される「共通の情報」を社会認識として捉え、その変遷を分析していく。

2. 調査手法

マスメディアは、これまでの多くの効果研究でもわかるように、人々がものごとに対して情報を入手する際、直接・間接的に何らかの影響を与えている。また、人々の関心を呼び起こすようなトピックを選び、それを取り上げている。このようなマスメディアの特性を考えた際に、日本社会で「個人情報」という言葉がどのようにマスメディア上で語られたかを検討することは、その当時の人々の個人情報に対する認識を捉えることとして有意義である。

具体的な調査対象として取り上げるのは、その発行部数の多さから日本の三大日刊紙と言われる『読売新聞』、『朝日新聞』、『毎日新聞』の記事である。『読売新聞』の「ヨミダス歴史館」、『朝日新聞』の「開蔵Ⅱビジュアル」、『毎日新聞』の「毎索」から、「全国版、東京朝刊」の「社会面」を対象に「個人情報」の1語のみの検索語で、全文検索を行った。また、各データベースによって記事の収録範囲が異なるため、本研究では共通的に記事が出現される1990年から2016年までの期間を分析範囲と設定した。抽出された新聞記事は合計4,003件（『読売新聞』1239件、『朝日新聞』1519件、『毎日新聞』1245件）であった。

まず、①「個人情報」関連記事の全体像と変化の傾向を把握するために、記事数変化の推移を検討する。②その上で、調査対象期間を5年ごとに分け^{注1}、トピック別に記事の変動を比較することで、記事内容の経年的変化を捉える。そのため、本研究では、テキストの計量分析フリーソフト「KHcoder」を用い、抽出される形態素を「名詞」と「形容詞」に限定し^{注2}、クラスター分析を行った。各クラスターに集まった言葉を解釈することでトピックを分類し、各トピックに該当する記事数の推移を検討する。最後に、③時期別に「個人情報」とともに現れる共起語をKHcoderの「関連語検索」機能を用い、時代別に抽出・検討を行う。このことで、「個人情報」という言葉が語られる文脈がどのように変遷していったのかを時系列的に調べていく。

3. 分析結果

(1) 記事数の経年的変化

ここでは、新聞記事数がいかに変化してきたのかについて検討する。各新聞社における個人情報関連記事数の推移を図1に示す。図1をみると記事数は1990年代には大きな変動はみられないが、2000年を前後に徐々に動き始め、2002年著しく増えたことがわかる。この2000年前後は、日本においてインターネット普及率が急増した時期であり^{注3}、それにつれ、ネット上における情報のやりとりへの関心が高まった時期でもある。特に、2002年は住民基本台帳ネットワーク（以下、住基ネットと表記）の稼働を背景に、個人情報保護への関心が高まった時期であり、以前約50件から150件の間にとどまった記事数も約300件ま

^{注1} ただし、1990年代は記事数が少なかったため、1990年から1999年を「1990年代」としてまとめた。また、「2010年代後半」は2015年と2016年だけを対象としている。

^{注2} KHcoderにおける品詞分類の中で、名詞は「名詞一般・サ変名詞・形容動詞の語幹」、形容詞は「形容詞」だけに限定して分析を行った。また、普段検出されない複合語などを強制的に検出するため、「タグ」を作成した。タグには「個人情報」「住基ネット」「背番号」「住民基本台帳」「個人情報保護法」「電話番号」「納税者番号」「匿名化」「情報公開」「情報保護」「住民基本台帳ネットワーク」「ウイニー」「アプリ」「スマートフォン」「監視カメラ」「防犯カメラ」「共通番号」「高齢者」「ろうえい」の単語が含まれている。

^{注3} 総務省の「通信利用動向調査」により、1997年6.4%に過ぎなかった世帯普及率は、1998年11.0%、1999年19.1%と徐々に増加し、2000年34.0%、2001年60.5%、2002年81.4%と、2000年以降日本社会でインターネットが著しく普及していったことがわかる。（<http://www.soumu.go.jp/jphoto/sintokei/statistics/statistics05b2.html>）

で急増している。その後、記事数は2005年「個人情報保護に関する法律」の全面実施を背景に再び急増し、約350件に達している。2005年に頂点に到達した記事数は、減少傾向に変わり、2012年までおさまっていく。しかし、2013年から2016年の間で、再度、増加傾向に戻っている。これはおそらく、この時期においては2016年から全面実施された「マイナンバー制度」をめぐる社会的議論が背景にあると考えられる。

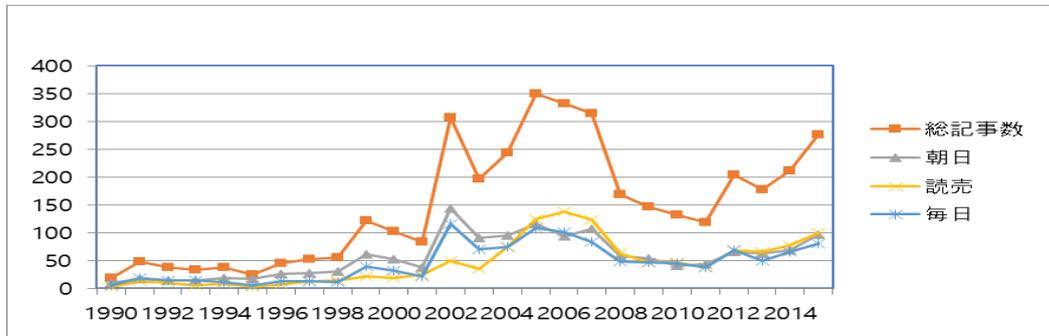


図1 個人情報に関する新聞記事数の推移

記事数の推移から、2002年、2005年、2016年など個人情報関連政策が行われた年度に、記事数も増加している傾向がみられた。新聞が政策実施と連動し、個人情報に関する話題を取り上げることで、読者への注意喚起を呼びかけていることが確認できる。

(2)内容における経年的変化

前節では新聞記事数の変化について検討したが、ここでは個人情報に関する新聞記事の中でどのように語られてきたのかを考察したい。そのために、まず、頻出語を対象に対応分析を行った。その結果を2次元散布図として表したものを図2に示す。

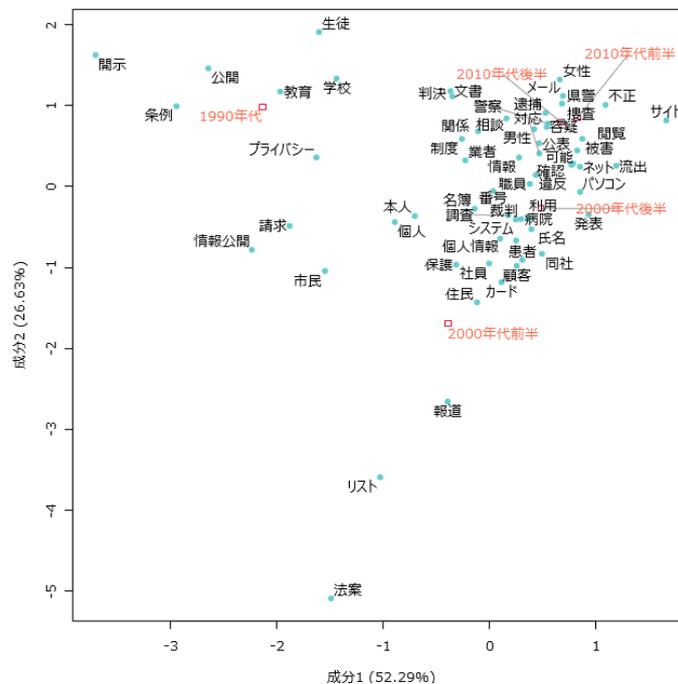


図2 頻出語における年代別対応分析

図2には、それぞれ特徴的な言葉が各年代の近くに示されている。1990年代周辺には、「学校」「生徒」「請求」「公開」「提示」などの言葉が現れている。この年代は、生徒による学校の内申書に関する情報公開や開示の申し出が頻繁に報道され

てきた時期である。2000年代前半には個人情報保護法の制定や実施の前の段階で、それと関連した議論が扱われた記事が多かったため、「個人情報」「閲覧」「行政」などが特徴的な言葉として表示されている。2010年代は、「警察」「逮捕」「容疑」「被害」という言葉が表示されており、個人情報関連の犯罪が頻発に発生していたことが推測できる。このように対応分析結果をみると、各年代別によって語られるトピックが異なっていることがわかる。

そのトピックごとの変動を調べるために、頻出語を対象にクラスター分析(Jaccard 係数^{注4}、Ward 法採用)を行った。各クラスターに集まった言葉からコーディングルール^{注5}を作成し、それぞれのコーディングルールを満たす記事を同一トピックとして分類した。その結果をまとめたのが表1である。クラスター分析の結果、8つのクラスターが得られた。各クラスターに集まった言葉の中から、「技術」「公開」「保護」「管理不足」「利用者管理主体」「不安」「犯罪」というトピックを抽出した。

表1 各トピックと年代におけるクロス分析^{注6}

	技術	公開	保護	管理不足	利用者管理主体	不安	基本権	犯罪	ケース数
1990年代	110 (23.35%)	172 (36.52%)	56 (11.89%)	75 (15.92%)	109 (23.14%)	70 (14.86%)	155 (32.91%)	160 (33.97%)	471
2000年代前半	288 (30.87%)	154 (16.51%)	120 (12.86%)	301 (32.26%)	213 (22.83%)	142 (15.22%)	229 (24.54%)	278 (29.80%)	933
2000年代後半	601 (45.88%)	156 (11.91%)	176 (13.44%)	638 (48.70%)	258 (19.69%)	123 (9.39%)	165 (12.60%)	409 (31.22%)	1310
2010年代前半	408 (48.28%)	106 (12.54%)	105 (12.43%)	422 (49.94%)	183 (21.66%)	109 (12.90%)	102 (12.07%)	412 (48.76%)	845
2010年代後半	186 (41.89%)	57 (12.84%)	50 (11.26%)	243 (54.73%)	95 (21.40%)	57 (12.84%)	49 (11.04%)	209 (47.07%)	444
合計	1593 (39.80%)	645 (16.11%)	507 (12.67%)	1679 (41.94%)	858 (21.43%)	501 (12.52%)	700 (17.49%)	1468 (36.67%)	4003
カイ2乗値	130.630**	173.812**	1.826	243.448**	4.274	20.449**	161.548**	111.053**	

増加傾向を見せたトピックは「技術」「管理不足」「犯罪」、減少傾向を見せたトピックは「公開」「基本権」が挙げられる。情報通信技術の発展とともに、人々の管理不足による個人情報をめぐる犯罪が増えている状況が読み取れる反面、知る権利やプライバシー保護などを背景とする「公開」や「基本権」と関連するトピックは減少していることがわかる。また、「保護」、「管理不足」トピックに対する記事数変動には有意な差はみられなかった。「不安」は2000年後半減少したが、2010年代また少し増加している。

より具体的に個人情報に対する社会認識を把握するために、各時代別に「個人情報」と共起する頻出語を抽出し、その共起語間の関係を、「共起ネットワーク図」から確認し^{注7}、表2にまとめた。

1990年代においては、「開示」「公開」「請求」などの言葉がとともに出現しており、「情報公開」要求などの文脈で多く使われていたことが推察できる。また、当時、学生による学校に対する内申書公開の要請などから「生徒」「学校」「教師」

注4 Jaccardの類似性尺度は0から1までの値をとり、関連が強いほど1に近づく性質を持つ。各文書において、出現する言葉すべての個数を、一致する言葉の個数で割って求められる数値であり、通常0.1以上から関連があると判断する。

注5 各コーディングルールは以下となる。

- ・技術:「インターネット」「ネット」「パソコン」「ソフト」「ファイル」「データベース」「携帯」のいずれかを含む記事
- ・公開:「情報と公開」「開示」「請求」のいずれかを含むか、「情報公開」という言葉を含む記事
- ・保護:「保護」「必要」を同時に含む記事
- ・管理不足:「流出」「紛失」「漏洩」「漏えい」のいずれかを含む記事
- ・利用者管理主体:「会社」「企業」「行政」「機関」「政府」「国」のいずれかと同時に「利用」または「管理」を含む記事
- ・不安:「不安」「危険」「不信」「恐ろしい」「怖い」のいずれかを含む記事
- ・基本権:「自由」または「プライバシー」を含む記事
- ・犯罪:「捜査」「県警」「容疑」「疑い」「逮捕」「違反」「詐欺」「悪用」のいずれかを含む記事

注6 各コードの集計は相互排他的に行われたものでなく、一つの記事が複数のコードを持つこともできるため、各年代におけるトピック別記事の合計は各年代の総記事より多くなり得る。

注7 KHooderの「共起ネットワーク図」の「サブグラフ」方式を選び、共起語の関係を確認した。「サブグラフ」とは、ノード(言葉)同士の関連が強いノードの集合体である。KHooderでは、サブグラフが自動的に色分けして出力されることで、よりノード集合体における意味解釈が容易である。ただし、本稿では紙面の制限上、共起ネットワーク図は示さず、その代わりに色分けされた言葉のグループを表2へ示す。

「教育」など、教育の現場と関係のある言葉の集まりが現れている^{注8}。2000年代前半においては、「住所」「氏名」「生年月日」など個人情報そのものを表す言葉と「政府」「法案」「メディア」「報道」「自由」「規制」など政府の個人情報保護法案発意に対するメディア側の反対が読み取れる^{注9}。

2000年代後半においては、「違反」「容疑」「事件」「県警」「捜査」といった言葉の集まりから、個人情報をめぐって何らかの事件が起きていたことが推測できる。そして「社会」「国民」「検討」からは、個人情報保護法に対する再検討を求める文脈が考えられる^{注10}。

2010年代前半には、「違反」「捜査」「容疑」に加えて「逮捕」という、法的拘束を意味する言葉も現れており、個人情報に対する犯罪がより深刻化されたことがわかる。2010年代後半には犯罪を表す言葉に加え、「不正」「アクセス」「会社」「サイト」「詐欺」「電話」「家族」などから不正アクセスや家族に装った電話詐欺など、その手法も読み取ることができる。

表2 共起ネットワーク図上で分類された頻出語のグループ

時期	頻出語グループ
1990年代	「保護、開示、条例、請求、決定、公開、情報公開、市民」/「要録、評価、同市」「関係、信頼」/「電子、インターネット」/「提訴、地裁、判決」/「申し立て、不服」/「高校、教委」/「全面、調査、答申、会長、教授」/「指導、生徒、教育、学校、教師、市立、中学、委員」
2000年前半	「管理、発表」/「氏名、住所、生年月日、電話番号」/「記者、会見」「住民基本台帳、住民」/「総務、ネットワーク」/「情報公開、請求、リスト、作成」/「行政、機関、対象」/「政府、法案、反対、国会、メディア、報道、自由、規制」
2000年後半	「情報、機関、対応、理由、保護、提供、施行、本人、必要」/「違反、容疑、事件、県警、捜査、警察」/「患者、病院」/「住民、自治体」/「社会、国民、検討」/「教授、指摘」/「インターネット、ネット」/「調査、対象」「ケース、多い」
2010年前半	「サーバー、アクセス」/「入手、業者」/「外部、流出」/「情報、保護」/「教授、必要、指摘」/「文書、資料、内部、国際」/「県警、逮捕、疑い、捜査、容疑、違反、事件」
2010年後半	「不正、アクセス、確認、電話番号、同社、違反、県警、逮捕、容疑、疑い、関係、取材、会社、サイト」/「閲覧、状態」/「プライバシー、地裁、訴訟、判決」/「インターネット、接続」/「特定、企業、政府、機関、対策、ナンバー、社会」/「注意、不審、詐欺、本部、電話、生活、家族、女性、専用、問い合わせ」/「全国、担当、保護、調査、委員、第三者、目的、本人、提供、教育」

4. 考察とまとめ

(1) 考察

では、以上の分析結果を手がかりに、日本社会における個人情報の認識変化について考察していきたい。まず、1990年代の、政府省庁・行政側に対する「情報公開」が多く報道されたことから、その文脈における個人情報の認識を考えていきたい。政府・行政機関が管理している情報を国民に公開することは、国民の「知る権利」を具体化する面で、重要な意味を持つ。この時期よく取り上げられた学校に対する内申書公開への要求は、学生自信の学校生活が学校側によって正当に評価されているのかを確認しようとする目的があり、ここには「知る権利」が裏付けられている。一方、たとえば、役人の汚職事件などにおいて、公的資金の私的使用などの内約を明らかにすることを求められることに対し、個人のプライバシーを理由として部分的公開が認められた事例もみられる^{注11}。つまり、情報公開における個人情報は、国民の「知る権利」と個人の「プライバシー保護」といった基本権の基礎的要素として語られていたと言える。

注8 たとえば、「大阪・高槻の中学女生徒の「内申書公開すべし」個人情報保護審が答申」『読売新聞』東京朝刊1991.03.01 p.30

注9 「表現の自由、制限の恐れ 個人情報保護法、報道への適用<解説>」『朝日新聞』東京朝刊2000.8.5 p.33

注10 「(Media Times Wide) 相次ぐ匿名発表 個人情報保護の「過剰反応」」『朝日新聞』東京朝刊2008.5.11 p.37

注11 「知事交際費を一部公開 市民団体側、内容に不満/愛知」『読売新聞』1997.05.23 東京朝刊 p.30

次に、2000年代は、技術的發展を背景に、2002年の「住基ネット」の稼働及び2005年の「個人情報保護法」施行などが行われた時期であった。2000年代前半においては、各データベースに格納され、ネットワーク上につながっていく個人情報に対する保護法案が立案される中で、個人情報の強制的保護に対し、報道の自由への規制を恐れたメディア側の反発がみられた^{注12}。さらに、2005年個人情報保護法が全面実施された以降は、個人情報をめぐる流出事件が多く報じられている。個人情報保護法を不履行または違反した際には、是正命令や罰金などの罰則を受けることから、法律順守に対する「強制性」を強調されていた。以上から、2000年代において個人情報は、「基本権」保障の要素としての認識に加え、「強制保護の対象」として認識されるようになったと言える。この認識は、2000年代末、各地において問題となった「個人情報保護に対する過剰反応」の原因としても捉えることができる。

2010年代に入ってから個人情報を取り巻く犯罪はより悪質化されており、その手法も様々であった。特にこの時期においては、ソーシャル・ネットワーク・サービス利用が盛んになっており、誰もが自分の情報をネットにのせることが容易になったことで、個人情報の保有先である行政機関や企業だけでなく、個人々が意識を持って自分の個人情報を管理する必要があるということが喚起されていた。

要するに、個人情報に対する認識は、「基本権の要素」から、「法律により保護される対象」としての認識を経て、今日には「個人により管理される対象」として変遷してきた。インターネットの普及を前後にし、個人情報に関する認識が変わってきたことは、この分析を行うまでもなく、簡単にわかると指摘される可能性もある。しかし、本研究ではそのような言説を、新聞記事から頻出する言葉を取り上げることで、あらかじめ個人情報をめぐる社会認識はどのような歩みを経て今日にいたるようになったのかを示すことにその意味が見出されると考える。また提示された結果については、必ずしも明確にその認識の転換が行われたとは限らない。1990年代の基本権の中での競合、2000年代の法律による保護という語りは今日においても依然と社会的課題として残されている。でも各時代において、どのような議論が主に行われ、人々の認識作りに影響を及ぼしたのか、その議論の密度を示すことは可能であったと考えられる。

②まとめ

以上の考察から、以下の知見が得られた。第一に、計量テキスト分析の手法を用いることで、個人情報が語られる文脈およびその社会情勢の変化を新しい手法から描写することができた。第二に、共起ネットワーク分析や関連語検索を通じて、出現頻度から構成される言葉の量的データに基づいた関連性を考えることで、質的にデータを捉えることも可能であった。これまで、個人情報に対する社会認識について、新聞記事を利用し把握しようとしている研究はほとんど見当たらなかったため、新しいアプローチで、以上の認識の変換を示したことから、本研究は有意義であると思われる。

しかし、残された課題もある。本研究においては、「個人情報」だけに焦点を当てて分析を行ったが、もともと個人情報とは個人のプライバシーから個人データ、個人情報として派生された言葉であるため、両方の認識を視野に入れた研究も必要とされる。また、新聞記事を利用することに当たって、妥当性の問題が残される。今回の分析は27年の長期間の時期を分析対象としており、その資料入手の容易性から新聞を選んだ。しかし、新聞は編集側の手によってフィルタリングされた現実であるため、必ずしも社会認識をそのまま反映しているとは限らない。他の資料、たとえば雑誌、またはネット掲示板などの書き込みのように他の言説空間との比較を行うことで、妥当性の向上を図っていくことが今後の課題となる。

参考文献

- 1) 樋口耕一(2014): 『社会調査のための計量テキスト分析—内容分析の継承と発展を目指して』、ナカニシヤ出版
- 2) 堀部政男(1988): 『プライバシーと高度情報化社会』、岩波新書
- 3) 堀部政男ほか(2010): 『プライバシー・個人情報保護の新課題』、商事法務
- 4) 宮下紘(2011): 「プライバシー・個人情報の新世代」、『駿河台法学 25(1)』、p111-136

注12 「個人情報保護法案 800人が「廃案を」——東京・有楽町」『毎日新聞』東京朝刊 2002.09.29 p.30