

日米中韓における情報倫理の比較研究 — personal information を題材として —

A Comparative Study of Information Ethics in Japan, the U.S., China, and Korea: the Case Study of Personal Information

吉見 憲二*
Kenji Yoshimi

Abstract

In this study, we analyzed articles and papers in Japan, the U.S., China, and Korea, all of which were retrieved from English databases. We employed quantitative text analysis due to examine our awareness of the expanding scope of issues within information ethics, which has led to ambiguity in the standards used.

Our findings enabled us to visualize and clarify the unique characteristics and trends in each of these countries. While it's challenging to claim deep insights due to limited data, our approach demonstrated some effectiveness in "confirming differences and commonalities," a prerequisite for considering the concept of information ethics.

This study specifically focused on one aspect of information ethics, "personal information." Therefore, in future research, we intend to validate the method's effectiveness and explore the broader concept of information ethics by applying the same approach to various analytical subjects.

I. はじめに

近年、ELSI (Ethical, Legal and Social Issues ; 倫理的・法的・社会的課題) という言葉に代表されるように先端的な科学技術の倫理面への課題に対する興味・関心が高まってきている。ELSIはもともと1990年にアメリカではじまったヒトゲノム計画を契機に広がった概念であるが、現在では生命科学分野に限らず情報技術を含めた幅広い分野に適用されている¹。例えば、大阪大学や中央大学では2020年代に入ってから相次いでELSIセンターが設立され、人工知能(AI)やIoT、ビッグデータといったテーマに関して取り組んでいる。

ELSIより前に同種の問題を扱っていた概念に「情報倫理 (information ethics)」がある。『図書館情報学用語辞典 (第5版)』における定義では、「情報にかかわって、人がいかによく生きる

* 成蹊大学経営学部 Faculty of Business Administration, Seikei University

¹ 岸本充生「ELSIとは何か」NHK解説室 <https://www.nhk.or.jp/kaisetsu-blog/400/477385.html>
2023年9月20日閲覧

かについて探求する応用倫理学の一分野」とされ、「プライバシー保護、知的財産権、情報やコンピュータネットワークの不正利用、情報セキュリティ、有害情報、デジタルデバイド等、情報にかかわる倫理的問題」が対象領域として挙げられている。他方で、村田・折戸（2021）では、「プライバシー、監視、デジタルアイデンティティ、ICTプロフェッショナルリズム、知的財産権、ジェンダーとコンピューティング、AI倫理、ロボット倫理、サイボーグ倫理」といったテーマが例として挙げられており²、その範囲が現在進行形で広がっていることが分かる。

情報倫理が包含する問題の拡大は現代の超情報化社会においては自然なことであるが、「情報倫理（information ethics）」という概念がある種の規範を扱うことを踏まえると、対象領域の無限定な拡大によってその依拠する基準が曖昧になることによる弊害が懸念される。ここでは例としてゲーム規制と研究倫理審査の2つの事例について取り上げる。

ゲームに関する規制については、日本国内では2020年に香川県が制定したネット・ゲーム依存症対策条例がその制定過程も含めて大きな議論を巻き起こし、現在進行形でその賛否が問われている³。一方で、韓国では2011年にゲーム時間の規制を行う「青少年夜間ゲームシャットダウン制」が導入されたことが話題となり、関連した調査研究も行われている⁴。しかしながら、当該制度は他の制度と統合されるかたちで2021年中に実質的に廃止された⁵。それに対して、中国では近年矢継ぎ早にゲーム規制が導入されている。2019年11月に発表された「未成年者のオンラインゲーム依存の防止に関する通知」では、未成年者のオンラインゲーム利用を「法定休日は3時間以内、それ以外の日は1時間半以内」に制限し、さらに、2021年8月に発表された「未成年者のオンラインゲーム依存を防止するための管理強化に関する通知」では、18歳未満を対象としたオンラインゲームのサービス提供が「金曜、土曜、日曜、祝日の午後8時から9時までの1時間のみ」と制限されている⁶。こうした状況下で、「未成年者のオンラインゲーム利用に対して、情報倫理を考える必要がある」といった言説が出た場合には、我々が参照すべき基準はどこになるだろうか。特定の国における特定の制度のみを基準とするようなことがあれば、現実問題の複雑さが考慮されず、「情報倫理（information ethics）」という概念が論者に都合よく利用されることになってしまう。

同じような懸念は、研究倫理審査における分野間の理解の差にも見ることができる。医学系分野から拡大してきた研究倫理審査は今では多くの研究分野において実践されているが、その質にばらつきがあることが問題として指摘されている⁷。情報倫理に関連した題材としてソーシャルメディアや監視カメラ等を研究に用いるケースが想定されるが、情報倫理に関する曖昧な理解は審査員によって何を「情報倫理上配慮が必要となる事項」とするかについて判断の不一致をもたらす可能性がある。

村田・折戸（2021）は情報倫理問題に対し、「あらかじめ決まったルールや、全面的に依拠できる道徳規範が存在するわけではなく、また必ずしも正解を探し出せるわけでもない⁸」としつ

² 村田・折戸（2021）p. ii

³ 条例の制定過程や抱える問題については山下（2023）で詳しく論じられている。

⁴ Choi et al. (2018) では、シャットダウン制が青少年のインターネット使用を削減できなかったと結論付けている。

⁵ 東亜日報「ネットゲーム強制シャットダウン制、きょうから廃止」 <https://www.donga.com/jp/article/all/20220101/3118525/1> 2023年9月20日閲覧

⁶ JETRO「未成年者のオンラインゲーム利用時間の制限を強化」 <https://www.jetro.go.jp/biznews/2021/09/a66cee87da19a2cb.html> 2023年9月20日閲覧

⁷ 吉見（2021）では、研究倫理審査委員会のばらつきを「倫理審査委員会3000個問題」として提起している。

⁸ 村田・折戸（2021）p. ii

つ、先行研究を踏まえて、「絶え間なく進行するICTの開発と利用とともに立ち現れてくるポリシーの空白を埋め続ける実践⁹⁾」として情報倫理を位置づけている。こうした観点を踏まえると、「情報倫理 (information ethics)」という概念をより蓋然性高く把握するためには、その対象範囲の明確化、他国の事例や法規制、研究分野における差異などを考慮する必要がある。

II. 先行研究

1. 国内における情報倫理の研究

茶園 (2002) は、情報倫理に関する国内の研究グループとして、電子情報通信学会の「情報文化と倫理研究会 (FACE : Forum for Advanced Communications Ethics)¹⁰⁾」や、「情報処理学会 (IPSJ) 電子化知的財産・社会基盤研究グループ (SIG-EIP)」、「情報倫理の構築プロジェクト (FINE : Foundations of Information Ethics)」を挙げている。特に、1998年から2003年にわたって展開された「情報倫理の構築プロジェクト (以下、FINEプロジェクト)」は、「倫理問題の実例や、各国の法規、各機関における倫理綱領などの一次資料に加えて、情報倫理に関するあらゆる研究文献を収集し、国内初の情報倫理資料センターとしての役割を果たすことを目的とし、多くの研究成果を生み出した¹¹⁾。FINEプロジェクトのような取り組みは、本研究の問題意識である「情報倫理 (information ethics)」が依拠する基準を考えるうえで、極めて有用である。しかしながら、こうした大型の研究プロジェクトを維持するためのコストは大きく、長期に亘って継続していくことは困難を伴う。

情報倫理という概念そのものを扱った研究はいくつか見られる¹²⁾ものの、特定のテーマを深く検討したものや広義の情報倫理教育を扱ったものが多数を占めている。萩原 (2023) は、情報倫理学が機能するために社会的文脈理解の必要性を示しつつ、倫理の普遍性をめぐる問題に対して、「(認識や規範を共有する) 可能性を探るに先立って、まずは互いの差異や共通性を確認する作業が不可欠である」と指摘している。村田・折戸 (2021) でも、グローバル化する現代情報社会における情報倫理の取り組みに「グローバルに受容可能で、かつローカルに有効性を持つもの (Globally Acceptable, and Locally Effective)¹³⁾」が必要であることに言及している。

2. 各国における情報倫理の研究

上述したような、情報倫理概念の相対的な差異を捉えようとする研究については、日本以外でも行われている。例えば、Myung and Lee (2013) は、サイバー倫理 (情報倫理) に関する韓国国内の研究をメタ分析し、研究分野における差異やそれぞれの重点的な研究領域について明らかにしている。Chang (2012) と Chang (2017) では、情報倫理に関する文化的な影響が検討されており、前者では中国の文化を踏まえた適切な情報倫理規定の在り方の提案が、後者では中国人学生とアメリカ人学生の情報倫理認識の比較が行われている。

こうした研究はあくまで一例ではあるものの、萩原 (2023) が指摘する「互いの差異や共通性を確認する作業」や村田・折戸 (2021) の「グローバルに受容可能で、かつローカルに有効性を

⁹⁾ 村田・折戸 (2021) p.58

¹⁰⁾ 2002年から「技術と社会・倫理研究会 (SITE)」に改称している。

¹¹⁾ オリジナルの研究報告書のリンクが消えていたため、二次資料として茶園 (2002) を参照した。

¹²⁾ 例えば、辰巳・原田 (1999)、水谷 (2000)、大谷 (2015) などが挙げられる。

¹³⁾ 村田・折戸 (2021) p.76

持つもの (Globally Acceptable, and Locally Effective)」の一環として理解することができる。

3. データベースを用いた情報倫理概念の比較に関する研究

吉見 (2022)、吉見 (2023) では、同様の問題意識から、データベースから取得できる情報を元に情報倫理概念の比較を試みている。吉見 (2022) では、世界最大級のオンライン学術データベースである Web of Science に収録されている論文を対象に、①トピック (タイトル, 抄録, 著者キーワード, Keywords Plus) に “information ethics” とそれぞれの国名 (Japan, China, Korea) が含まれているもの、②誌名が明示的に “information ethics” と関連している論文誌でトピックに国名 (Japan, China, Korea) が含まれているものを抽出したところ、最終的に得られた論文は、日本を対象とした論文12本、中国を対象とした論文12本、韓国を対象とした論文4本のみであった。加えて、吉見 (2023) では、世界中の大手出版社からの新聞、ニュース Web サイト、ブログの全文を収録したデータベースである ProQuest Newsstand に収録されている全文が取得できる新聞記事を対象に “information ethics” をキーワードとして抽出したところ、該当したのは U.S. Newsstream から25本、Asian Newsstream から24本のみであった。

いずれの研究も英語で取得できる世界最大規模のDBを利用しているが、明示的に国名を含んで “information ethics” について論じている記事や論文はごく少数であった。

4. 先行研究の課題と本研究の目的

先行研究からは、蓋然性の高い情報倫理概念の共有のために、まずは国や研究分野を踏まえて差異や共通性を相対的に捉える必要性が提起されている。しかしながら、直接的に “information ethics” という単語から情報倫理概念の実態について把握することは、論文DB上からも新聞DB上からも困難であった。

そこで、本研究では手法の有効性を示すことを主な目的とし、「情報倫理 (information ethics)」という概念そのものではなく、情報倫理が包含する具体的な問題に対して、計量テキスト分析の手法を用いて「差異や共通性を確認する」ことを試みる。

III. 分析手法

1. 分析対象

本研究では、特に日米中韓の国際比較の観点から英語で発信されている情報に着目する。そのため、新聞DBと論文DBから取得できる英文情報を分析対象とした。具体的には、新聞DBとして前述の ProQuest Newsstand を利用し、各国の収録されている英文記事を収集した。ただし、データベースへの収録期間が一致していないため、2013年9月1日から2023年8月31日までの10年間の記事を対象とした。論文DBに関しては、こちらも前述の Web of Science Core Collection を利用し、トピック検索の結果から該当した論文の abstract を収集した。データ収集に使用したデータベースは表1の通りである。

表1 データ収集に使用したデータベース

	日本	米国	中国	韓国
新聞 DB	The Japan News	New York Times	China Daily	The Korea Times
論文 DB	Web of Science	Web of Science	Web of Science	Web of Science

先行研究で取り上げたように、「情報倫理 (information ethics)」という概念については、直接的に言及されていないことが多い。そのため、本稿では各国比較を行うことを優先し、情報倫理に関する課題の中でも対応する英語訳が明確であり、分析に十分なデータを有するトピックに限定し、調査を行うこととした。候補となるトピックは「デジタルデバインド (digital divide)、ネットいじめ (cyber-bullying)、インターネット依存 (internet addiction)、ゲーム障害 (gaming disorder)¹⁴、プライバシー (privacy)、個人情報 (personal information)」の6つとし、それぞれの該当件数について表2にまとめている。

なお、ProQuest Newsstandでの検索にあたっては、記事全文が取得できるものとし、ドキュメントのタイプはニュースに限定している。加えて、関係のない記事を避けるために、「主題」を当該の項目と一致するものとした。そのため、キーワードだけが含まれた関連性の少ない記事は排除されている。Web of Scienceの検索機能では、先行研究と同様にトピック (タイトル、抄録、著者キーワード、Keywords Plus) に各キーワードとそれぞれの国名 (Japan, United States, China, Korea) を含むものとした。それぞれの該当記事/論文の件数は表2の通りとなった¹⁵。

最終的に、新聞DBと論文DBの双方で50件以上を示した“personal information”を対象以降の分析を実施する。

表2 キーワードと該当の記事/論文数

キーワード	新聞 DB				論文 DB			
	US	CH	KR	JP	US	CH	KR	JP
digital divide	43	46	16	0	321	301	76	31
cyber-bullying	19	22	17	4	169	112	49	15
internet addiction	5	9	6	5	46	350	124	71
gaming disorder	5	12	25	5	28	136	75	35
privacy	1490	185	120	26	1953	959	265	241
personal information	344	269	182	90	186	221	95	82

2. 分析手順

本研究では、計量テキスト分析の手法を用いて各国間の比較を行うことを試みた。まず、国ごとに頻出上位30語の抽出と共起ネットワークの可視化を行った。ただし、情報倫理に関する話題に注目するため、品詞を名詞、固有名詞、形容詞に限定した。加えて、より重要度が高い単語を抽出するために、共起ネットワークの描写対象を全体の2割以上の記事/論文に登場した単語とし、出現数ではなく、記事/論文ごとの登場の有無をカウントした。こうした条件によって、特定の記事/論文に大量に登場する単語があっても、その影響を排除することができる。共起の基準としてJaccard係数を採用し、原則として新聞記事では0.4以上の共起関係、論文abstractでは0.3以上の共起関係に着目した¹⁶。

¹⁴ 表記ゆれを考慮し、game disorder, gaming disease, game disease, gaming addiction, game addiction を検索対象語句とした。

¹⁵ それぞれの検索条件に該当したものであるため、“personal information”が直接言及されていない記事や論文も含まれている。

¹⁶ Jaccard係数は類似性の指標であり、単語Aと単語Bにおいて、 $|A \cap B| / |A \cup B|$ の式によって求められる。単語間の共起関係を描写する基準として、Jaccard係数以外にもDice係数やSimpson係数が存在するが、より保守的な基準を採用した。

次に、各国間の比較を行うために、各国の記事と論文をそれぞれ統合したデータについて分析を行った。こちらは単語の共起関係ではなく、各国に特徴的な単語の把握を意図したものであるため、頻出上位100語についてコサイン類似度¹⁷を用いて類似性の高い単語をそれぞれ可視化した。

最後に、これまでの分析で特徴的な単語に対して、国ごとの登場傾向についてカイ二乗分析を行った。一連の分析を通じて、各国間の比較を視覚的に示すこと、特徴のある単語に関して統計的な差異があるか検証することを試みている。

分析ソフトについては、フリーのテキストマイニングソフトウェアであるKH Coder¹⁸を使用し、形態素解析についても付属のStanford POS taggerを使用した¹⁹。

本研究では、研究者が利用できる商用データベースからデータを取得しており、検索条件を明示している。さらに、分析ソフトウェアもフリーのソフトウェアが利用できるため、再現性が高い手法となっている。

IV. 分析結果

1. 各国の分析結果

(1) 米国

米国の新聞記事における頻出上位語と共起ネットワーク分析の結果はそれぞれ表3と図1に示している。なお、視認性に問題があったため、共起の基準を0.4から0.5に引き上げている。頻出上位語では、新聞記事特有の単語に加えて、“company” “privacy” “security” “law”といった単語が見られた。さらに、共起ネットワークでは、“FACEBOOK” “GOOGLE”といった具体的な企業名や“social” “media”への言及があることが特徴的であった。

米国の論文abstractにおける頻出上位語と共起ネットワーク分析の結果はそれぞれ表4と図2に示している。新聞記事の結果と同様に、“privacy” “law”はこちらでも上位に登場していたが、“company” “security”は見られなかった。一方で、“consumer”や“policy”が上位に登場している。共起ネットワークでは、“privacy”が頻出語の中でも共起関係が多い単語となっていた。

(2) 中国

中国の新聞記事における頻出上位語と共起ネットワーク分析の結果はそれぞれ表5と図3に示している。“company” “privacy” “security” “law”は米国と同様に頻出上位語となっていた。加えて、“regulation”は米国には見られなかった単語である。共起ネットワークでは、米国のように具体的な企業名は見られなかったものの、“user”と“app”の共起が見られた。

中国の論文abstractにおける頻出上位語と共起ネットワーク分析の結果はそれぞれ表6と図4に示している。頻出上位語、共起ネットワークともに論文に用いられる単語が中心になっていた。それ以外では、“health”が目立つ程度であった。

¹⁷ コサイン類似度はベクトル空間モデルによって共起関係を計算する手法で、ベクトルの内積から類似度を0から1までの連続した数値で求めることができる。コサイン類似度の説明として新原（2022）などがある。

¹⁸ KH Coder <https://kncoder.net/> 2023年9月20日閲覧
なお、KH Coderの使用手法や活用事例は樋口（2020）にまとめられている。

¹⁹ 強制抽出の対象語句として、“personal information” “United States” “New York Times” “New York”の4語を設定した。

表3 米国記事の頻出上位語

#	抽出語	出現数	比率	#	抽出語	出現数	比率
1	information	299	86.9%	16	time	200	58.1%
2	personal_information	299	86.9%	17	month	195	56.7%
3	datum	287	83.4%	18	user	192	55.8%
4	other	287	83.4%	19	service	191	55.5%
5	people	284	82.6%	20	law	190	55.2%
6	company	280	81.4%	21	access	186	54.1%
7	year	266	77.3%	22	number	186	54.1%
8	last	247	71.8%	23	way	184	53.5%
9	more	247	71.8%	24	security	179	52.0%
10	Mr	227	66.0%	25	government	178	51.7%
11	personal	225	65.4%	26	photograph	178	51.7%
12	many	218	63.4%	27	week	171	49.7%
13	new	215	62.5%	28	such	166	48.3%
14	privacy	211	61.3%	29	United_States	165	48.0%
15	New_York_Times	210	61.0%	30	public	164	47.7%

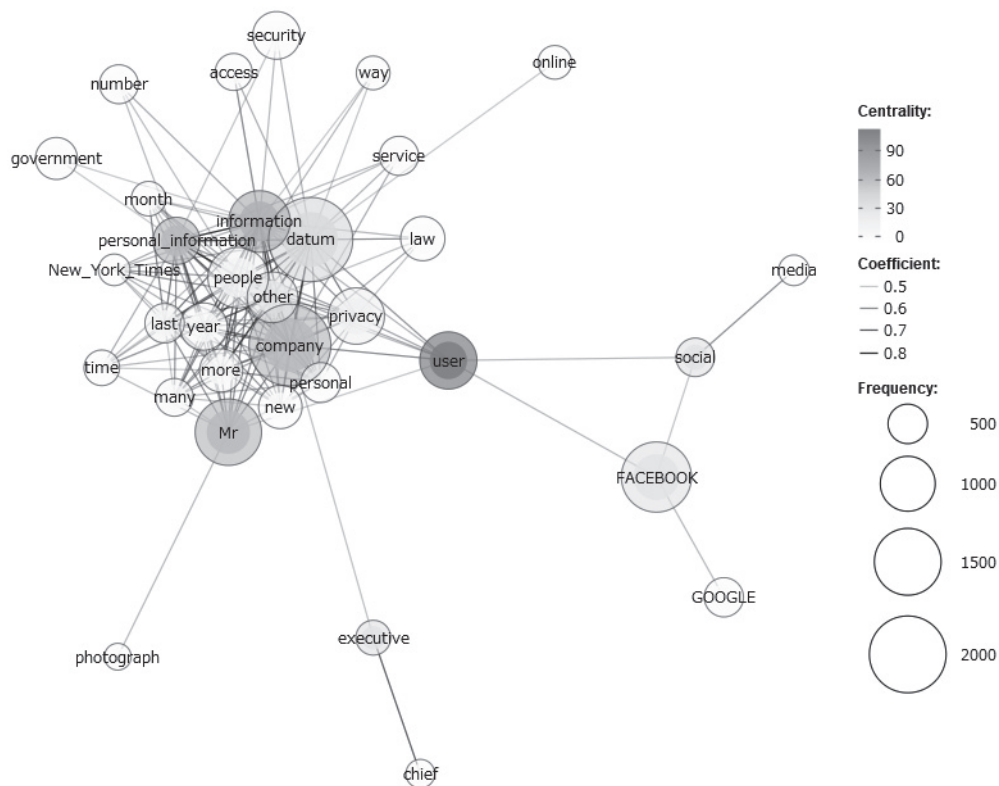


図1 米国記事の共起ネットワーク

表4 米国論文abstractの頻出上位語

#	抽出語	出現数	比率	#	抽出語	出現数	比率
1	personal_information	174	94.6%	16	consumer	43	23.4%
2	United_States	146	79.3%	16	protection	43	23.4%
3	datum	99	53.8%	18	concern	42	22.8%
4	information	94	51.1%	18	personal	42	22.8%
5	privacy	88	47.8%	20	article	41	22.3%
5	study	88	47.8%	20	individual	41	22.3%
7	result	69	37.5%	20	policy	41	22.3%
8	analysis	59	32.1%	23	public	39	21.2%
9	other	57	31.0%	24	conclusion	37	20.1%
10	such	54	29.3%	24	technology	37	20.1%
11	use	51	27.7%	26	risk	36	19.6%
12	finding	47	25.5%	27	implication	35	19.0%
13	law	45	24.5%	27	practice	35	19.0%
13	method	45	24.5%	27	social	35	19.0%
13	research	45	24.5%	30	online	34	18.5%
				30	significant	34	18.5%

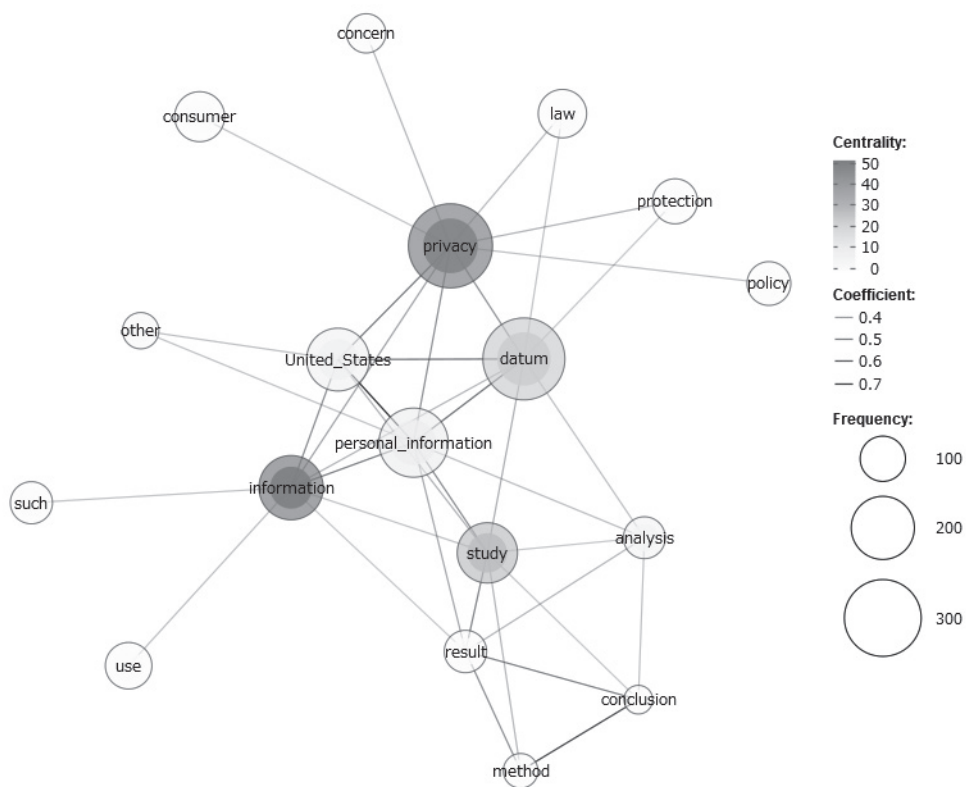


図2 米国論文abstractの共起ネットワーク

表5 中国記事の頻出上位語

#	抽出語	出現数	比率	#	抽出語	出現数	比率
1	personal_information	263	97.8%	16	user	127	47.2%
2	CHINA	212	78.8%	17	internet	126	46.8%
3	information	206	76.6%	18	privacy	125	46.5%
4	such	203	75.5%	19	service	124	46.1%
5	people	198	73.6%	20	new	123	45.7%
6	law	173	64.3%	21	number	123	45.7%
7	datum	165	61.3%	22	many	113	42.0%
8	other	157	58.4%	23	online	113	42.0%
9	protection	156	58.0%	24	regulation	112	41.6%
10	year	154	57.2%	25	case	110	40.9%
11	more	146	54.3%	26	country	110	40.9%
12	security	144	53.5%	27	Law	109	40.5%
13	public	143	53.2%	28	technology	107	39.8%
14	company	131	48.7%	29	BEIJING	102	37.9%
15	personal	127	47.2%	30	use	99	36.8%

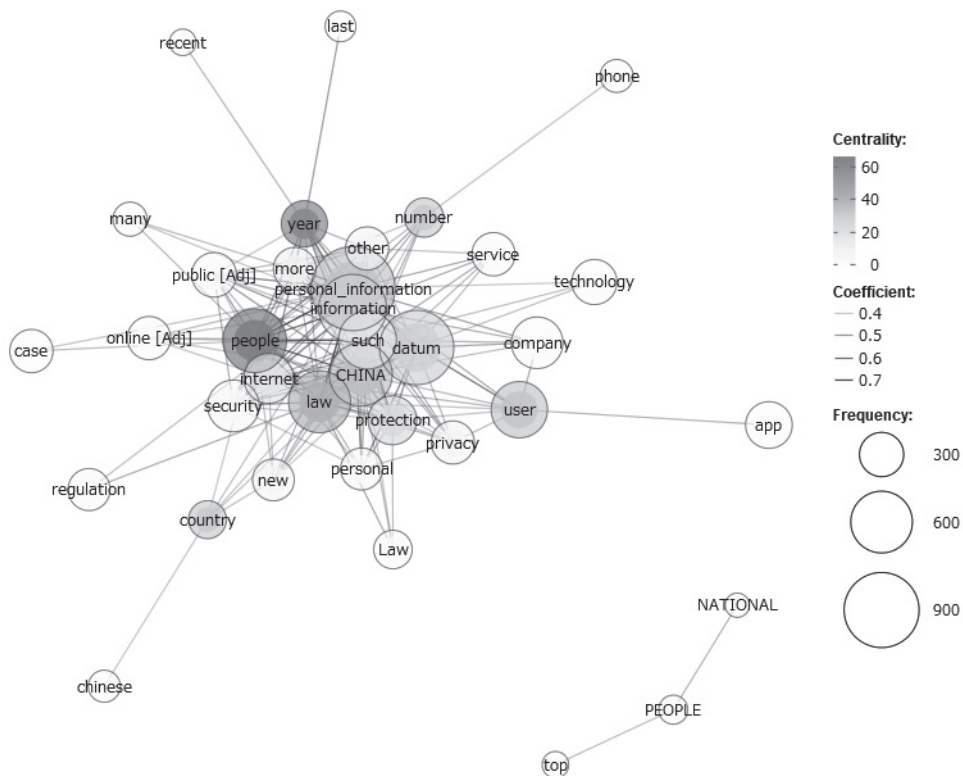


図3 中国記事の共起ネットワーク

表6 中国論文abstractの頻出上位語

#	抽出語	出現数	比率	#	抽出語	出現数	比率
1	personal_information	213	97.7%	16	other	56	25.7%
2	China	172	78.9%	17	significant	56	25.7%
3	study	144	66.1%	18	effect	53	24.3%
4	datum	121	55.5%	19	level	53	24.3%
5	result	120	55.0%	20	model	53	24.3%
6	method	93	42.7%	21	different	52	23.9%
7	information	91	41.7%	22	survey	52	23.9%
8	analysis	87	39.9%	23	protection	51	23.4%
9	conclusion	70	32.1%	24	such	51	23.4%
10	factor	70	32.1%	25	people	50	22.9%
11	chinese	67	30.7%	26	age	47	21.6%
12	health	59	27.1%	27	questionnaire	47	21.6%
13	privacy	58	26.6%	28	development	46	21.1%
14	risk	57	26.1%	29	research	46	21.1%
15	background	56	25.7%	30	higher	45	20.6%

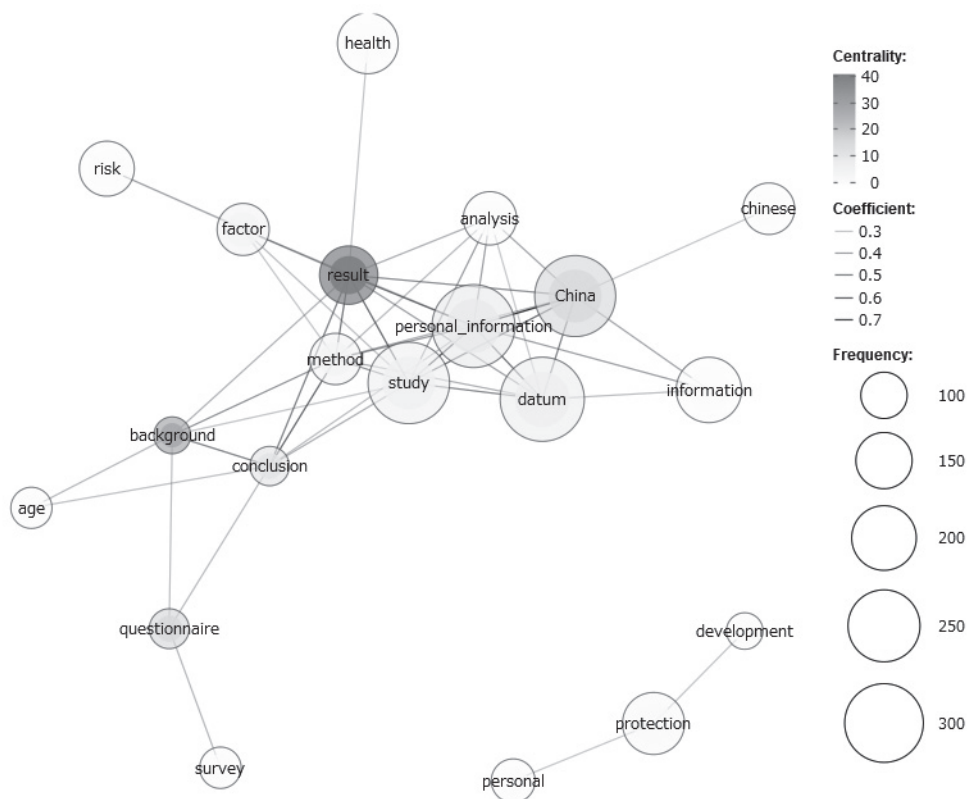


図4 中国論文abstractの共起ネットワーク

(3) 韓国

韓国の新聞記事における頻出上位語と共起ネットワーク分析の結果はそれぞれ表7と図5に示している。“company” “law” は米中と同様に頻出上位語となっていたが、“privacy” “security” は頻出上位に登場しなかった。さらに、韓国の記事における特徴的な単語として、“credit” “card” “financial” が挙げられる。これは韓国で2014年に発覚したクレジットカードの顧客情報流出事件²⁰が背景にあると考えられる。共起ネットワークにおいても、クレジットカード流出事件に関連したサブグループの存在が見られた。

韓国の論文abstractにおける頻出上位語と共起ネットワーク分析の結果はそれぞれ表8と図6に示している。中国の論文と同様に、“health” “medical” といった健康・医療に関する単語が頻出上位に登場していた。

(4) 日本

日本の新聞記事における頻出上位語と共起ネットワーク分析の結果はそれぞれ表9と図7に示している。米国と同様に、視認性に問題があったため、共起の基準を0.4から0.5に引き上げている。“company” “security” “law” が頻出上位に登場した一方で、他国には登場していた“privacy” はなかった。さらに、“government” の出現割合が他国と比べても高くなっている。共起ネットワークにおける“name” “address” の共起関係も他国には見られないものであった。

日本の論文abstractにおける頻出上位語と共起ネットワーク分析の結果はそれぞれ表10と図8に示している。論文に関する単語が中心となっており、特徴的な単語があまり見られなかった。記事には登場しなかった“privacy” が頻出語であったものの、他国と比べても共起ネットワークにおける中心性は低かった。

2. コサイン類似度による各国比較

ここまで概観してきた通り、各国の頻出上位語と共起ネットワークについてそれぞれ類似点と相違点が見られた。加えて、記事と論文abstractの比較では、あまり多くの一致点は見られなかった。しかしながら、単純な頻出上位語の比較では各国の比較は容易ではない。そこで、コサイン類似度を基準に各国に特徴的な単語を可視化したものが図9と図10である。

大まかな傾向としては、図1～8で示した結果と大きな差異はなく、各国の頻出上位語が概ね特徴的な単語として登場していた。米国では記事でも論文abstractでも“privacy” が特徴的な単語となっており、それに対して、日本では記事でも論文abstractでも“government” が特徴的な単語となっていた。逆に、中国と韓国では記事と論文abstractであまり一貫した傾向が見られなかった。

²⁰ REUTERS 「韓国で1億人超分の個人情報流出、カード会社への怒り渦巻く」 <https://jp.reuters.com/article/13n0kv271-south-korea-card-info-idJPTYEA0K04Q20140121> 2023年9月20日閲覧

表7 韓国記事の頻出上位語

#	抽出語	出現数	比率	#	抽出語	出現数	比率
1	personal_information	182	100%	16	personal	72	39.6%
2	information	128	70.3%	17	case	71	39.0%
3	Korea	127	69.8%	17	law	71	39.0%
4	datum	110	60.4%	19	user	69	37.9%
5	such	107	58.8%	20	last	64	35.2%
6	company	104	57.1%	20	time	64	35.2%
7	other	103	56.6%	22	country	63	34.6%
8	official	92	50.5%	23	financial	62	34.1%
9	people	90	49.5%	24	name	61	33.5%
10	year	89	48.9%	25	credit	60	33.0%
11	government	85	46.7%	26	National	58	31.9%
11	number	85	46.7%	27	card	57	31.3%
13	service	83	45.6%	28	measure	56	30.8%
14	customer	74	40.7%	28	new	56	30.8%
15	firm	73	40.1%	30	phone	55	30.2%
				30	system	55	30.2%

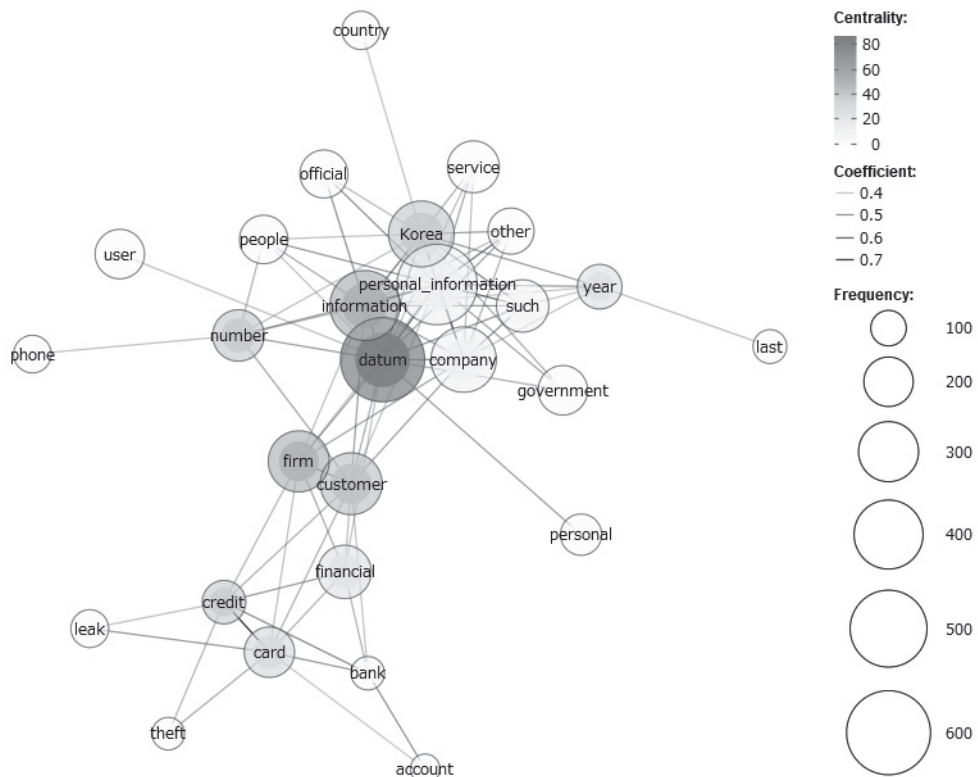


図5 韓国記事の共起ネットワーク

表8 韓国論文abstractの頻出上位語

#	抽出語	出現数	比率	#	抽出語	出現数	比率
1	personal_information	86	91.5%	15	survey	25	26.6%
2	Korea	85	90.4%	17	analysis	24	25.5%
3	study	71	75.5%	17	factor	24	25.5%
4	datum	57	60.6%	19	public	22	23.4%
5	result	47	50.0%	19	purpose	22	23.4%
6	information	38	40.4%	21	conclusion	21	22.3%
7	South	35	37.2%	21	research	21	22.3%
7	such	35	37.2%	23	finding	20	21.3%
9	system	31	33.0%	23	policy	20	21.3%
9	use	31	33.0%	23	service	20	21.3%
11	method	30	31.9%	26	level	19	20.2%
11	privacy	30	31.9%	26	protection	19	20.2%
13	health	27	28.7%	26	technology	19	20.2%
13	korean	27	28.7%	29	age	18	19.1%
15	medical	25	26.6%	29	country	18	19.1%
				29	risk	18	19.1%

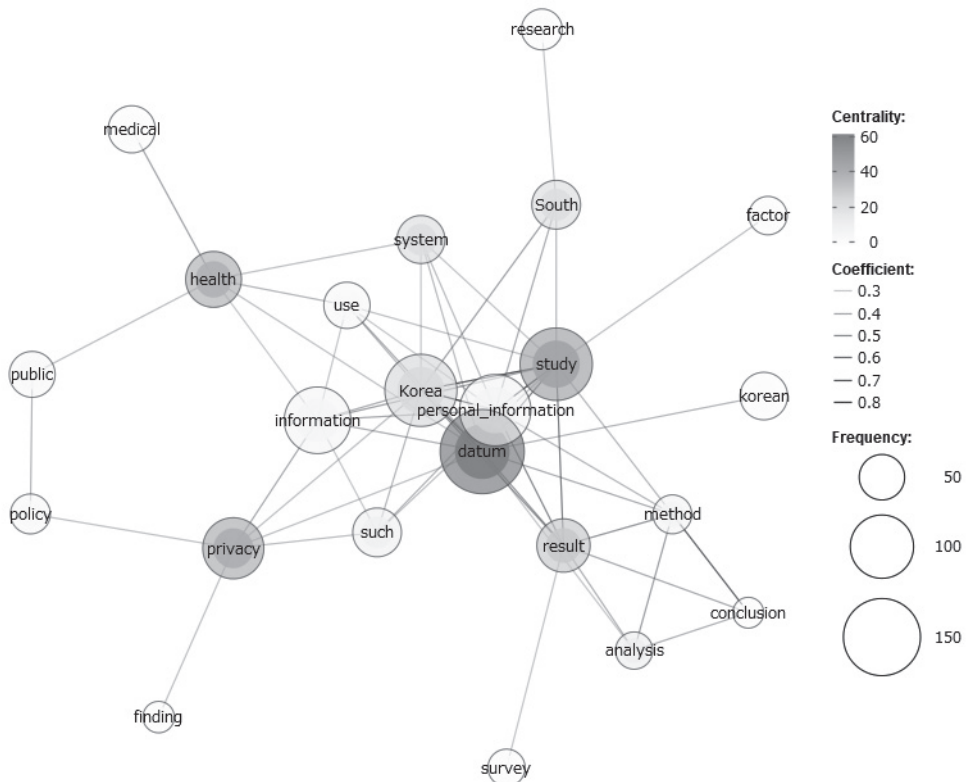


図6 韓国論文abstractの共起ネットワーク

表9 日本記事の頻出上位語

#	抽出語	出現数	比率	#	抽出語	出現数	比率
1	personal_information	90	100.0%	16	use	42	46.7%
2	information	81	90.0%	17	many	41	45.6%
3	such	75	83.3%	17	measure	41	45.6%
4	other	72	80.0%	19	new	41	45.6%
5	government	69	76.7%	20	Tokyo	41	45.6%
6	datum	63	70.0%	20	service	40	44.4%
7	people	62	68.9%	22	protection	39	43.3%
8	system	58	64.4%	23	individual	38	42.2%
9	Japan	56	62.2%	24	official	38	42.2%
10	company	54	60.0%	25	possible	38	42.2%
11	name	48	53.3%	26	security	38	42.2%
11	number	48	53.3%	27	law	37	41.1%
13	year	47	52.2%	28	organization	37	41.1%
14	case	46	51.1%	28	time	37	41.1%
15	personal	45	50.0%	30	public	36	40.0%

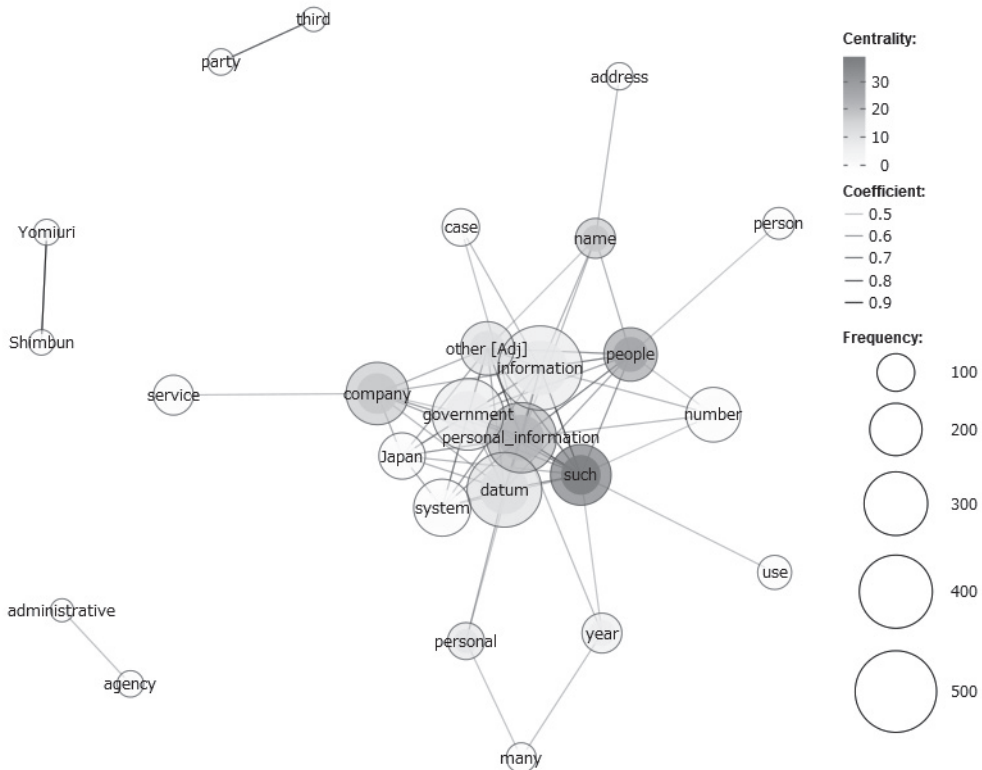


図7 日本記事の共起ネットワーク

表10 日本論文abstractの頻出上位語

#	抽出語	出現数	比率	#	抽出語	出現数	比率
1	personal_information	74	93.7%	15	year	20	25.3%
2	Japan	72	91.1%	18	country	19	24.1%
3	result	38	48.1%	18	development	19	24.1%
4	datum	37	46.8%	20	public	18	22.8%
4	study	37	46.8%	20	response	18	22.8%
6	such	32	40.5%	20	survey	18	22.8%
7	information	28	35.4%	23	government	17	21.5%
7	method	28	35.4%	23	personal	17	21.5%
7	protection	28	35.4%	23	questionnaire	17	21.5%
10	japanese	25	31.6%	26	analysis	16	20.3%
11	other	24	30.4%	27	conclusion	15	19.0%
11	use	24	30.4%	27	guideline	15	19.0%
13	privacy	22	27.8%	27	health	15	19.0%
14	system	21	26.6%	27	issue	15	19.0%
15	paper	20	25.3%	27	policy	15	19.0%
15	research	20	25.3%	27	service	15	19.0%

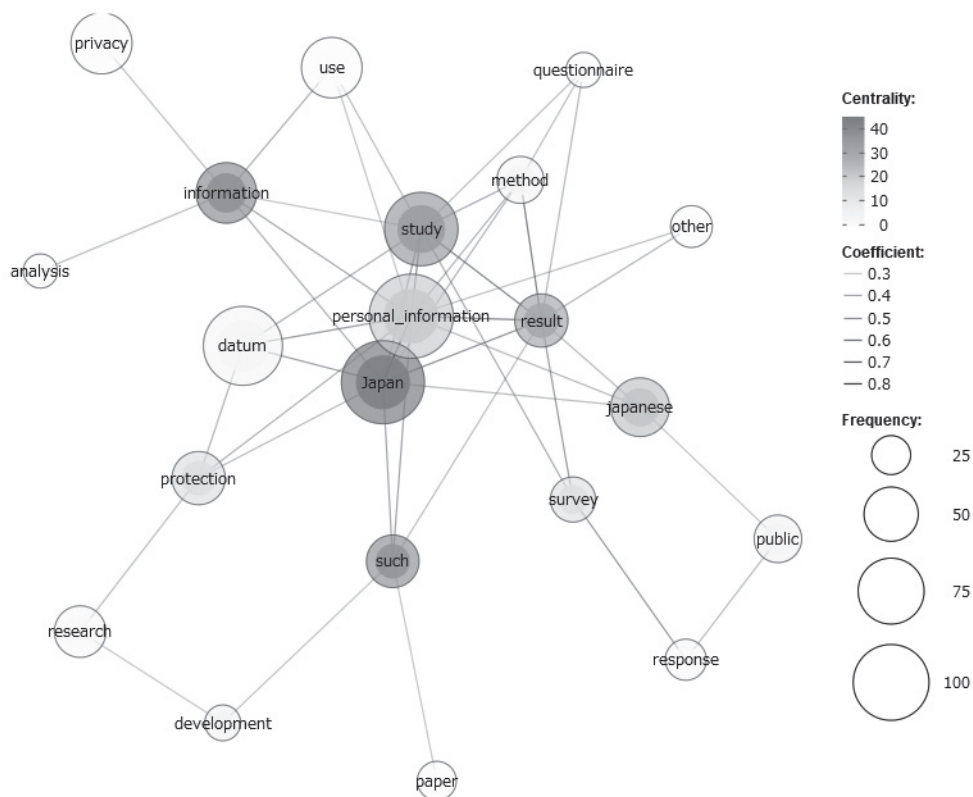


図8 日本論文abstractの共起ネットワーク

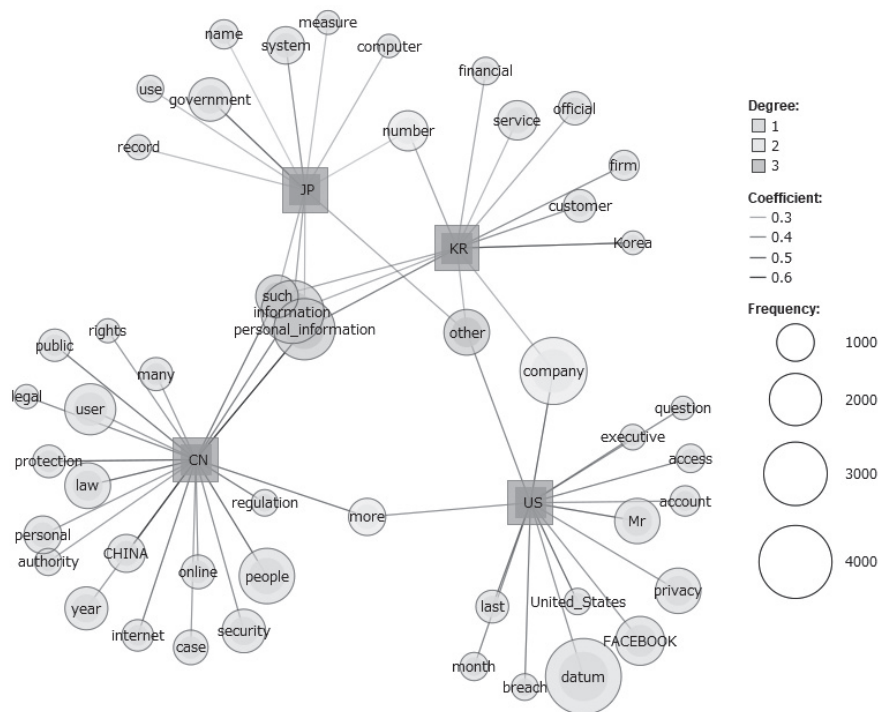


図9 コサイン類似度（記事）による各国比較

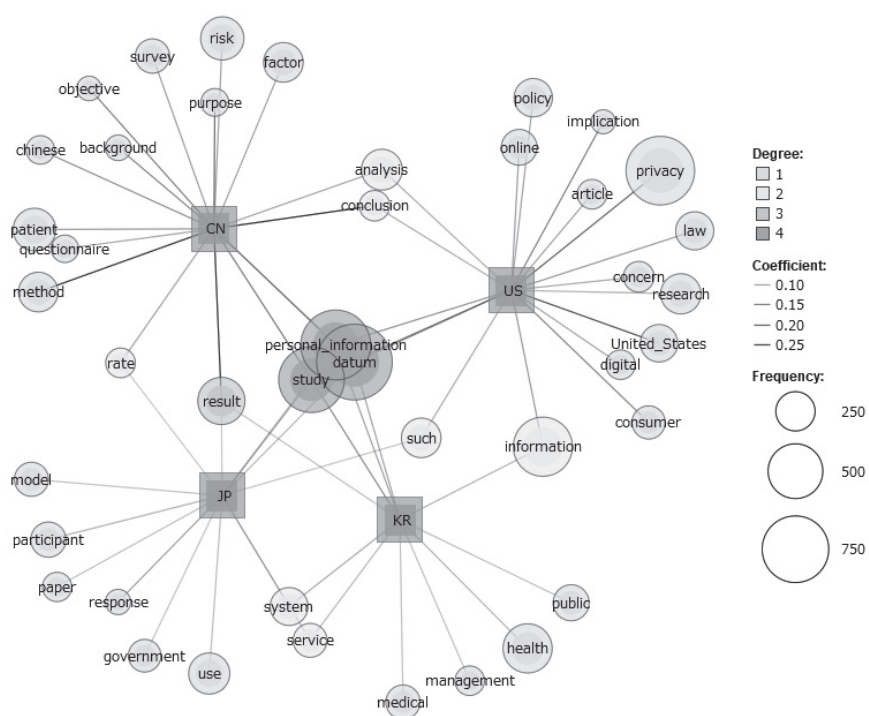


図10 コサイン類似度（論文abstract）による各国比較

3. 特徴的な単語に関する各国比較

これまでの結果で取り上げた特徴的な単語のうち、多くに登場してきた“government”“company”“privacy”“security”“law”“policy”“regulation”の7単語に着目し、各国の登場傾向についてクロス集計表を踏まえて比較したものが表11と表12である。なお、登場傾向の差異が明確になるように、記事単位ではなく、文単位での登場数をここではカウントしている。

図9、10と同様に、表11、12でも日本では“government”への、米国では“privacy”への言及がそれぞれ有意に多く、“personal information”と共起する象徴的な単語となっていた。また、米国では“company”との共起も多く、より企業との関係において「個人情報（personal information）」を意識していることが示唆されている。それに対し、日本では記事でも論文でも“government”への共起が多く、“privacy”への共起が少ないという傾向が一貫して示されており、国としてのスタンスの違いが顕在化していた。

中国では、記事において“security”“law”“regulation”の登場比率が他国に比べて有意に高かったものの、論文では同様の傾向は見られなかった。このことから、記事と論文において捉え方に差異があるかもしれない。韓国は特有の傾向があまり見られず、日米中の中間に位置づけられるものが多かった。

表11 記事のクロス集計表

	government	company	privacy	security	law	policy	regulation	ケース数
US	496 (5.65%)	1793 (20.44%)	934 (10.65%)	742 (8.46%)	565 (6.44%)	231 (2.63%)	104 (1.19%)	8772
CN	175 (8.18%)	308 (14.39%)	218 (10.19%)	329 (15.37%)	450 (21.03%)	58 (2.71%)	221 (10.33%)	2140
KR	169 (12.05%)	244 (17.39%)	86 (6.13%)	120 (8.55%)	136 (9.69%)	44 (3.14%)	62 (4.42%)	1403
JP	262 (19.91%)	208 (15.81%)	51 (3.88%)	86 (6.53%)	114 (8.66%)	34 (2.58%)	45 (3.42%)	1316
合計	1102 (8.08%)	2553 (18.73%)	1289 (9.46%)	1277 (9.37%)	1265 (9.28%)	367 (2.69%)	432 (3.17%)	13631
カイ2乗値	346.972**	52.349**	81.877**	112.990**	435.678**	1.233	477.168**	

表12 論文のクロス集計表

	government	company	privacy	security	law	policy	regulation	ケース数
US	50 (3.02%)	44 (2.66%)	299 (18.09%)	28 (1.69%)	103 (6.23%)	83 (5.02%)	39 (2.36%)	1653
CN	46 (2.04%)	14 (0.62%)	202 (8.97%)	40 (1.78%)	64 (2.84%)	62 (2.75%)	51 (2.27%)	2251
KR	24 (2.92%)	9 (1.09%)	84 (10.21%)	21 (2.55%)	25 (3.04%)	34 (4.13%)	15 (1.82%)	823
JP	36 (4.97%)	11 (1.52%)	58 (8.01%)	9 (1.24%)	14 (1.93%)	25 (3.45%)	14 (1.93%)	724
合計	156 (2.86%)	78 (1.43%)	643 (11.80%)	98 (1.80%)	206 (3.78%)	204 (3.74%)	119 (2.18%)	5451
カイ2乗値	17.189**	28.905**	92.102**	4.017	40.778**	14.118**	1.024	

IV. 考察

本研究では、世界最大規模の新聞DBと論文DBを利用し、情報倫理に関する問題の中でも「個人情報（personal information）」に関連した内容について各国比較を試みた。結果より、日本における“government”への言及の多さと、米国における“privacy”への言及の多さ、韓国におけるクレジットカードの顧客情報流出事件の影響の大きさ、中国における記事と論文の傾向の違い、といった観点を得ることができた。これらは必ずしも深いレベルでの文化的差異を明らかにしたものではないが、議論のための相対的な立場を認識する上で有用な情報となっている。

一方で、本研究の課題として、分析対象とする情報の不足が挙げられる。今回は50件以上の記事/論文を対象としたが、頑健な結果を導くためにはサンプルサイズが十分ではないだろう。先行研究で“information ethics”に直接言及した記事や論文が少なかったことや、情報倫理に関連した問題の多くが十分な分析対象となる件数を保持していなかったことも大きな課題である。複雑な国内問題に関しては、英語での発信が限定的になっている可能性があるため、より精緻な結果を得るためには翻訳によるトランスレーションバイアスを考慮したうえで、各国のローカル言語での情報発信についても対象に含めることが求められる。

加えて、今回用いた分析手法は情報縮約の手法であり、因果関係を示すものではないため、その結果の扱いについては慎重にならなければいけない。採用する条件や基準によって結果が大きく変わらないかについても検討の余地があるだろう。

しかしながら、本研究では各国の特徴的な単語や傾向についてはある程度一貫した結果が得られている。さらに、同一基準を適用した場合の各国の差異について視覚的に提示できたことは、よりリーズナブルなアプローチとして広い範囲に適用が可能である。分析対象や分析手法を今後洗練させていく必要はあるものの、当初の目的である「差異や共通性を確認する」ことについては一定の成果を上げることができたと言える。

V. まとめ

本研究では、情報倫理が包含する問題の拡大とそれによる依拠する基準の曖昧さを問題意識とし、日米中韓の各国から英語で発信された情報を対象に、DBから取得された記事と論文に対して、計量テキスト分析の手法を用いて分析した。結果より、各国の特徴や傾向を可視化し、明らかにすることができた。情報量の不足から深い洞察が得られたとは言い難いものの、情報倫理概念を共有する前提となる「差異や共通性を確認する」ための手法として、ある程度の有効性を示すことができた。

本研究ではあくまで「個人情報 (personal information)」という単一の情報倫理に関する問題しか扱うことができていない。そのため、今後の研究では、さまざまな分析対象に対して同様のアプローチを試みることで、手法の有効性を示すとともに、情報倫理概念の全体像に対して検討していきたい。

利益相反について

本論文に関して、開示すべき利益相反事項はない。

参考文献

- 大谷卓史 (2015) 「情報倫理学の3つの起源 サイバネティックス・倫理綱領・応用倫理学」『情報管理』58巻2号, pp.139-144
- 辰己丈夫・原田康也 (1999) 「新しい「情報倫理」の目指すもの」『情報処理学会論文誌』Vol.40, No.3, pp.990-997

- 茶園利昭 (2002) 「情報倫理教育に関する一考察 (IV)」『日本橋学館大学紀要』1巻, pp.79-92
- 日本図書館情報学会用語辞典編集委員会 (2020) 『図書館情報学用語辞典 (第5版)』丸善出版
- 萩原優騎 (2023) 「社会リテラシーとしての再帰的近代化論 —情報倫理学の社会的文脈への理解を深めるために—」『東京海洋大学研究報告』第19号, pp.15-31
- 樋口耕一 (2020) 『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【第2版】』ナカニシヤ出版
- 水谷雅彦 (2000) 「インターネット時代の情報倫理学」越智貢・土屋俊・水谷雅彦編『情報倫理学 電子ネットワーク社会のエチカ』ナカニシヤ出版, pp.3-48
- 村田潔・折戸洋子 (2021) 『情報倫理入門: ICT社会におけるウェルビーイングの探求』ミネルヴァ書房
- 山下洋平 (2023) 『ルポゲーム条例 なぜゲームが狙われるのか』河出書房新社
- 吉見憲二 (2021) 「情報社会における倫理審査と倫理審査委員会3000個問題」『情報処理学会論文誌』Vol.62, No.12, pp.2119-2126
- _____ (2022) 「情報倫理に関する研究の日中韓比較」『研究報告電子化知的財産・社会基盤 (EIP)』2022-EIP-95(14), pp.1-6
- _____ (2023) 「新聞DBを用いた情報倫理に関連した記事の比較」『研究報告電子化知的財産・社会基盤 (EIP)』2023-EIP-101(8), pp.1-4
- 新原俊樹 (2022) 「コサイン類似度を用いた行政文書タイトルの同一判定」『レコード・マネジメント』82巻, pp.61-63
- Chang, C. L. (2012) "How to build an appropriate information ethics code for enterprises in Chinese cultural society," *Computers in Human Behavior*, vol.28, pp.420-433
- _____ (2017) "Confucian ethics perspective the effect of course style on the information ethics cognition of students," *Management Decision*, vol.55, pp.427-449
- Choi, J., Cho, H., Lee, S., Kim, J. & Park, E.C. (2018) "Effect of the Online Game Shutdown Policy on Internet Use, Internet Addiction, and Sleeping Hours in Korean Adolescents," *Journal of Adolescent Health*, vol.62, pp.548-555
- Myung, J., Lee, H, T. (2013) "A Meta-Analysis of Cyberethics Research and Critical Evaluation from a Perspective of Information Human Rights," *Informatization Policy*, vol.20, no.1, pp.3-21. (in Korean)