ユーザブルセキュリティ研究に向けた情報セキュリティ・プライバシーに関する問題セットの構築

吉川 諒[1]　徐 安然[1]　ゼファン シュラーメク[1]　矢谷 浩司[1]

概要：情報セキュリティやプライバシーについて知っておくべき知識は多様であり，それらについて学ぶことは欠かせない．ユーザブルセキュリティの研究においても，実験参加者等の知識を測るために情報セキュリティ・プライバシーに関する問題を出題することがあり，インターネットに関する知識を網羅的に扱った問題セットの需要は高い．そこで本研究では，情報セキュリティ・プライバシーに関する網羅的な問題リストの作成を目指した．まず，過去 10 年のユーザブルセキュリティに関する論文を分類し，9 の大分類と 40 の小分類を作成した．そして，分類カテゴリに基づいて 90 問の正誤問題を作成し，クラウドソーシング調査（N=900）で正答率を測定した．調査の結果から，問題への正答率と SeBIS のスコアの相関や，年代や居住国，性別による正答率の差異が明らかになった．

Building a Question Set of Online Safety Awareness for Usable Security Research

RYO YOSHIKAWA[1]　ANRAN XU[1]　ZEFAN SRAMEK[1]　KOJI YATANI[1]

1. 研究の背景

インターネットを安全に，適切に利用する上では，情報セキュリティやプライバシーに関する知識が欠かせない．また，情報セキュリティ教育などの場では，受講者や生徒の持つ知識を適切に把握するためのテストも必要となる．インターネットは日常の至る所で用いられており，従って求められる知識も多様となっている．そのため，広範な分野に関する知識を提供したり，知識の定着を確認するための教材が必要となる．また，インターネットに関する知識の測定は，ユーザブルセキュリティ研究の分野においても欠かせないものとなっている．ユーザブルセキュリティの研究では，研究参加者が事前に持っている知識を測定することで，参加者の属性や，調査前後での知識の定着を明らかにすることがある．この際，幅広い分野について扱った問題セットがあれば，研究者は研究テーマに即した問題を選んで利用することができる．

過去の研究では，ユーザの知識や行動を明らかにする指標が開発されてきた．Security Behavior Intentions Scale (SeBIS) [7] は代表的な例である．SeBIS は，ユーザの情報セキュリティに関連する行動について尋ねる 16 問の設問からなり，4 分野（Device securement, Password generation, Proactive awareness, Updating）の特性を測定することができる．これらの指標は，インターネットに関する一側面についての能力を明らかにするために有用であると言える．しかし，インターネットに関する知識を幅広く扱っているわけではなく，教育などの用途を想定したものではない．

そこで本研究では，情報セキュリティ・プライバシーに関する包括的な問題セットを作成し，検証することを目指した．そのためにまず，過去 10 年のユーザブルセキュリティに関する論文（471 編）を収集し，分類することを通じて主要なテーマを明らかにした．結果として，Password & Authentication, Threat, Online security, Fake, Literacy, Privacy & Data management, Data collection, Sensing & IoT, Interface という 9 つの大分類と，それに付随する 40 の小分類を作成した．続いて，これらのテーマをすべて包括する 90 問の問題を作成した．問題の検証のため，クラウドソーシングのプラットフォームを用いた正答率調査を行った．本調査では，様々な国，性別，年代の参加者

[1] 東京大学 Interactive Intelligent Systems Laboratory

を対象することで，属性に応じた情報セキュリティ・プライバシーの認知度の差を検証した．さらに，作成した問題の正答率が，ユーザの知識を測る題材として妥当であるか判断するため，既存の指標である SeBIS のスコアと問題の正答率の相関を調査した．ロジスティック回帰による分析では，問題の正答率と SeBIS のスコアに相関が見られ，特に SeBIS のサブカテゴリのうち，“Proactive” は強い相関を示した．本論文では，正答率調査等の結果について議論した上で，作成した問題セットの評価と，今後の方針について述べる．

2. 関連研究

2.1 情報セキュリティ・プライバシーに関する習慣を測る指標

インターネットの利用者が，情報セキュリティやプライバシーに関するリスクをどのように捉え，振る舞っているかを測定するための指標が開発されてきた．Egelman ら [7] は，インターネット利用者のセキュリティ慣行を測定するための Security Behavior Intentions Scale (SeBIS) を開発した．SeBIS は，device securement，password generation，proactive awareness， updating の 4 つのサブスケールで構成されており，そのサブスケールに関連する合計 16 問のリッカート尺度の設問から構成される．Egelman ら [6] は，SeBIS のサブスケールの指標のスコアが，セキュリティに関する行動と関連していることも検証した．また，Harvart ら [12] は，情報セキュリティやプライバシーに関する姿勢や知識について問うアンケートを開発した．

日本では，総務省 [16] が，青少年のインターネットリテラシーを測定するために，ILAS という指標を開発した．この指標では，違法有害情報リスク，不適切利用リスク，プライバシー・セキュリティリスクの 3 つの分野についてのリテラシーを測定することを目指しており，スマートフォン依存や SNS いじめといった，特に青少年において問題になりやすい事柄について扱っている．

コンピュータの操作などからセキュリティ習慣を推測することを試みた研究もある．Forget ら [8] はコンピュータの操作の記録を長期的に記録する，Security Behavior Observatory (SBO) を開発した．この研究では，収集した情報をもとに，セキュリティインシデントに繋がりうる行動について検証することを目指した．Canfield ら [2] は SBO のデータを用いて，フィッシング検知の能力を推定できる可能性があることを明らかにした．

これらの指標の多くは，若年層など特定の属性の人を対象としていたり，情報セキュリティ・プライバシーに関する特定の側面について着目している．一方で，インターネットが日常で幅広く用いられていることを踏まえれば，特定の分野にとらわれずに，情報セキュリティやプライバシーに関して幅広く扱う問題セットも必要である．そこで本研究では，既存の指標より幅広い分野について知識を問う問題セットの作成を目指した．

2.2 情報セキュリティ・プライバシーの捉え方の文化的・年代的な差異

特定の地域や文化における情報セキュリティ・プライバシーの慣行を調査した研究もある．Chen ら [3] はガーナにおけるインターネットセキュリティの捉え方や慣行を調査し，先進国と似た特徴があることを明らかにした．Sambasivan ら [4] は南アジアの女性の，保有するデバイスへのアクセス権を他人と共有する習慣について調査した．これらの研究からは，地域や文化の違いによりセキュリティ慣行が大きく異なることがわかる．

世代によってインターネットセキュリティに関する認知度が異なることを調査した研究もある．James ら [14] は，65 歳以上の高齢者の情報セキュリティの認知度と，どのように情報を収集しているかについて調査した． この調査では，情報セキュリティに関する情報をインターネットでは収集しないといった高齢者の特徴が明らかとなった． Frik ら [9] は高齢者の IoT に対する態度を検証し，IoT デバイスの使いにくさや知識不足が情報セキュリティに懸念を及ぼすことがわかった．また，セキュリティ上の懸念がある場合，高齢者はデバイスを使わないことで対処することも明らかとなった．このように高齢者を対象とした研究がある一方で，若年層の慣行を調査した研究もある．Cranor ら [5] は，10 代の人とその親で，インターネット技術やプライバシーに関する意思決定や，プライバシーの捉え方が異なることを明らかにした．また，親世代では，意思決定の際に誤った認識を基にしてしまうことがあると分かった．その他にも，若年層に特有とも言える脅威として，ネットいじめやセクスティングなどがあることが明らかとなっている [1,10]．

これらの研究からは，世代や文化，性別などによって情報セキュリティ・プライバシーに関する知識や習慣などが大きく異なることがわかる．インターネットリテラシーに関する普遍的な指標を作る上では，このような属性による差を十分考慮する必要がある．そのため，本研究では，様々な年代，国，性別の人を対象とした調査を行い，参加者の属性に応じた特徴を検証した．

3. 分類カテゴリと問題文の作成

3.1 ユーザブルセキュリティに関する論文の調査

包括的な問題セットを作成するためには，近年の情報セキュリティ・プライバシーの主要なトピックを把握する必要があるが，それらについて網羅的に扱った先行調査は少ない． また，情報セキュリティ等の知見は頻繁に変化するため，最新の動向を問題セットに反映することが重要である．そこで本研究では，近年のユーザブルセキュリティ研

究の主題分析を通じて，情報セキュリティやプライバシーに関する独自のカテゴリを作成することを目指した．分類の対象として，USENIX SOUPS と ACM CHI の過去 10 年分の論文（2013–2022 年に発表されたもの）を選んだ．これらの会議ではユーザブルセキュリティやヒューマンコンピュータインタラクションに関する論文が幅広く発表されており，近年の情報セキュリティ・倫理に関する主要なトピックを網羅的に把握するのに適切であると考えた．

まず，CHI の論文は，ユーザブルセキュリティに全く関係ないものもあるため，予め分類対象とする論文を絞り込んだ． CHI のプログラムのページで論文を絞り込み，発表セッションのタイトルを確認して関連しうる論文を収集した．この際，“Privacy”, “Social Media”, “IoT”, “Transparency” , “Security” を含むセッションについては無条件で分類対象とし，これらが含まれないセッションであっても，関連度が高いセッションを広く対象とした．SOUPS についてはすべての論文を分類対象とし，合計で 471 編（SOUPS: 259 編, CHI:212 編）の論文を収集した．ここではセキュリティ等に関する論文を漏れなく集めることを目的としたため，エンドユーザのセキュリティ・プライバシーに直接関連しない論文も含まれていた．

3.2 論文の分類とカテゴリの作成

続いて，論文をもとにしたカテゴリの作成と論文の分類を行った．最初に，著者のうち 3 名が論文を概観し，複数の論文に共通するキーワードを記録した．先述の通り，471 編の論文には，一般のインターネット利用者に関連しない論文が含まれていた．例えば，組織内の IT 管理者に関する内容や，情報セキュリティの理論的な内容に関するもの，特定のユーザ集団に注目した研究などである．これらは一般利用者向けの問題作成に適さないことから，後の分類作業において分類対象から除外することで合意した．そして，類似の内容の統合などを通じて分類カテゴリの草案を作った．この草案は 9 の大分類で構成されており，さらに細分化した 38 の小分類を作成した．

次に，分類対象の論文を，カテゴリの草案（38 の小分類）に分類する作業を行った．この分類作業は，あらかじめ作成したカテゴリが，分類対象の論文の内容を代表するキーワードとなっているか確認することを目的として行った．まず，著者のうち 2 人が独立して分類を行った．この際，1 つの論文の内容が複数のカテゴリに該当しうる場合は，それぞれの著者が最適と考える 1 つのカテゴリを選んだ．分類作業はカテゴリの網羅性の検証を目的としており，最低一つのカテゴリに論文が当てはまれば網羅性を裏付けることができるため，分類カテゴリの選択が両者で異なることは問題ないと判断した．

1 回目の分類作業の完了後，分類を行った 2 名の著者が分類結果について議論した．ここで，明らかな分類ミスやカテゴリの解釈の齟齬が見つかったため，これらを修正した．更に，多くの論文が含まれているカテゴリについては，細分化することで合意した．そして，細分化前のカテゴリに含まれる論文は，修正後のカテゴリに基づいて再度分類を行った．

最終的に，9 の大分類と，40 の小分類を構築した．最終的に，分類者 A は 332 編，分類者 B は 334 編の論文を分類対象として捉え，各カテゴリに分類した．また，カテゴリへの論文の分類の一致度は 93% , Cohen の Kappa は 0.93 であった．なお，Kappa は以下の数式に基づいて計算した．

$$\kappa = \frac{P_o - P_e}{1 - P_e} = \frac{\sum_i^N p_{ii} - \sum_i^N p_{i\_} * p_{\_i}}{1 - \sum_i^N p_{i\_} * p_{\_i}}$$

p_{ii} ... 分類作業者 A,B が両方カテゴリ i に分類する確率
$p_{i\_}$ ... 分類作業者 A がカテゴリ i に分類する確率
$p_{\_i}$ ... 分類作業者 B がカテゴリ i に分類する確率

表 1 は全カテゴリと，そのカテゴリに分類された論文の数を表している．論文分類の一致度が高いことから，作成したカテゴリは近年の情報セキュリティ・プライバシーに関するトピックを適切に要約していると言える．この結果を踏まえ，今回作成したカテゴリに基づいて問題を作成することとした．

3.3 問題文の作成

作成したカテゴリ一覧に基づいて，正誤問題の作成を行った．問題の作成に際しては，カテゴリ内の論文の内容を参考にして作問した．作成した問題の種類は大きく 2 つに分けられる．1 つはある用語の定義の正誤を問うもので，もう 1 つは，問題文で述べられている状況について，情報セキュリティやプライバシーの観点から適切かどうかを問うものである．問題文の意図を明確にして，回答者によって捉え違いが起きないようにするため，次のことに留意した．

- 明確に表現する．
  例えば，「限定公開の SNS であっても，他人が写っている写真を自己判断で掲載するのは望ましくない.」という問題は，「望ましい」という語が示す点が曖昧なため，適切ではない．そのため，「他人が写っている写真を自己判断で SNS に掲載することは，**プライバシー上のリスクがありうる**」のように明確に表現する．
- すべての場合に当てはまるように注意する．
  例えば，「スマートスピーカは収集した情報を製造元に送信する.」という問題は，すべてのスマートスピーカに当てはまるとは限らないため不適切である．そのため，「スマートスピーカで収集された情報が製造元に送信されることは**絶対に**ない.」などと表現する．

“Out of Scope” と “Specific Users” のカテゴリは，一般

通番	カテゴリ名	分類者 A	分類者 B	問題数
Password & Authentication		**54**	**54**	**17**
A1	Password Policy & Management	20	20	2
A2	2FA	4	4	3
A3	Password Memorization	5	5	2
A4	Alternative Authentication	4	5	2
A5	Biometric Authentication	13	12	2
A6	Implicit Authentication	2	2	2
A7	Usability of Authentication	3	3	2
A8	FIDO	3	3	2
Threat		**44**	**41**	**19**
B1	Phishing& Spam & Social Engineering	12	10	6
B2	Malware & Ransomware	1	1	3
B3	Account compromise	4	5	2
B4	Shoulder / Physical Hacking	9	9	2
B5	Data Breach	6	8	2
B6	Harrasment & Stalking	8	5	2
B7	Cyber Bullying	4	3	2
Online Security		**12**	**13**	**5**
C1	Secure Connection & Communication	8	9	3
C2	Software Update	4	4	2
Fake		**7**	**7**	**8**
D1	Deep Fake & Disinformation	3	3	2
D2	False Information	4	4	6
Literacy		**66**	**66**	**10**
E1	Understanding for Technology	15	15	2
E2	Online Safety Education	8	8	0
E3	Digital Wellbeing	4	2	2
E4	Users' Literacy & Practice	29	32	2
E5	SNS Literacy	5	4	2
E6	Institution's Practice / Policy	5	5	2
Privacy & Data management		**55**	**54**	**15**
F1	Selfie & Privacy Share	16	16	5
F2	Decaying	3	3	2
F3	Access Control	5	4	2
F4	Application Grant	18	18	2
F5	Surveillance & Invasion	10	10	2
F6	Cloud	3	3	2
Data Collection		**36**	**32**	**6**
G1	Infer & Personalization	7	5	2
G2	Ads	7	6	2
G3	Consent & Transparency	22	21	2
Sensing & IoT		**26**	**24**	**6**
H1	IoT Security	5	6	2
H2	IoT Privacy Collection	18	17	2
H3	Healthcare	3	1	2
Interface		**32**	**39**	**4**
I1	Dark Pattern	9	12	2
I2	Transparent Mechanism	6	11	0
I3	Warning & Alert	17	16	2
	分類対象外	139	141	–

表 1: 分類者 A,B があるカテゴリに分類した論文の数と，そのカテゴリに関連する問題の数の一覧.

のインターネットユーザ向けの啓発と関係ないことから，問題を作成しなかった．同様に，“Online Safety Education” と “Transparent Mechanism” のカテゴリも，問題作成の題材として適さないことから，問題作成の対象から除外した．その他のサブカテゴリについては最低 2 問の問題を作成したが，“False Information” や “Shoulder / Physical Hacking” といったカテゴリは，広範なトピックを扱うため，問題を追加し，合計で 90 問を作成した．表 1 の右列に，作成した問題数を併記した.

4. 正答率調査

4.1 調査内容

すべての調査は，クラウドソーシングサービスを通じて，オンラインフォームを用いて行われた．調査内容は前半・後半の 2 部に分けられる.

まず，前半では，作成した問題を出題して正答率を測定した．この際，90 問すべてを出題すると参加者の負担となり，正答率にも影響を及ぼしうるため，20 問を出題した．一方で，回答者ごとに解いた分野に偏りがあると，適切に正答率を測れなくなる可能性がある．そこで予め，各分野から満遍なく問題を収集した 20 問のセットを 9 セット作成し，その中からランダムに 1 セット選んで参加者に提示した．また，この 20 問に加えて，2 問のダミー問題を出題した．ダミー問題は，「この問題には『正しい』と答えてください」といった答えが明確なものであり，参加者が集中して回答しているかを検証するために出題した．出題した全 22 問は，出題順をランダムにして参加者に提示された.

続いて後半では，参加者の性別，年代，居住国を尋ねたほか，SeBIS [7] のスコアを測定した.

4.2 参加者

多くのユーザを想定した正答率調整を行うため，複数の国から参加者を募集することを目指した．本研究ではアメリカ，イギリス，スペイン，日本を調査の対象国とした．これらの 4 カ国は地理的な場所が離れており，文化的な特徴などを表すホフステッド指数 [13] も異なる. そのため，これらの国を対象とすることで，特定の国や文化に偏らない正答率を収集することが可能になる.

参加者の収集には，クラウドソーシングサービス（日本の参加者：クラウドワークス，その他の国の参加者：Prolific）を用いた．各国から 225 人，合計で 900 人が参加した．調査への回答を完了した参加者に，1.21US ドルに相当する額を，それぞれのクラウドソーシングサービスでの通貨で支払った．また，同じ参加者が複数回調査に参加すると，1 回目に回答した問題や SeBIS の設問から，2 回目以降の問題の正答を推測できる可能性があるため，1 回のみ調査に参加できるよう制約を設けた．参加者の属性等について，次節以降で述べる.

5. 結果

5.1 参加者の属性

本調査では，900 名の参加者を募集した．このうち，4 名の参加者が，回答の質を担保するためのダミー問題に誤答した．また，12 名の回答が不完全であった．そのため，これらの参加者のデータは今後の分析から除外した．表 2 は分析対象の参加者の属性である．

男性	440
女性	426
その他	11
回答しない	7
合計	884

(a) 参加者の性別．

20 代	266
30 代	283
40 代	199
50 代	88
60 代以上	48
合計	884

(b) 参加者の年代．

日本	217
スペイン	221
イギリス	223
アメリカ	223
合計	884

(c) 参加者の居住国．

表 2: 分析対象の参加者の属性の分布．

5.2 問題の正答率

図 1 は全 90 問の正答率のヒストグラムである．また，表 3 に，正答率が最も高かった 3 問と最も低かった 3 問を例示した．図 2 は，各カテゴリ内の問題の平均正答率を，参加者の性別・年代・居住国ごとに散布図にしたものである．

5.3 問題の正答率と参加者の属性の関係

図 2 に示した通り，参加者の性別や年代等の属性によって，平均正答率に差があることが明らかになった．そこで，これらの属性によって正答率がどの程度変わるか定量的に分析するために，ロジスティック回帰分析を行った．

表 2a に示されている通り，18 名の参加者は，性別について「その他」や「回答しない」と選択した．回答数が少なく，統計処理に適さないと考えられるため，これらの参加者の回答はロジスティック回帰分析に含めなかった．

ロジスティック回帰分析では，以下のようなモデルを作成した． まず，目的変数は 2 値であり，問題に正解したら 1, 不正解であれば 0 を取る．また，説明変数は以下のように設定した．

(1) SeBIS の 4 つのサブカテゴリ．

- SeBIS の 4 カテゴリ (device securement, password generation, proactive awareness, updating) [7] 毎のスコアを利用した．原論文に従い平均スコアを算出した上で，ロジスティック回帰で扱うためにスコアから 3 点を減算し，値が-2 から 2 を取るようにした．

(2) 性別．

- 先述の通り，性別の回答が男性/女性以外の参加者は回帰分析の対象から除外した．男性である場合に 1，女性である場合に 0 を取るダミー変数を設けた．

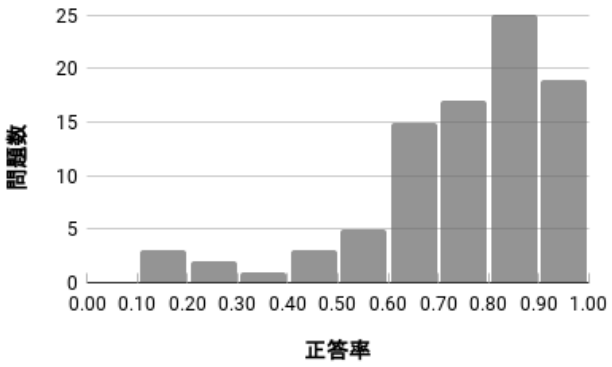
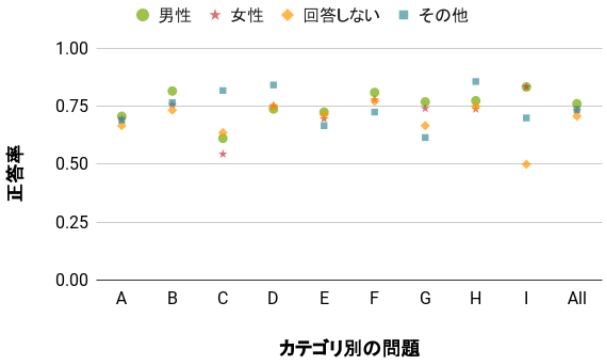
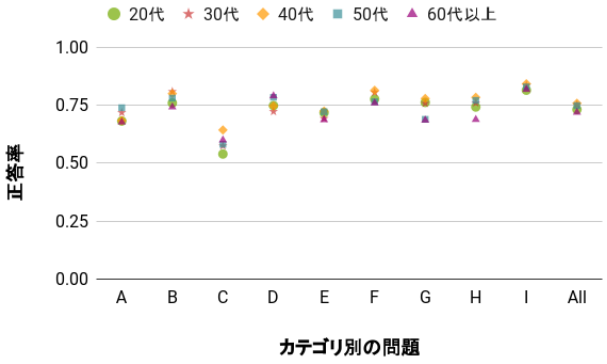


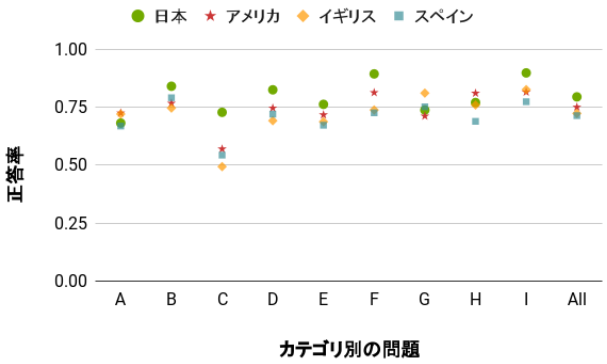
図 1: 作成した 90 問の正答率のヒストグラム．



(a) 正答率の平均を，回答者の性別ごとにプロットしたもの．



(b) 正答率の平均を，回答者の年代ごとにプロットしたもの．



(c) 正答率の平均を，回答者の居住国ごとにプロットしたもの．

図 2: 各カテゴリに属する問題の正答率の平均を，参加者の属性別にプロットした散布図．

(3) 世代.

- 20 代, 30 代, 40 代, 50 代それぞれである場合に 1 をとるダミー変数（合計 4 変数）を設けた．60 代以上の参加者の場合はすべての変数が 0 を取る．

(4) 居住国.

- 日本，イギリス，スペインそれぞれの居住者である場合に 1 をとるダミー変数（合計 3 変数）を設けた．アメリカ居住の参加者の場合はすべての変数が 0 を取る．

上記の内容に基づき，1 人の回答者につき，回答したすべての問題（20 問）ごとに説明変数と目的変数を含めたデータを作成した．そのため，データサイズは分析の対象とした回答者数（866 人）× 回答数（20 問）=17320 行となった，このうち，分類カテゴリごとの分析を行う場合は，そのカテゴリの問題に関する行だけを抽出して分析を行った．

まず，すべての変数を投入してロジスティック回帰分析を行った．続いて，モデルの精度を上げるため，赤池情報量規準（AIC）に基づくステップワイズ法を実施した． 表 4 は，それぞれのカテゴリの問題の平均正答率と，ステップワイズ法を行った際の回帰係数を表している．

6. 考察

6.1 問題のカテゴリ別の正答率について

まず，各カテゴリ内の問題の正答率について着目したところ，問題のカテゴリによって，正答率の傾向が異なることがわかった．例えば，カテゴリ G(“Data Collection”) について，年代によって正答率が大きく異なることがわかる．また，カテゴリ F(“Privacy & Data management”) では，居住国によって大きく正答率が異なっている．

特定の属性の人に着目すると，問題のカテゴリごとに得手不得手があることがわかった．例えば，30 代の参加者はカテゴリ A については有意に高い正答率である一方で，カテゴリ D では有意に低い正答率となっている．また，スペインの参加者は複数のカテゴリで正答率が有意に低い一方，カテゴリ B では正答率が有意に高くなっている．

これらの結果から，あるユーザのインターネットリテラシーを測定するためには，問題を網羅的に出題する必要があることが示唆される．問題の内容によって，大きく正答率が異なるからである．従って，知識を測定するための指標は，幅広い分野の問題によって構成される必要があるといえる．

6.2 SeBIS のスコアと問題セットの平均正答率の関連について

ロジスティック回帰分析では, SeBIS の 4 つのサブスケールのスコアも変数として加えた．結果として，“Proactive awareness” のサブスケールのスコアは，カテゴリ C 以外のすべてのカテゴリの正答率と有意な正の相関があった．“Proactive awareness” は，フィッシング対策などの，予防的な行動に関する慣行を問う内容であり，情報セキュリティに関して積極的な行動を取る人は，関連知識も豊富であると示唆される．また，“Updating” のサブスケールのスコアは，暗号化通信やソフトウェア・アップデートなどを扱うカテゴリ C の正答率と正の相関があった．この結果からは，作成した問題セットが，ユーザのリテラシーを測定するための妥当な題材となりうることが期待される．

しかし，一部の SeBIS のサブスケールについて，本研究で作成した問題セットの，対応するカテゴリの正答率と相関が見られないこともあった．例えば，“Password generation” のサブスケールのスコアは，パスワードや認証について扱ったカテゴリ A の正答率との相関が確認されなかった．この理由として，知識として得ていることと実際に行うことの間に隔たりがあることが考えられる．我々の問題セットでは，取るべき行動について問うているのに対し，SeBIS ではユーザの実際の行動について質問している．例えば，複雑なパスワードを設定するのが望ましいとわかっていつつ，利便性の観点から簡単なパスワードを使う場合があることが指摘されている [11,15].

一方で，SeBIS のスコアと問題セットの正答率が負の相関を持つ，つまり SeBIS のスコアと正答率が食い違う例は確認されなかった．この結果からは，作成した問題セットが，インターネットリテラシーを測定する上で妥当な題材となりうることが期待される．

6.3 参加者の属性に応じた正答率の差

参加者の性別については，女性がカテゴリ B と全カテゴリの平均において正答率が有意に低かったものの，他のカテゴリについては性別の差が正答率に有意な差をもたらすことはなかった．

参加者の年代について，カテゴリ G では，20-40 代の人は 50 代以上の参加者と比べて正答率が有意に高いことがわかった．カテゴリ G はターゲティング広告やプライバシーポリシーなどについて扱うもので，比較的最近の内容であることが正答率に影響をもたらした可能性がある．しかしその他のカテゴリでは，特定の年代の正答率が平均と大きく乖離する傾向はほぼ見られなかった．

先行研究を踏まえると，高齢世代での正答率が下がることなどが想定されたが，本調査ではそのような傾向は確認されなかった．その理由の一つとして考えられるのが，参加者をクラウドソーシングサービスで募集したことである．クラウドソーシングを日常的に利用している人は，年齢によらずインターネットに関する知識を一定水準で有している可能性があり，結果として年代別の正答率の差が顕著に現れなかったと考えられる．

参加者の居住国に着目すると，様々な面で正答率に大きな差があった．特に，日本の参加者は，多くのカテゴリに

正答率ランキング	問題文	正解	正答率
1	ウェブブラウザでの検索内容が，表示される広告の内容に反映される場合がある．	正しい	0.99
2	パスワードを忘れると不便なため，”1234”のような覚えやすいものを用いるのが良い．	誤り	0.98
3	インターネットに掲載された情報は，数十年といった長期間にわたり残る場合がある．	正しい	0.98
...			
88	Cookie には個人が特定できる情報が含まれ，サイト管理者に送られる．	誤り	0.18
89	位置情報に基づいてスマートフォンのロックを解除する技術が開発中である．	正しい	0.12
90	インターネットアカウントの認証方法の FIDO は，指紋などの生体情報を通信でサービス提供者に送るので，セキュリティ上のリスクがある．	誤り	0.11

表 3: 本研究で作成した 90 問のうち，正答率が上位の 3 問と下位の 3 問．

		目的変数（平均正答率）									
		A	B	C	D	E	F	G	H	I	All
	Intercept	**0.624** (**p<.001**)	**1.011** (**p<.001**)	***-0.216*** (**p=.04**)	**0.906** (**p<.001**)	**0.54** (**p<.001**)	**1.121** (**p<.001**)	0.242 (p=.20)	**1.112** (**p<.001**)	**1.081** (**p<.001**)	**0.896** (**p<.001**)
SeBIS	Device	0.081 (p=.06)				**0.117** (**p=.04**)	**0.169** (**p=.002**)	0.148 (p=.052)	0.148 (p=.06)	**0.246** (**p=.02**)	**0.097** (**p<.001**)
SeBIS	Password		**0.129** (**p=.02**)		*-0.151* (p=.06)		*-0.124* (p=.07)				
SeBIS	Proactive	**0.144** (**p=.02**)	**0.211** (**p<.001**)	0.192 (p=.07)	**0.316** (**p<.001**)	**0.208** (**p=.01**)	**0.346** (**p<.001**)	**0.366** (**p<.001**)	**0.206** (**p=.04**)	**0.335** (**p=.02**)	**0.238** (**p<.001**)
SeBIS	Update	0.100 (p=.06)		**0.455** (**p<.001**)							
性別	女性		***-0.286*** (**p<.001**)				*-0.182* (p=.06)				***-0.134*** (**p<.001**)
年代	20 代					0.173 (p=.14)	0.182 (p=.13)	**0.473** (**p=.02**)			
年代	30 代	**0.169** (**p=.048**)	**0.237** (**p=.01**)		***-0.290*** (**p=.03**)		0.183 (p=.12)	**0.425** (**p=.04**)			
年代	40 代				*-0.263* (p=.10)			**0.653** (**p<.001**)			
年代	50 代	0.197 (p=.15)									
居住国	日本		**0.609** (**p<.001**)	**1.076** (**p<.001**)	**0.725** (**p<.001**)	**0.431** (**p<.001**)	**0.79** (**p<.001**)			**0.903** (**p<.001**)	**0.363** (**p<.001**)
居住国	イギリス						***-0.483*** (**p<.001**)	**0.423** (**p=.01**)	*-0.279* (p=.11)		***-0.134*** (**p=.01**)
居住国	スペイン	***-0.216*** (**p=.01**)	**0.203** (**p=.04**)			*-0.177* (p=.15)	***-0.551*** (**p<.001**)		***-0.672*** (**p<.001**)	*-0.362* (p=.1)	***-0.195*** (**p<.001**)
	残差逸脱度	3979.9	3718	1204.8	1700.6	2281.4	2796.1	1276.7	1222.2	697.12	19366
	AIC	3993.9	3732	1212.8	1712.6	2293.4	2816.1	1290.7	1232.2	707.12	19380

表 4: ステップワイズを用いたロジスティック回帰分析の結果．目的変数は，カテゴリ A-I に属する問題の平均正答率と，全問題の平均正答率である．説明変数は，性別，年代，居住国と，SeBIS のスコアである．空欄は，ステップワイズ法を行う過程で省かれた変数を表す．また，太字は有意な係数，斜体は負の係数を表す．

おいて有意に高い正答率を示した．一方でスペインの参加者は，カテゴリ B などで正答率が有意に低かった．日本での正答率が高くなった要因として，調査参加者の特性や，日本での教育の特徴などが考えられる一方で，調査の設計が影響を及ぼした可能性もある．本調査では，日本での参加者募集のみ，他の 3 国とは違うクラウドソーシングサービスを用いたほか，問題なども日本語で出題した．問題文は日英それぞれの母語話者が検証し，内容や表現が同じものになるように確認しており，影響は限定的であると考えられる．しかし，問題セットが多くの国で利用されるためには，翻訳を介することによる正答率への影響について検証する必要がある．

7. 本研究の制約と今後の展望

本研究には複数の制約が存在する．まず，本研究では，クラウドソーシングによって参加者を募集した．そのため，参加者は一般平均よりも，インターネットを使い慣れており，関連知識を有していた可能性がある．例えば，今回の調査では年代の違いによって顕著な正答率の差は見られなかったが，これは高齢者であっても，知識を有している参加者であったことが理由である可能性がある．また，調査では 4 カ国から参加者を募集したが，すべての国が先進国であった．このように，年代や居住国の面で，十分に幅広い属性のインターネット利用者が調査に参加したとは言い難い．今後の調査では，インターネットをあまり使わない人など，より多様な人を対象とした調査を行い，認知度の差異などを検証する．

また正答率調査を通じて，作成した問題の正答率に大きな差があることがわかった．正答率が大きく異なる問題が混在していることで，幅広い分野に関する知識を測定する上で支障が出ることが懸念される．また，どの程度の難易度の問題が，ユーザの知識を測る上で適切かも明らかではない．今後の研究では適切な難易度を定め，各問題をその難易度に揃えることが必要になる．

正答率が低かった問題は，問題の文言や答えが不適切であったり，過度に専門的な内容であることで，参加者が正解に至らなかった可能性も考えられる．正答率調査では，作成した 90 問のうち 14 問において，正答率が 6 割を下回った．本研究では基礎的な事項についての問題作成を目指しているため，正答率が過度に低い問題は目的に合致しない．そこで，情報セキュリティ・プライバシーの知識を有する人に問題の検証を依頼し，不適切な問題や極端に専門的なものを省くなどの改善を行う予定である．

8. 結論

本研究では，情報セキュリティ・プライバシーに関するトピックを網羅的に扱う問題セットの作成を目指した．そのために，まずユーザブルセキュリティに関する近年の論文の分類を行い，11 のカテゴリと 42 のサブカテゴリを作成した．次に，このカテゴリに基づいて合計 90 問の正誤問題を作成した．ロジスティック回帰分析によって，SeBIS の複数の指標が，本研究で作成した問題の正答率と相関することがわかった．今後の調査では，より広範な属性の参加者を集め，今回作成した問題セットが，インターネットに関する知識の測定に普遍的に役立つかを検証する．

謝辞 本研究は，株式会社メルカリとインクルーシブ工学連携研究機構の共同研究である，価値交換工学の成果の一部です．また，調査にご参加頂いた皆様と，貴重な指摘を下さった矢谷研究室の皆様に深く感謝いたします．

参考文献

[1] Zahra Ashktorab and Jessica Vitak. Designing cyberbullying mitigation and prevention solutions through participatory design with teenagers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, p. 3895–3905, New York, NY, USA, 2016. Association for Computing Machinery.

[2] Casey Canfield, Alex Davis, Baruch Fischhoff, Alain Forget, Sarah Pearman, and Jeremy Thomas. Replication: Challenges in using data logs to validate phishing detection ability metrics. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pp. 271–284, Santa Clara, CA, July 2017. USENIX Association.

[3] Jay Chen, Michael Paik, and Kelly McCabe. Exploring internet security perceptions and practices in urban ghana. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, pp. 129–142, Menlo Park, CA, July 2014. USENIX Association.

[4] Jessica Colnago, Lorrie Faith Cranor, Alessandro Acquisti, and Kate Hazel Stanton. Is it a concern or a preference? an investigation into the ability of privacy scales to capture and distinguish granular privacy constructs. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, pp. 331–346, Boston, MA, August 2022. USENIX Association.

[5] Lorrie Faith Cranor, Adam L. Durity, Abigail Marsh, and Blase Ur. Parents' and Teens' perspectives on privacy in a Technology-Filled world. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, pp. 19–35, Menlo Park, CA, July 2014. USENIX Association.

[6] Serge Egelman, Marian Harbach, and Eyal Peer. Behavior ever follows intention? a validation of the security behavior intentions scale (sebis). In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, p. 5257–5261, New York, NY, USA, 2016. Association for Computing Machinery.

[7] Serge Egelman and Eyal Peer. Scaling the security wall: Developing a security behavior intentions scale (sebis). In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, p. 2873–2882, New York, NY, USA, 2015. Association for Computing Machinery.

[8] A. Forget, S. Komanduri, A. Acquisti, N. Christin, L.F. Cranor, and R. Telang. Security behavior observatory: Infrastructure for long-term monitoring

of client machines. Technical Report 14-009, CyLab, Carnegie Mellon University, July 2014.

[9] Alisa Frik, Leysan Nurgalieva, Julia Bernd, Joyce Lee, Florian Schaub, and Serge Egelman. Privacy and security threat models and mitigation strategies of older adults. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, pp. 21–40, Santa Clara, CA, August 2019. USENIX Association.

[10] Christine Geeng, Jevan Hutson, and Franziska Roesner. Usable sexurity: Studying People's concerns and strategies when sexting. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, pp. 127–144. USENIX Association, August 2020.

[11] Ameya Hanamsagar, Simon S. Woo, Chris Kanich, and Jelena Mirkovic. Leveraging semantic transformation to investigate password habits and their causes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–12, New York, NY, USA, 2018. Association for Computing Machinery.

[12] Franziska Herbert, Florian M. Farke, Marvin Kowalewski, and Markus Dürmuth. Vision: Developing a broad usable security & privacy questionnaire. In *Proceedings of the 2021 European Symposium on Usable Security*, EuroUSEC '21, p. 76–82, New York, NY, USA, 2021. Association for Computing Machinery.

[13] Hofstede Insights. Country comparison. `https://www.hofstede-insights.com/`.

[14] James Nicholson, Lynne Coventry, and Pamela Briggs. "if it's important it will be a headline": Cybersecurity information seeking in older adults. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–11, New York, NY, USA, 2019. Association for Computing Machinery.

[15] Rick Wash, Emilee Rader, Ruthie Berman, and Zac Wellmer. Understanding password choices: How frequently entered passwords are re-used across websites. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pp. 175–188, Denver, CO, June 2016. USENIX Association.

[16] 総務省. 青少年がインターネットを安全に安心して活用するためのリテラシー指標等に係る調査. `https://www.soumu.go.jp/use_the_internet_wisely/special/ilas/`.2023 年 5 月閲覧.