IMES DISCUSSION PAPER SERIES

ユーザブルセキュリティ研究の現状と展望

金岡 晃（かなおか あきら）・井上紫織（いのうえ しおり）

Discussion Paper No. 2026-J-8

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

日本銀行金融研究所

〒103-8660 東京都中央区日本橋本石町 2-1-1

IMES Discussion Paper Series 2026-J-8
2026 年 9 月

ユーザブルセキュリティ研究の現状と展望

金岡 晃（かなおか あきら）*・井上紫織（いのうえ しおり）**

要　　旨

ユーザブルセキュリティとは、システム利用者の認知や行動特性を踏まえてセキュリティ技術等の実効性を高めることにより、情報および情報システムを不正なアクセス等から保護することである。継続的にセキュリティ事案が発生する中、これを実現する方法や知見を探求することを目的とした研究領域である「ユーザブルセキュリティ研究」の重要性が高まっている。これまでに、利用者認証やアクセス制御といった、情報システムに実装されるセキュリティやプライバシー保護に関する技術要素の実効性を高める観点から、利用者の操作画面や操作手順の設計について研究が進められてきた。さらに、近年では、フィッシング攻撃や偽情報による金融詐欺など、人間の認知や行動特性を悪用した攻撃への対策へと研究対象が拡大している。多様な研究対象に取り組むうえで、適切な研究手法を選択し組み合わせるアプローチが必要なほか、心理学や社会学など多様な学問の知見が求められる。本稿では、こうしたユーザブルセキュリティ研究における手法や具体的な研究事例を解説するとともに、研究の透明性の確保に向けた近年の動向を整理し、今後の研究における課題と方向性を展望する。

キーワード：ユーザブルセキュリティ、ユーザブルプライバシー、量的研究、質的研究、研究の透明性
JEL classification: D81、L86、O33

* 東邦大学理学部教授（E-mail: akira.kanaoka@is.sci.toho-u.ac.jp）
** 日本銀行金融研究所企画役（E-mail: shiori.inoue@boj.or.jp）

本稿は、日本銀行金融研究所からの委託研究論文である。本稿の作成に当たっては、山田明教授（兵庫県立大学）から有益なコメントをいただいた。ここに記して感謝したい。ただし、本稿に示されている意見は、筆者たち個人に属し、日本銀行の公式見解を示すものではない。また、ありうべき誤りはすべて筆者たち個人に属する。

目　次

１．はじめに

インターネットなど高度な情報システムの発展と普及は、社会に利便性の向上をもたらした一方で、新たなセキュリティ上の脅威も生み出している。その対策として、暗号化技術やアクセス制御、利用者認証といった、セキュリティやプライバシー保護のための高度な技術や機構が開発・実装されてきた。しかし、これらの機構は、利用者による警告画面の読み飛ばしや、セキュリティ設定の誤認、アプリケーションの開発者による過剰な権限付与など、人間の認知や行動特性に起因する問題により想定どおりに機能しない場合があり、結果的にセキュリティ事案に繋がるケースも発生している。

こうした状況を背景に、システム利用者の認知や行動特性を踏まえてセキュリティ技術等の実効性を高めることにより、情報および情報システムを不正なアクセス等から保護すること（ユーザブルセキュリティ）や、これを実現する方法や知見を探求することを目的とした研究領域である「ユーザブルセキュリティ研究」の重要性が高まっている。

この研究領域では、特にセキュリティ機構そのものを、人間が簡単に正しく理解し、間違えずに操作できるよう、主にセキュリティ機構の「使いやすさ」に着目した取り組みが進められてきた。例えば、情報システムの利用者の操作画面（User Interface：UI）や操作手順における改善を通じて、セキュリティを人と技術の両側面から捉える新たなアプローチが提案されている。

さらに近年、サイバー攻撃の巧妙化に伴い、悪意をもって人間の判断や認知を歪めようとする攻撃が増加しており、ユーザブルセキュリティ研究の対象は人の理解や意思決定、行動そのものへと広がっている。例えば、銀行やオンラインサービスからの正当なメールやメッセージを装い情報システムの利用者の個人情報を盗み取るフィッシング攻撃や、人間の心理的な弱みや油断を利用して情報を聞き出したり操作したりするソーシャルエンジニアリング、利用者に意図しない行動を取らせることを狙った偽情報やダークパターン[1]の悪用などが挙げられる。これらの攻撃は、利用者の判断を歪めようとする攻撃であり、技術的に強固なセキュリティ機構の開発のみでは十分に対処できない課題となっている。このような状況を踏まえ、ユーザブルセキュリティ研究の分野では、人々が安全かつ主体的に社会生活において情報技術を利用できるように、セキュリティやプライバシーの社会的な在り方や人間に対する影響を探究する研究が進められている。こうした研究の中では、利用者に対するセキュリティ教育や啓発活動、

[1] ダークパターンとは、Web サービスやアプリケーションの UI 設計において、情報や選択肢の提示の仕方に細工を施すことにより、情報システムの利用者が十分に意識しないまま特定の選択を行うよう誘導したり、望ましくない行動を取らせたりする可能性のある設計手法を指す。

分かりやすいプライバシーポリシーの提示など、技術にとどまらず社会的な取り組みを含めた検討も行われている。

本稿では、ユーザブルセキュリティ研究を便宜的に「狭義」と「広義」に分類し、まず第 2 節において、それぞれの定義と研究対象の範囲を明確化する。第 3 節では、代表的な研究手法の一つである量的研究について、基本的な分析手法と分析で利用される代表的なアンケートを整理したうえで、具体的な研究事例を示す。第 4 節では、もう一つの代表的な研究手法である質的研究について、基本的なデータ取得方法と分析手法、妥当性・信頼性確保の方法を整理したうえで、具体的な研究事例を示す。第 5 節では研究の透明性確保に向けた対応について、コミュニティ内の議論の状況と課題を整理する。第 6 節では、ユーザブルセキュリティ研究の今後を展望し、第 7 節で議論全体を簡潔にまとめる。

２．研究の概要

（１）狭義のユーザブルセキュリティ研究

情報システムを安全に使うためには、セキュリティに関する仕組みが必要であるが、その仕組みが複雑すぎると、人間が操作を間違えたり、誤った判断をしたりしてしまい、セキュリティが低下する場合がある。狭義のユーザブルセキュリティ研究とは、こうした課題を解決するために、人間がセキュリティ機構を簡単に正しく理解し、間違えずに操作できるよう、同機構の「使いやすさ」を向上させることを主眼に置いた研究領域である[2]。

研究対象としては、各種のセキュリティ機構、例えば、利用者が本人であることを証明するために用いるパスワード認証や顔認証、指紋認証などの「利用者認証」のほか、特定のファイルや情報にアクセスできる人を制限する「アクセス制御」、データを安全に保つための「暗号ライブラリー」、悪意あるサイトであることなどの危険を知らせる「警告インターフェース」、スマートフォンアプリケーションで利用する機能を許可するかどうかを決める「アプリケーション権限設定」などが挙げられる。研究では、これらのセキュリティ機構が適切に理解され、簡単かつ正確に操作されるよう、利用者の操作画面や操作手順の在り方が検討されている。また、研究を進める中では、人がどのように情報を理解し、どのように操作を進めるかといった行動特性を踏まえることが重要となる。

以下では、こうした狭義のユーザブルセキュリティ研究の具体例をいくつか

---

[2] 本研究領域は、1996 年に Zurko らが提唱した User-Centered Security（UCS）に端を発する（Zurko and Simon 1996）。当初はシステムの管理や運用において利用者を主体として捉える概念であったが、その後、セキュリティ機構の認知負荷や操作性を重視する狭義のユーザブルセキュリティ研究へと発展した。

紹介する。

イ．スマートフォン向けモバイル OS におけるアプリケーション権限設計

初期のスマートフォン向けモバイル OS（オペレーティングシステム）では、アプリケーションをスマートフォンにインストールする時に、そのアプリケーションがスマートフォン内で利用できる機能や情報に対するアクセス権（パーミッション）の許可を、利用者に一括で求める方式が採用されていた。しかし、その後の利用者調査や実験の結果、この方式にはいくつかの問題があることが明らかとなった。まず、利用者の多くがパーミッションの説明をよく読まないことが多く、どのアクセス権を許可しているのか、また許可することによってどのようなリスクが存在するのかを完全には認識していないケースが散見された。また、アプリケーションの開発者が、特定の機能を実現するために必要以上の権限を要求する過剰権限の問題が発生していることも確認された。

例えば、Felt らは、アプリケーションをインストールする際に表示されるパーミッション画面について調査し、利用者の行動を観察する実験を実施した。その結果、多くの利用者が権限表示に十分な注意を向けないほか、内容を深く理解しないまま許可を与えてしまう場合があることを確認した（Felt *et al.* 2012）。また、過剰権限の問題について構造的な要因を分析したところ、アプリケーション開発者が必要な機能を実現するために API[3]を利用する際、各 API を利用するためにアプリケーションが必要とする権限に関する情報が、API 提供者が公開する開発者向けドキュメントに十分に記載されていないことが明らかになった。これにより、開発者が API 利用に必要な権限を正しく把握することが困難になり、必要最小限の権限のみを要求するアプリケーションを設計することが難しくなりうることが報告された（Felt *et al.* 2011）。

こうした課題を踏まえ、特定のスマートフォン向けモバイル OS では、位置情報や連絡先、カメラ、マイクなど、セキュリティ漏洩のリスクが高い権限については、アプリケーションのインストール時に一括でパーミッションを求めるのではなく、利用者がそれらの機能を利用しようとするときに求める方式へと変更した（Google［2026］）。

これらの知見を踏まえると、スマートフォン向けモバイル OS におけるパーミッション設計を検討するうえでは、利用者のパーミッションの判断およびアプリケーション開発者のパーミッション要求の設定を支援する仕組みを実装することが重要であるといえる。具体的には、利用者の判断支援の観点からは、実際に機能を利用しようとする場面でパーミッションを求めることにより、利用

[3] API（Application Programming Interface）は、ソフトウェアやプログラム間で互いに機能やデータを連携・通信するための仕組みである。

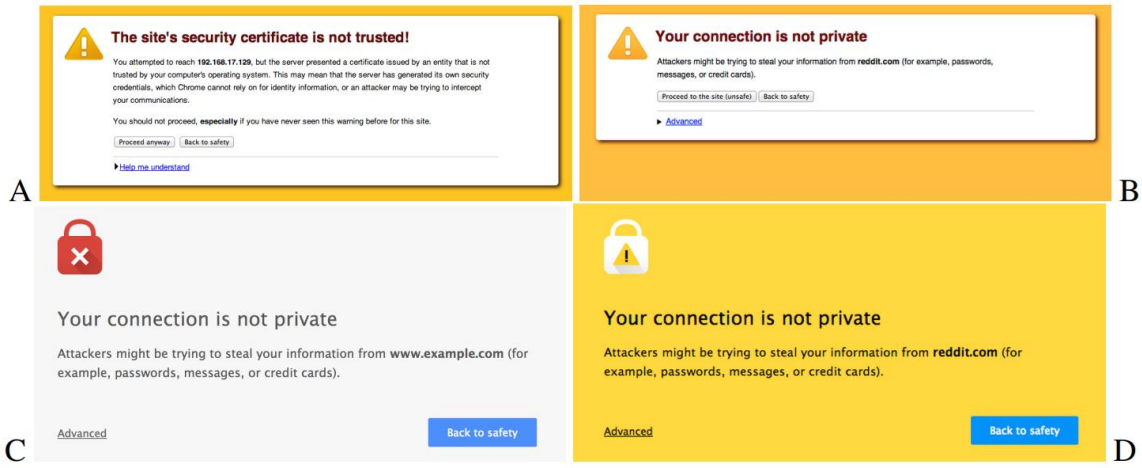
者が「いま何をしようとしており、そのためにどの情報や機能へのアクセスが必要なのか」を把握しやすいようにすることや、利用者がいったんパーミッションの可否を判断した後に、その判断を自由に変更または撤回できる設計とすることが考えられる。また、アプリケーション開発者の支援の観点からは、スマートフォンの API の仕様やその説明ドキュメントを分かりやすく整備することや、デフォルト設定でパーミッションを必要最小限なものとなるように設計することにより、開発者側の意図しない誤った設定を防ぐことなどが考えられる。

ロ. Web ブラウザにおけるセキュリティ警告画面の設計

インターネットブラウザに表示される警告メッセージについては、早くからその効果が高くないことが指摘されてきた。代表的な研究として、Sunshine らが 2009 年に行ったオンライン調査では、インターネット上で安全にデータをやり取りするための通信プロトコルである SSL（Secure Sockets Layer）に関する警告について、利用者に認識されているにもかかわらず、安全な行動に向けた利用者の判断に十分に活かされていない点を明らかにした。そのうえで、新たな警告デザインのほか、安全でないウェブサイトへのアクセスをブロックする仕組みなど、警告に依存せずにセキュリティを高めるアプローチを提案した（Sunshine *et al.* 2009）。その後、Akhawe らは、主要な Web ブラウザを対象とした大規模な調査を実施し、2,500 万件を超える実際の SSL 警告表示に対する利用者行動を計測し、どのような設計が警告に対する利用者の遵守行動につながるかを定量的に評価した（Akhawe and Felt 2013）。

これらの研究結果を踏まえ、ブラウザ開発に関わる研究チームは、利用者の警告表示に対する「理解と遵守」を改善することを目的に、警告表示の大きさや色、レイアウト、アイコン、文面の工夫により可読性を上げ、利用者が次に取るべき安全な行動を分かりやすく示すような設計を提案した（図 1 を参照、Felt *et al.* 2015）。本提案に基づく設計は主要 Web ブラウザの一つにおいて実装され、利用者による警告遵守率が改善したことが報告されている（Kovacs 2015）。以降もブラウザの警告に関する仕組みは改善を重ねている。例えば、HTTP 通信に対する安全性表示の見直し（Schechter 2017）や、危険なダウンロードに関する警告表示における説明表現や警告内容の分かりやすさの改善のほか、リアルタイムで悪性コンテンツを検知する仕組みが導入されるなど、利用者の警告に対する「理解と遵守」を高めるための工夫が継続的にアップデートされている。

図 1 SSL 警告に関する実験で用いられた警告表示



（注）A は Google Chrome 36 で使用された警告画面であり、危険な選択肢（次のページに進む）と安全な選択肢（前のページに戻る）が並列に表示されている[4]。これに対して、B~D は利用者の警告表示に対する理解と遵守を改善することを目的に提案・検証された警告画面の案。B は、A からデザインは変更せずに、次のページに進んだときに具体的に生じうる被害に関する記載へと警告文書を変更した案。C は、警告文書は B と同様としたうえで、背景色を灰色とし、安全な選択肢を選択しやすい構造とした案。D は C の背景色を一般的に警告色とされる黄色に変更した案。実験の結果、利用者による順守率が最も高かった C が Google Chrome 37 の警告画面として採用された。

（資料） Felt *et al.* 2015 の Figure 4.

ハ．メッセージングにおけるエンドツーエンド暗号化（E2EE）の大規模実装

PC やスマートフォンの利用者間で行うリアルタイムのコミュニケーションを仲介するメッセージングプラットフォームでは、従来、メッセージの送信者の端末とプラットフォームのサーバとの間の通信は暗号化されていたものの、メッセージはプラットフォームのサーバ上で復号され、受信者に送信される構成が一般的であった。

一方で、エンドツーエンド暗号化（End-to-End Encryption：E2EE）は、送信者から受信者までの通信を完全に保護するための暗号化方式で、通信の送信者と受信者のみがメッセージ内容を復号できるように設計されており、通信サービスの提供者や中継者を含む第三者は内容にアクセスすることができない。ただし、E2EE の実用化に当たっては、単に暗号プロトコルを実装するだけでなく、利用者が誤用なく安全を得られるようにするユーザブルセキュリティの視点が重要となる。

---

[4] Google Chrome は、Google LLC の商標。

2010 年代前半に整理された、安全なメッセージ通信（セキュアメッセージング）に関する知見を体系的に整理した論文[5]では、セキュアメッセージングを実現するためのシステム設計に求められる枠組みとして、①利用者が特別な知識や技術を持っていなくても簡単に導入できること（導入のしやすさ）、②メッセージ通信を安全に行うためのシステムやアプリケーションが、いつでもどこでも使える状態であること（可用性）、③暗号化したメッセージを受信者本人だけが読み取れるよう、利用者が意図した相手と実際に通信していることを確認するための技術や手法が用意されていること（認証手段の確立）、④メッセージのやり取りが中断されても、それを問題なく再開できる仕組みがあること（メッセージとセッションの継続性）といった、ユーザブルセキュリティの視点を取り入れた要件が示された（Unger *et al.* 2015）。

こうした流れを受け、2016 年には、大規模なメッセージングプラットフォームで、暗号化や復号に用いる鍵管理を自動化することで、利用者が通常の操作感のまま E2EE を利用できる環境が実現した[6]。これにより、利用者間でのチャットや通話、複数利用者によるグループ会話や画像・動画等の添付データについても E2EE による保護が標準設定となり、通信中やサーバに保存されている間にもメッセージ等の内容が暗号化されるようになった。これを契機に、大規模プラットフォームにおいて、E2EE を利用者の操作性に影響を与えないかたちで標準設定として導入することが本格化し、他のサービスにおいても、E2EE が標準設定あるいは任意選択の機能として提供されるなど採用が相次いだ。このように、利用者の操作性を維持しつつ鍵管理の負担を最小化することにより、利用者が暗号を誤用することなく適切に日常利用できる状況が実現可能となった。

（２）広義のユーザブルセキュリティ研究

広義のユーザブルセキュリティ研究とは、人々が安全かつ主体的に社会生活において情報システムを利用できるように、セキュリティやプライバシーの社会的な在り方や人間に対する影響に主眼を置いた研究領域であり、これまで Human-Centered Security and Privacy と呼ばれてきた研究領域に重なる。研究対象となるのは、セキュリティ機構そのものではなく、それを利用する人間の理解と

---

[5] このように、既存の研究や知見を体系的に整理し、特定の分野を網羅的に理解するための視点や課題を提供することを目的とした論文は SoK（Systematization of Knowledge）論文と呼ばれる。

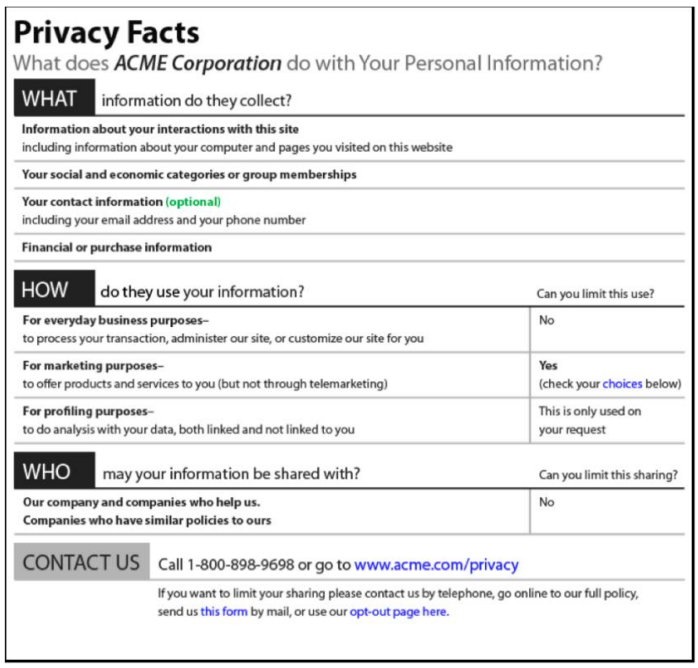
[6] 具体的には、メッセージングプラットフォームの通信インフラに、Signal Protocol と呼ばれるオープンソースの暗号通信方式を採用することにより、E2EE を実現した。Signal Protocol では、通信のたびに自動で新しい暗号鍵を作る仕組みである二重ラチェットアルゴリズム（Double Ratchet Algorithm）等を用いることで、過去の通信内容の漏洩を防ぐ前方秘匿性（PFS：Perfect Forward Secrecy）や将来の秘匿性を高度に担保する。

行動や意思決定に影響を与える要因である。具体的には、フィッシング攻撃（Dhamija, Tygar and Hearst 2006）やソーシャルエンジニアリング、偽情報、ダークパターン（Gray *et al.* 2018、Mathur *et al.* 2019）、セキュリティ教育や啓発、プライバシーポリシーの理解支援などである。研究では、心理学や社会学、教育学などを含む多様な学問の視点を取り入れ、人間の理解、行動、意思決定に影響を与える手法や教育プログラム、情報提示の在り方について設計し、それを評価・検証することなどが行われている。

イ. Privacy Nutrition Label

従来のセキュリティ機構では十分に扱われてこなかったプライバシー情報の収集・利用状況を、利用者に分かりやすく示すフォーマットとして、Kelley らが提案した Privacy Nutrition Label が知られている（図 2 を参照）。これは「食品の栄養成分表示」に着想を得て、視覚的・構造的な要素を取り入れることで、モバイル端末などの情報システム利用時に、OS やソフトウェア、アプリケーションがどのように利用者の個人情報を収集・使用・共有するのかを、利用者が直感的に理解しやすいよう工夫されたものである（Kelley *et al.* 2009）。その後、同様の表示形式は、主要なモバイルアプリケーション配信プラットフォームにおけるアプリケーション公開時のプライバシー表示や、アプリケーションによるデータ取扱いに関する開示制度にも採用された（Apple 2026、Frey 2022）。

図 2 Privacy Nutrition Label の例



（資料） Kelley *et al.* 2009 の Figure 3.

ロ. SNS 上の情報操作および偽情報拡散の現状と対策

SNS（ソーシャルネットワーキングサービス）におけるフェイクニュースなどの偽情報への対策は、フィルタリングによる技術的な防御だけでなく、人間の考え方や行動の特性を考慮した予防的な対策を講じることが有用である。

こうしたアプローチの一つとして、心理学の「接種理論（Inoculation Theory）[7]」を応用したプレバンキングとよばれる手法が挙げられる。本手法は、SNS 利用者に偽情報の特徴や手口をあらかじめ周知することで、人々が偽情報に騙されにくくする手法である。偽情報を広める際に使われる主なテクニックとして、恐怖をあおる、敵対関係を強調する、専門家になりすますといった手口がある。例えば、心理学者である Basol らや Roozenbeek らの研究では、クイズ形式を用いた体験型学習や短い動画を予め利用者に見せることによって、利用者が偽情報で用いられる操作的な表現や誘導の仕方を見分ける力や偽情報に対処できるという自信を高められることが示された（Basol, Roozenbeek, and van der Linden 2020、Roozenbeek *et al.* 2022）。これらの効果は、大規模な動画共有プラットフォームを用いた実証実験でも確認された。近年では、こうした研究の知見をもとに、SNS に関する政策や SNS を提供するプラットフォームにおいて、プレバンキングを用いた対策の実装が広がっている。

３．量的研究によるアプローチ

（１）量的研究の概要

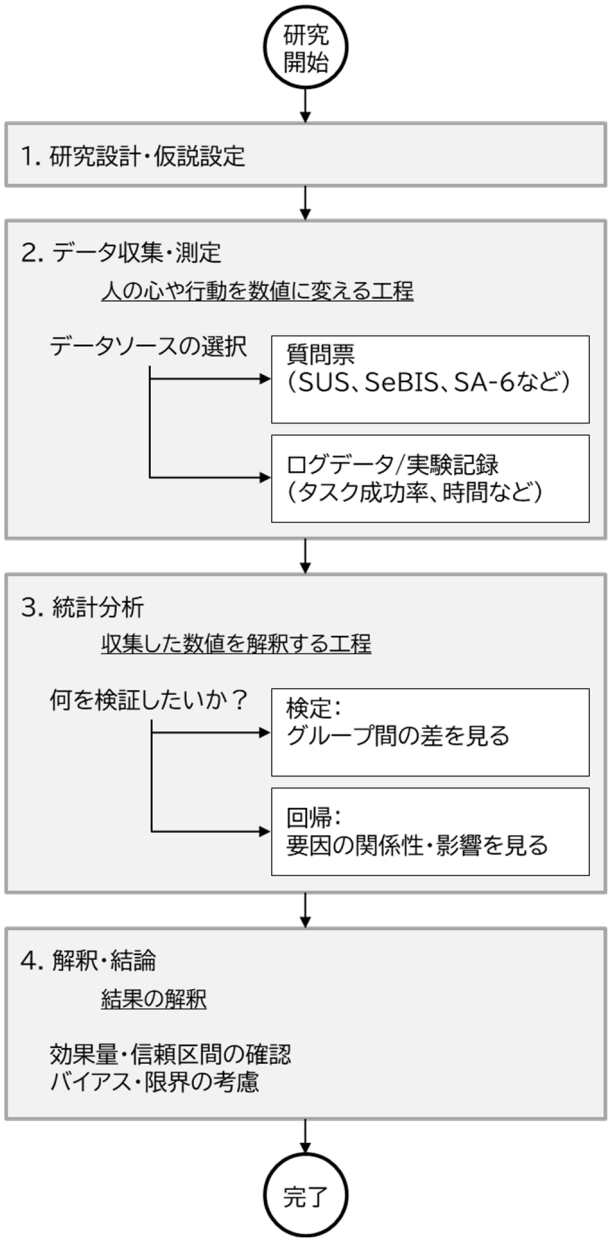
本節では、ユーザブルセキュリティ研究で用いられる代表的な研究手法の一つである量的研究について説明する。量的研究とは、人の理解や行動、意思決定を数値化して収集し、統計的手法により一般化可能な知見を導く方法である。ユーザブルセキュリティ研究では、実験や調査により収集したデータや、情報システムの運用ログ長期時系列データを用いて、一定の仮説を設定したうえで、統計的な検定や回帰分析を行い、その仮説の正しさを評価することにより、各種要因が人間の行動等に与える影響の大きさを評価する（図 3 を参照、検定や回帰分析の具体的な手法については補論の（１）（２）を参照）。

量的研究では、用いるデータの特性に応じて、仮説を明確に設定し、評価に用いる指標を予め定義するとともに検証結果の影響の大きさ（効果量）やその結果の不確実性を適切に推定することが重要となる。本節では、よく用いられる分析手法について、それを用いる場合の前提条件と、留意点について整理する。

---

[7] 接種理論とは、将来的に受ける可能性のある説得や誤情報に対し、その弱い形や代表的な論点をあらかじめ提示・学習させることで、後の影響を受けにくくするという心理学の理論である。

図 3 量的研究の流れ



（２）量的研究に用いられる分析手法

データには、カテゴリや順序、間隔、比などのように多様な種類が存在し、それぞれの特性（尺度）に応じて、適用可能な分析手法や評価手法が異なる。例えば、リッカート尺度は、アンケートなどでよく用いられる評価手法の一つで、選択肢が「1: 非常に否定的」から「5: 非常に肯定的」などのように段階で分かれている。このようなデータは、一般的に「順序尺度[8]」として扱われ、中央値や

[8] データの順番や大小関係に意味があるが、値同士の差には数値的な意味がない尺度のこと。

分布の形状（値がどこに集中しているか、データがどう偏っているかなど）が重視される。ただし、ユーザブルセキュリティ研究では、これを便宜的に「間隔尺度[9]」とみなして平均値や分散（データのばらつきの程度）で評価する場合もある。この場合、回答が選択肢の上限や下限に集中して、正確に平均値や分散を評価できなくなる天井効果や床効果が発生することがありうるほか、回答が偏った分布を持つ場合や離散的であまり均等でない場合には、平均値や分散からでは正確な状態を反映しにくくなるため、留意が必要となる。

つづいて、実験を計画する際の具体的な検討事項を整理する。まず、どのような仮説を立て、実験結果をどのような指標で評価するのかについて明確にする必要があるほか、信頼できる結果を得るためには、どれくらいのサンプル数が必要になるかを見積もることも重要である。実証実験においては、同じ参加者に複数の条件を検証させる「実験参加者内設計」とするか、あるいは参加者をグループ分けしてそれぞれが異なる条件を検証する「実験参加者間設計」とするかを決定する。後者の場合、参加者のグループ分けの方法について、無作為に行うか、年齢や性別などを考慮し層別化するかといった点についても検討する必要がある。さらに、実験を進める中でデータをどのタイミングで測定・収集するか、どのような参加者情報を収集するのかなどについても事前に決定する。これらの検討事項について、ユーザブルセキュリティ研究では、ある条件を先に体験したことで参加者が慣れたり学習したりしてしまい、その後の条件の結果に影響が出る「学習効果」や、先に体験した条件の印象や経験が後の条件にも影響を及ぼす「持ち越し効果」を避けるため、「実験参加者間設計」を採用する研究が多い。結果の評価指標としては、タスク完遂率（何人がタスクを完了したか）や、エラー率（どれだけミスが起きたか）、セキュリティ警告の理解度などを用いる場合が多く、これらに加えて、補助的にアンケートによる自分のスキルや理解度に対する自信の評価や、簡単な知識テストで測る理解度の確認などが用いられることが多い。

なお、実証実験では特定の利用者層だけでユーザブルセキュリティの効果が強まる場合や、異なる集団で効果に違いがみられる場合がある。例えば、高齢者層が若年層に比べてパスワードの記憶に困難を感じやすい傾向がある場合には、年齢層とパスワードの記憶のしやすさに交互作用[10]があると考えられる。このような状況を、回帰分析を用いてモデル化することで、年齢層による影響の有無やその大きさの違いを明らかにし、ユーザブルセキュリティの効果が特定の集団に集中しているのか、それとも集団間で異質性がみられるのかについて検証す

9 データの順序に加え、数値同士の差にも意味がある尺度で、距離感を測ることができる。

10 ある要因（例：警告のデザイン）が結果に与える影響が、別の要因（例：利用者の年齢や知識量）によって変化すること。

る。ユーザブルセキュリティ研究では、年齢、性別、技術的スキルなどの実験参加者属性がもたらす交互作用を研究した報告が増えており、特定の利用者層で効果が強まるかを明確に示すことが重視されている。

また、ユーザブルセキュリティ研究では実験参加者数が数十名規模にとどまる事例も多い。標本サイズが小さく、データの背後に正規分布などを仮定することが困難な場合には、再標本化（リサンプリング）に基づく手法が利用されている。これは、観測されたデータをコンピュータ上で繰り返し抽出し直し、擬似的な分布を生成することで統計的推論を行うものである（具体的な手法については補論の（３）を参照）。再標本化は実務的価値が高いが、ユーザブルセキュリティ研究における活用事例は限定的であり、今後活用が進むことが期待される。

（３）量的研究に採用されるアンケート

量的研究において、利用者や開発者の考えを反映したデータを新たに収集したい場合、アンケートを実施する場合がある。アンケートの策定・選定に当たっては、測定したい対象を明確にし、それに基づいた内容となっているか（構成概念妥当性）、そのアンケートが一貫して正確な結果を出せるか（信頼性）、アンケートの対象者の属性や言語が違っていても同様の結果が出るか（測定不変性）といった点が満たされることが重視されている。仮に、既存の英語版のアンケートを流用する場合には、まず原文を日本語に翻訳し、その日本語を再び元の言語に戻して（逆翻訳）内容に違いがないかを確認したうえで、実験参加者による事前の試行テスト（パイロット測定）を実施し、その品質を検証・確認する。必要に応じて、確認的因子分析（Confirmatory Factor Analysis：CFA）[11]という手法を用いて、アンケートの内容が測定したい項目を正確にカバーしているかを点検する。以下では、ユーザブルセキュリティ研究で利用される主なアンケートとして、セキュリティ技術の使いやすさや受容度等を測定するための利用者へのアンケートや、安全性の高い製品を開発する能力に対する自信に関する開発者へのアンケートを紹介する。

イ. 使いやすさに関するアンケート

使いやすさを主観的に評価するために広く使われている、10 個の質問項目から構成されるアンケートとして、SUS（System Usability Scale）がある（別表 1 を参照、Brooke 1996）。各質問項目に対して、1 から 5 までの回答の選択肢があり、その数値を一定のルールに基づき採点して、使いやすさを数値化して評価する。

[11] CFA は、収集したデータが、理論的に想定される変数をどの程度反映しているかを検証する分析手法。具体的には、理論的に結果に影響を与えると予測される変数とその影響の大きさについて、実際のデータがどの程度一致しているかを統計的に評価する。

具体的には、奇数番目の質問項目は回答値から 1 を引き、偶数番目の質問項目は 5 から回答値を引き、得られた数値の合計値に 2.5 を掛けることで、0 から 100 のスコアに変換する。スコアが高いほど、利用者が使いやすいと評価していることを表しており、経験則として、68 前後が平均的な水準、80 以上は高評価として扱われることが多い。

評価に当たっては、SUS のスコアと併せて、タスク完遂率（評価対象の製品がタスクをどの程度やり遂げたか）やタスク完了に要した時間、タスクのエラー率といった他の客観指標を用いることが一般的である。

ロ. 利用者による受け容れの程度に関するアンケート

ある製品やサービスがどの程度利用者に受け容れられているかを測るためのアンケートとして、9 つの質問項目からなる Van der Laan Acceptance Scale がある（別表 2 を参照、Van der Laan, Heino, and Waard 1997）。「役立つか役立たないか」、「快適か不快か」など、対になる言葉の中から自分の感じ方に近いほうを選ぶ形式で回答する。SUS と併用することで、製品やサービスの使いやすさと、利用者がその製品やサービスをどの程度受け容れる意向があるか（受容）との関係を調べるのに役立つ。ユーザブルセキュリティ研究では新しい認証方式や安全性を高めたセキュアな UI を評価する際などに、利用者から受け容れられるかを評価するのに用いられる。

ハ. セキュリティ行動の実行意図や意識に関するアンケート

情報システムの利用者が、推奨されるセキュリティ行動をどの程度実行しようと考えているか、すなわち行動意図を測定するために用いられるアンケートとして、16 個の質問項目からなる SeBIS（Security Behavior Intentions Scale）がある（別表 3 を参照、Egelman and Peer 2015）。質問項目は、①パスワード生成、②デバイス保護、③更新、④プロアクティブ意識に関する 4 つの観点を測定するものとなっている。具体的には、①では安全なパスワードを作成する習慣に関する意識、②では自分のスマートフォンやパソコンなどを守るための意識、③ではソフトウェアやアプリケーションを適切に更新することへの意識、④ではセキュリティ上のリスクを未然に防ぐための姿勢や意識をそれぞれ測定する。本アンケートは、ユーザブルセキュリティ研究においても、セキュリティ教育やプレバンキングを行う前後で、利用者の行動や意識にどのような変化があったかを検証するうえで活用されている。また、パスワードの変更履歴やデバイスの利用記録などのログデータと関連付けることで、行動の傾向を考察するのにも広く用いられる。

SeBIS の質問項目には「逆転項目」と呼ばれる、意図的に回答の方向性が逆に

なるように設計した項目が含まれる。そのため、回答の集計や分析時に処理を間違えないよう注意する必要がある。また、測定結果については、全体のスコアだけでなく、上記の 4 側面のそれぞれに関するスコアも報告・活用することが推奨される。

利用者のセキュリティに対する考え方や感じ方を、より少ない質問項目で測定できるアンケートとして、ユーザブルセキュリティ研究の中で新たに提案された SA-6（A Self-Report Measure of End-User Security Attitudes）がある（別表 4 を参照、Faklaris, Dabbish, and Hong 2019）。質問項目が 6 つと少ないことから調査対象者の負担が少なく、セキュリティ教育やセキュリティ機構の改善といった取り組みの着手前後における、利用者のセキュリティに対する意識や認識、自身の安全を守れると思う程度の変化を効率よく測定し、取り組みの改善に役立てることができる。また、SeBIS などの他のアンケートと併用することで、利用者のセキュリティに対する考え方や感じ方が、実際の行動や態度にどのような影響を与えているかを探る手掛かりを得ることもできる。

さらに、利用者がインターネットやデジタルデバイスを使うときに、どのようなセキュリティ対策をとっているか、またどの程度セキュリティ意識を持っているかを幅広く測定するためのアンケートとして、ESBS（Extended Security Behavior Scale）が知られている（別表 5 を参照、Sawaya *et al.* 2024）。ESBS については、オンライン調査を通じた検証が繰り返し行われており、その信頼性が示されている。具体的には、人々の回答結果と実際の行動が関連していることが分かっており、例えば、ブラウザで利用しているウェブサイトの種類や、セキュリティ意識について測定された結果が、ESBS の結果から予測される内容と一致することが報告されている。

ニ. 開発者の自己効力感に関するアンケート

情報システムの開発者が、安全性の高いソフトウェアを開発し実装する能力についてどの程度の自信（自己効力感）を持つかを測定するための尺度としては、SSD-SES（Secure Software Development Self-Efficacy Scale）が用いられる（別表 6 を参照、Votipka, Abrokwa, and Mazurek 2020）。本アンケートは、安全性の高いソフトウェアの開発を進めるためのステップや作業の枠組みに基づいて、個々のタスクを抽出し、それぞれのタスクに対する自己効力感を多角的に調査することにより、正確な評価が可能となっている。SSD-SES は、社内教育の効果について、「安全性を意識したコードを書けるようになったか」といった観点から評価したり、ソフトウェア開発におけるレビュー体制の効果について、「セキュリティ面でのレビュー体制を導入した後、開発者がどのように成長したか」といった点について定量的に評価したりするうえで有用である。

ホ. 社会的望ましさのバイアス軽減を企図したアンケート

一般的に、本当の考えや行動を隠して、社会的に好ましい答えを書こうとする傾向を社会的望ましさ（Social Desirability）のバイアスという。セキュリティの分野でも、例えば、本当はパスワードを複数のサイトで使い回しているものの、セキュリティ意識が高いように見せかけるために、使い分けていると回答することなどが、このバイアスに相当する。こうしたセキュリティやプライバシーにおける社会的望ましさに起因するバイアスに特化して、その影響を評価するためのアンケートに、SP-SDS（Security & Privacy Social Desirability Scale）がある（別表 7 を参照、Abels, Smith, and Ortloff 2025）。自己申告データのバイアスを補正するために、その他の調査結果とともに、共変量として SP-SDS のスコアを分析モデルに投入して分析することが考えられる。あるいは、SP-SDS のスコアに基づいて回答者をいくつかのグループに分けて分析することも有効と考えられる。SP-SDS は近年提案された新しい尺度であるため、今後の活用が期待される。

（４）量的研究の事例

イ. 新しい認証方式にかかるユーザビリティに関する実証研究

Lyastani らは、安全な認証方式としてオンラインサービスにおける利用拡大が進む FIDO2 のうち、パスワードを用いない単要素認証（1FA）方式[12]について、利用者による受け容れやすさ（受容性）と使いやすさを、量的研究のアプローチにより検証した（Lyastani *et al.* 2020）。本研究は、実験参加者 94 名[13]をランダムに 2 つのグループに分ける参加者間設計を用いて行われた。一方のグループ（以下「グループ 1FA」）は FIDO2 ベースのパスワードレス認証を、もう一方のグループはパスワードを入力する従来手法による認証を用いて、検索サービスおよび SNS を模した 2 つの模擬 Web サイトへの登録とログインを試行させた[14]。その直後に、アンケート調査を実施した。

アンケートでは、システムの使いやすさを測るための SUS と、受容性を測る

---

12 FIDO（Fast Identity Online）アライアンスが策定した、公開鍵暗号方式による認証規格である。認証には、指紋や顔認証といった生体認証や、セキュリティキー等の物理デバイスを用いる。単要素認証（1FA）は、このうちの 1 つの要素のみを用いて認証する方式であり、利用者 ID やパスワードを用いた従来の認証方式を併用しない方式。なお、従来の FIDO では認証に専用デバイスが必要であったが、FIDO2 ではスマートフォンや PC の生体認証機能を利用できるため、必ずしも専用デバイスが必要ではなくなった。

13 実験参加者は大学キャンパスで募集された、属性別内訳は男性 41 名、女性 53 名、平均年齢は約 24.9 歳であった。

14 実験参加者は、最初に実験に関する説明動画を視聴した。その後、グループ 1FA の参加者は、FIDO2 に関する説明と注意説明を受け、専用デバイスの使い方に関する動画を視聴した。

ための Van der Laan らの受容性スケールのうちの 5 項目が採用された[15]。アンケート結果に基づき、2 つの認証方式のちがいについてt検定を用いた有意差検定による評価が行われたほか、受容性の決定要因については、誤差の不均一性を考慮したロバスト回帰を用いた分析が行われ、それぞれの決定要因について、相対的重要度分析[16]で説明率が評価された。その結果、FIDO2 の 1FA は従来のパスワードを用いた認証方式よりも使いやすく、かつ受容性が高いとの評価が得られた[17]。一方で、認証用デバイスの紛失時や盗難時の、アカウント復旧やデバイスの扱いに対する懸念や、認証用デバイスを使用可能な場面が限られること、技術への不信や誤解といった新たな懸念も抽出された。

ロ. クッキー（利用情報）に関する同意の実効性検証

企業等のウェブサイトや広告主が、利用者やその利用状況に関してデータを収集したり他の組織等に共有したりすることについては、プライバシー保護との観点でかねてより課題となってきた。インターネット広告の業界団体である Interactive Advertising Bureau Europe（IAB Europe）は、企業等によるデータ収集・共有に関する個人データの処理や利用者による同意の在り方に関するガイドラインである「透明性と同意に関するフレームワーク（Transparency & Consent Framework : TCF）」を策定している。TCF は、EU の一般データ保護規則（GDPR）[18]および e プライバシー指令（ePrivacy Directive）[19]を遵守できるように設計されている。Matte らは、利用者のデータ利用に関する同意を取得し、それを広告エコシステム全体で共有するための TCF に準拠した技術仕様と仕組みである「同

---

15 これらに加えて、補助データとして、利用者が新しい技術やシステムに対して、どの程度積極的かつ意図的に関わろうとするかという個人の性格特性（テクノロジーへの親和性）を測定するアンケートである Affinity for Technology Interaction(ATI)と、新しいテクノロジーに対する実験参加者のプライバシー懸念を測定するアンケートである Privacy Concern における質問の一部が採用された。また、情報系の知識や学習経験の有無についても質問項目に含めた。

16 重回帰分析等において、各独立変数が従属変数の分散を説明する割合を算出する手法。

17 SUS はグループ 1FA の平均 81.74 が従来のパスワードを用いたグループ（グループ Pass）の 71.77 を有意に上回り（$p < 0.001, d = 0.85$）、どちらも「Good」水準だが 1FA がより高評価であることが示された。受容性はグループ 1FA 4.29、グループ Pass3.41 で有意差あり（$p < 0.001, d = 1.35$）となった。ロバスト標準誤差を用いた回帰分析では、SUS（標準化係数 b≈.02）と群ダミー（1FA, b≈.76）が受容性に有意であること、情報系バックグラウンドはわずかに負の効果（より批判的）だが説明率は小さいことが示された。

18 2018 年 5 月 25 日に施行された個人データの保護と利用に関する EU の包括的な法令であり、EU 域内の個人データの処理・移転に携わる世界中の組織に適用される。

19 EU が電子通信分野における個人のプライバシー保護を目的として定めた指令で、特にウェブサイトの Cookie（クッキー）利用に関する同意取得や、迷惑メール、通信の秘密などを厳しく規制し、GDPR を補完・具体化する役割を持つ。Cookie 法とも呼ばれる。

意管理プラットフォーム（CMP）」について、GDPR や e プライバシー指令といった現行の法規制に則って正しく運用されているかについて、法学的観点から評価するとともに、定量研究の手法を用いて検討・評価した（Matte, Bielova, and Santos 2020）。

具体的には、GDPR と e プライバシー指令の要件に基づき、①利用者が選択を行う前に、勝手に「同意した」として記録していないか（選択前同意の保存の有無）、②データ収集を「拒否する」方法が明確に提供されていないということはないか（拒否手段の不提供の有無）、③選択肢が最初から「同意あり」に設定されていないか（事前確認の有無）、④実際に「拒否」したのに、その選択が反映されずに「同意」として記録されていないか（選択の不遵守）の 4 点について検証を行った。本研究では、欧州中心のウェブ 28,257 サイトを調査したところ、TCF に準拠していることを示すバナーがある 1,426 サイトを自動検出した。このうち、バナー内容にしたがった同意や拒否の選択がある 560 サイトを精査したところ、少なくとも 54%に相当する 304 のサイトで、上記いずれかの問題が発生していることが確認された[20]。また、利用者がデータ収集に「拒否」を選んでも、一部のサイトでは実際にデータが収集・送信される「同意のエスカレーション」や、拒否後にもトラッキングリクエスト（監視用のデータ）が増加する現象も確認された。

検証に当たり、著者らは①標準的な API、②ウェブサイト間で共有される、利用者の設定情報を保存する仕組みである共有クッキー、③ウェブアドレスに埋め込まれる情報である URL パラメータ、④特定のページから別のページへと遷移される仕組みであるリダイレクト、など TCF が規定する同意共有メカニズムを網羅的に観察し、各ウェブページで記録された同意情報や、第三者への送信状況を分析した。分析には、Cookinspect と呼ばれる計測基盤を用いた[21]。まず、ウェブサイトをロードし、TCF が規定する「同意情報を記録するための共通データ形式」と、それを取得・共有するための「技術的なインターフェース」の仕様に基づき、各サイトの同意情報を取得、記録した。そのうえで、取得したデータの中にサイトや広告業者の識別情報、利用者が許可した処理の種類や範囲が含まれているかを調べた。

---

[20] 選択前同意の保存を行うサイトは TCF 準拠サイト全体のうち 141 サイトであった。また、精査を行った 560 サイトのうち、事前確認を行うサイトが 236 サイト、選択の不遵守に相当するサイトが 27 サイトであった。

[21] Web サイトにおける Cookie の設置状況やプライバシー保護の遵守状況を自動調査するための動的解析プラットフォーム。ブラウザ自動操作ツールを介して実際の Web サイトをクロールし、Cookie の有効期限、属性（HttpOnly, Secure 等）、およびサードパーティへのデータ送信挙動を大規模かつ継続的に計測・可視化することを可能とする。

判定に当たっては、5 つあるデータ利用目的に関する項目のうち、1 項目でも利用者が操作する前の時点で「同意」したものとして扱われていた場合には、選択前同意に該当し、法的な違反の疑いがあるものと判定した。拒否手段の不提供と事前確認は、実際のサイトを目視で確認することにより判定した。選択の不遵守については、全ての項目を拒否したにもかかわらず、その直後に、すべての項目について同意したとしてデータが保存されている場合に該当すると判定した。

本研究の特徴は、人間の行動の結果として生成される成果物（Web サイト）を大量に収集し分析することで、制度や設計原理が社会にどの程度浸透しているかを測定している点にある。ユーザブルセキュリティ研究では、実際に利用者を勧誘してタスク実施やアンケート回答をしてもらうケースが多い一方で、このように実環境を大規模に観測する量的研究も重要な位置を占めている。

４．質的研究によるアプローチ

質的研究は、人々の行動や意思決定の背後にある動機や考え方を、実際の生活や仕事の様子の観察（参与観察）やインタビュー、文書や行動の記録（ログ）の解釈などを通じて理解し、説明しようとする研究手法である。ユーザブルセキュリティ研究では、セキュリティやプライバシーに関する警告や権限要求をどのように受け止めて対応しているのか、セキュリティ機構に関してどのような誤解があって、誤った行動をとってしまうのかといった、心理的な背景を探ることが主な役割となる。また、これらの心理面に、職場や家庭といった特定の環境におけるルールがどのように影響しているのか、ダークパターンをどのように認識し判断しているかなど、量的研究だけでは捉えにくい構造についても明らかにする。本節では、質的研究を行ううえで必要となる情報の代表的な取得法と分析法を概説し、結果の妥当性や信頼性を確保するうえで必要となるアプローチを紹介する。

（１）質的データの取得方法

イ．半構造化インタビュー

半構造化インタビューは、実験参加者の考えを深く理解するために、事前に質問の大まかな流れや内容を決めた質問ガイドを用いつつ、実験参加者の回答に応じて質問を掘り下げるなど柔軟に進めていくインタビュー手法である。半構造化インタビューを行う際には、いくつかの工夫が必要となる。例えば、望ましいと考えられる回答を選んでしまう社会的望ましさの影響を緩和するためには、実際に起きたことの確認と、その背景にどのような意図や考え方があったのかについて、それぞれを切り分けて聴取することや、過去の具体的な経験を思い出させるように誘導したうえで、それを踏まえて仮想的な状況に対する反応を尋

ねることが有効である。また、インタビューを、インターネット等を介してリモートで実施する場合には、確実にインタビューの内容を記録して後から振り返るために録音品質を高く保つことや、正確にインタビュー内容を書き起こすことが必要となる。質問ガイドは、インタビューを円滑かつ体系的に進めるために、導入、ウォームアップ、中心的な質問、深掘り、振り返りという段階的な構成とすることが望ましい。具体的には、導入部分では自己紹介や簡単な質問から始め、ウォームアップの質問で参加者が話しやすい雰囲気を作った後、インタビューの中心となる質問へと移る。その後、適宜、回答の背景に関する深堀りや感情面の確認を行い、最後にインタビュー全体の振り返りへと段階的に進むことで、参加者の経験や考えを段階的に把握できるようにする。

ユーザブルセキュリティ研究では質的研究の手法として半構造化インタビューが最も広く用いられている。半構造化インタビューによって、研究者は警告画面の理解やプライバシー設定の意思決定の過程を詳細に把握することができ、そこで得られた知見は、後続の量的研究での質問設計や測定項目の精緻化、UI 設計の改善案の試作（プロトタイピング）などに活かされることが多い。

ロ. 行動観察に基づくエスノグラフィックアプローチ

エスノグラフィックアプローチは、もとは人類学で使われる調査法で、人々の生活や文化について、実際に彼らの日常生活や行動を間近で直接観察や記録をし、その情報を活用して解釈する手法である。

ユーザブルセキュリティ研究では、インターネットを使うときやセキュリティのための設定を行う場面など、普段の作業の中で利用者がどのように行動しているのかを現場で俯瞰的に観察するほか、人々の行動を規定する社内規程のような文書や日々の実践の記録、自作操作手順書といったアーティファクト（研究プロセスで生成されたあらゆる成果物）を収集したり、調査対象者に付き添って実際の作業をその場で観察するシャドーイングを行ったりすることにより実践される。また、調査対象者の作業中にインタビューを行うことで、具体的な活動の中で何を考えながら行動しているのかを聞き取るコンテクスチュアル・インクワイアリの手法も使われている。

研究の内容によっては、データのやり取りや操作が、規制やデータ保護の観点から制約されている場合もある。こうしたケースでは、現場で直接観察する参与観察ではなく、少し距離を置いて観察する非参与観察を行いつつ、半構造化インタビューを組み合わせて、観察できない行為を後から振り返って質問する回想法や、一連の活動を全体的に振り返るアクティビティ・マッピングで補う。

観察事項を記録するフィールドノートでは、どのような出来事が起き、どのようなことが語られたかだけでなく、どのような環境や規則、システム設定のもと

で観察された事象かについて、切り分けて記録する。

ハ. 専門家による認知的ウォークスルー

認知的ウォークスルーは、製品や技術の使いやすさについて、専門家が利用者の立場にたって、特定のタスクを達成できるかについて順を追って確認することにより判定していく評価手法である。本手法は、実験参加者のデータを直接収集する必要がないため、研究の実施コストが低いほか、プライバシーへの配慮といった倫理面の検討負担も小さいが、評価する専門家自身の先入観や、設計を行った人による思い込みが解釈に影響しやすい点には留意が必要である。

例えば、セキュリティに関する UI を評価する場合には、注意と理解、決定と行動という利用段階ごとに検証をし、利用者が仮に誤った行動をしてしまったときに、その結果がどの程度重大な影響をもたらすのか、そのことをどの程度利用者が認識できるのかなど、多角的に検証することが望ましい。

ニ. 複数の専門家の匿名意見に基づく Delphi 法

Delphi 法は、複数の専門家が匿名で意見を出し合い、その集計結果に基づき再度意見を出し合うことを繰り返すことにより、最終的な合意形成を行う方法である。本手法は、匿名性により、互いの意見の影響を受けにくいように工夫されているほか、全体としてどのような意見が多かったかをまとめた資料を基に、次のラウンドで再評価するプロセスを何度か繰り返すことで、最終的により多くの人が納得できる結論にたどり着けるように設計されている。本手法は、例えば、セキュリティ技術の開発において、その機能に関する設計要件や評価項目を検討・策定するのに向いている。

本手法を用いる場合、参加する専門家は、多様なバックグラウンドの人材から偏りなく選択することが重要である。また、最終的な合意形成時の意見の中央値や分散、最終合意に至った根拠をしっかりと開示することが望ましい。

ホ. 自由記述式アンケート

自由記述式のアンケートは、短時間で多様な事例や情報を集められる手法として有用で、ユーザブルセキュリティ研究では、大規模ウェブ調査に組み込まれる形で広く利用されている。本手法を単独で用いるのではなく、必要に応じて半構造化インタビューなど他の質的研究や、ログ分析といった量的研究と組み合わせて利用することが多い。

利用者から具体的な回答を引き出すことができるような設問文とし、社会的望ましさを緩和するために「正解はありません」「普段通りの正直な状況を教えてください」といった前置きの文を添えることが行われている。また、回答内容

の公開可否について確認する質問項目を設けておくことが推奨されている。

（２）実験参加者の募集

半構造化インタビューやエスノグラフィックアプローチといった、利用者の協力を前提とした研究では、実験参加者をどのように集めるかは、研究の質に影響を与える重要な要素の一つである。

一般の利用者を幅広く募集する場合には、インターネットを通じて参加者を募るクラウドソーシングが広く活用される。本手法は、短期間で多様なバックグラウンドをもつ多数の参加者を効率的に集めることができる一方で、一人一人の詳細な状況や考え方を深く調べるような調査には向かないことがある。

そこで、クラウドソーシングにより参加意向を示した候補者から、2 段階のスクリーニングにより必要な条件を満たしている人を選ぶ場合が多い。第 1 段階では、調査対象となる特定の行動の実践有無を確認して候補者を絞り込む。つづいて、第 2 段階では、回答の注意深さや回答に要した時間のほか、同一の IP アドレスやデバイスから複数の回答がよせられるなどしたものを、異なる参加者の回答として重複カウントしていないかといった点から、回答の質を確保するための絞り込みを行い、最終的な実験参加者を確定する。さらに、研究の一環として特に詳しく話を聞きたい場合は、スクリーニングを通過した人たちの中からインタビューに協力してもらう参加者を募集することがある。

実験参加者を集めるもう一つの手法として、スノーボール・サンプリングが挙げられる。本手法は、まず一人の実験参加者を決定し、その参加者に次の参加者を紹介してもらうことを繰り返して対象者を広げていく非確率的なサンプリング手法である。この手法は、クラウドソーシングでは集めにくい高齢者や、特定の技術を扱う専門家、特定の職業に従事している人など、日常的に接触する機会が少ない属性の人たちを他の手法に比べて集めやすく、ユーザブルセキュリティ研究では、対象を絞った特定の母集団に向けた調査において採用される傾向がある。紹介制であるため、同じような特性や考えをもつ人が集まりやすく、偏りが出やすい点には留意が必要である。これに対しては、紹介の連鎖がなるべく多様性を持つように管理したり、既に選ばれた人々とは違う特徴をもった人を意図的に追加で探したりする対策が取られることがある。例えば、企業内での研究では、直属の同僚同士を紹介することを避け、利害関係の薄い部門からも協力者を募集するなどの工夫が施されることがある。

（３）質的データの分析手法

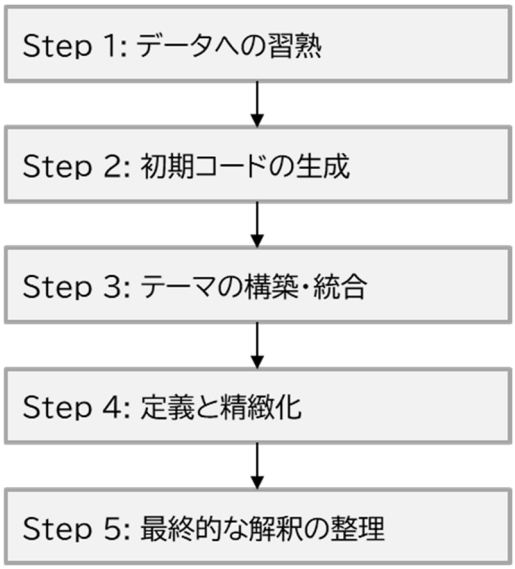
以下では、（１）で示した手法により収集した質的データの分析手法について整理する。

イ. テーマティックアナリシス

インタビューやアンケートなどで集めた大量の文字データの分析には、テーマティックアナリシスとよばれる手法が用いられる（図 4 を参照、Braun and Clarke 2006）。分析は以下の流れで行われる。まず、集めたインタビューやアンケートの結果を通読したうえで、それぞれの意見や記述に対して、何を意味しているのかを示す初期コードと呼ばれるフレーズを付す[22]。つづいて、類似した概念の初期コードをテーマごとに統合し、各テーマの定義やテーマ間の境界を精緻化する。最後に、それぞれのテーマの連関性を踏まえて、全体として一貫性のある解釈へと整理していく。ユーザブルセキュリティ研究では、本手法を用いて、行動、認知、組織、制度などの観点からテーマを整理することで、システムの機能や UI の設計に関連する具体的な仕様や、それを実現するための設計要件を導き出すプロセスに活用される[23]。

テーマティックアナリシスには、複数のアプローチがあり、代表的な手法としてコードブック方式とリフレクシブ方式が挙げられる。コードブック方式は、事前に分析に用いるコードとその定義を決めておき、それを複数の人が使用することで、互いの結果について信頼性（相互信頼性、Inter-Rater Reliability：IRR）を確保しながら分析を進めることができる。

図 4 テーマティックアナリシスの概要



22 例えば、「アプリケーションの使いやすさが重要」という意見に対しては、「使いやすさ」や「ユーザビリティ」という初期コードを付す。

23 データから概念を抽出する質的研究手法としては、グラウンデッド・セオリー・アプローチ（Grounded Theory Approach：GTA）が広く知られている。GTA が新たな理論の構築を最終目的とするのに対し、テーマティックアナリシスは、理論構築に縛られず、特定の事象（利用者の行動や心理など）における共通パターンを柔軟に抽出・整理することに特化した手法である。

一方、リフレクシブ方式は、研究者が分析を進めながら自由にコードを生成し、適宜調整しながら使用するため、特定のルールに縛られず、柔軟にデータを解釈することができる。本方式において、結果の妥当性を確保するために、研究者同士が互いの研究計画や結果について深く話し合い合意形成に努めるピア・デブリーフィングや、詳細な研究プロセスの記録（監査可能性の担保）、主要解釈を実験参加者に提示してフィードバックを得るメンバーチェック、理論的飽和[24]の説明、研究結果に反するデータや事例の提示による公平性担保、研究結果に影響を与えうる研究者自身の特性や経歴等の開示などを行っている。

なお、テーマティックアナリシスの分析結果を量的研究の分析に採用する場合には、意見や記述に付すコードについて、それぞれの出現頻度や共起関係を数量化して評価する。その際、単語単位の頻度だけに偏らないよう、意味単位の粒度を明確にするように留意するとともに、それぞれのコードに該当する代表的な意見や記述について、それぞれ賛否の両面を踏まえて評価する。

また、テーマティックアナリシスの研究結果の信頼性を確保するうえでは、指標[25]などを用いて、コードの一貫性の妥当性を確保していくことも行われている。

ロ. アフィニティダイアグラミング

多くの意見を整理し、要点を抽出する際に有効な手法として、アフィニティダイアグラミングが挙げられる。本手法では、それぞれの意見や発言について、短い言葉やフレーズに纏めた 1 枚のカードを作成する。つづいて、類似した意見や考えが記載されたカードごとにグルーピングする。ユーザブルセキュリティ研究では、セキュリティ機構の設計要件や設計原則の抽出に利用されている。例えば、多要素認証の導入に関する利用者調査の結果を整理する場合、「SMS のコード入力が面倒」「スマホを常に持ち歩いていない」「認証のタイミングが予測できない」といった個別の不満をグルーピングし、それに対して「利用者が認証のタイミングを事前に予期できる通知を行う」や「業務フローを中断させない認証デバイスを選択する」といったグループ名を付与する。

その際、グループ名は名詞にとどめず、行動や条件を含む短い文で表現すると設計に直結しやすい。また、多様な提案を促し、本手法の有用性を高めるために、

---

24 研究対象のデータをさらに増やしても、新たな知見が導かれたり異なる結果が得られたりすることがない状態。それ以上のデータ収集は必要ないと判断し、研究の一区切りとみなす基準となる。

25 複数の分析者による判断の一致度を測る指標には、2 人の分析者によるカテゴリ分類に用いる「Cohen のカッパ係数」がある。これは、2 人の分析者が同じ対象を分類した際の一致度を評価する指標であり、単純一致率ではなく、偶然一致する可能性を考慮して補正した一致度を示す。3 人以上に拡張した指標として「Fleiss のカッパ係数」、順序尺度におけるカテゴリ間のズレの大きさを考慮する「重み付きカッパ係数」などがある。

複数のメンバーがワークショップ形式で取り組むこともある。

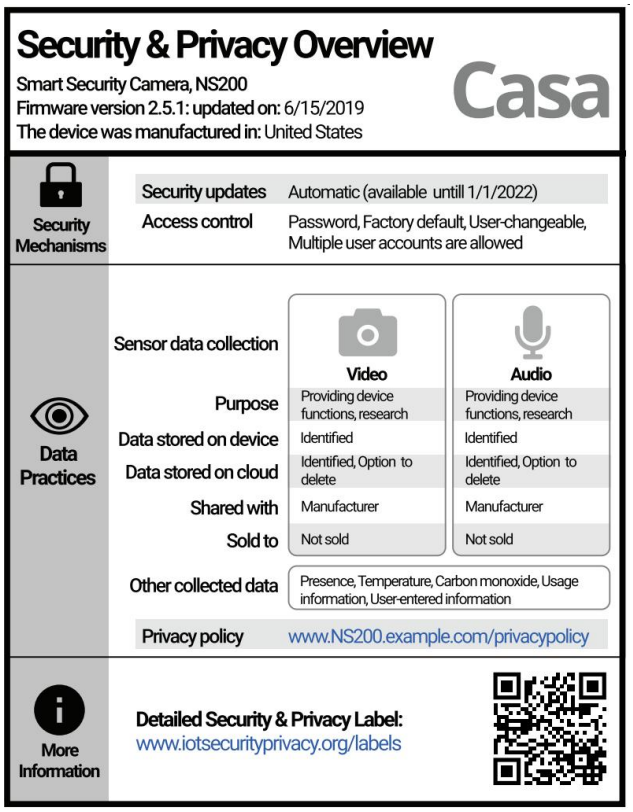
（４）質的研究の事例

イ．IoT 機器向けプライバシーラベルの検討

Emami-Naeini らは、消費者が安心して IoT 機器を購入できるよう、機器に付すプライバシーラベルを提案し、それに記載すべき具体的項目について、質的研究を用いて研究した（図 5 を参照、Emami-Naeini *et al.* 2020）。

まず、22 名のプライバシー分野の専門家[26]に対してラベル候補項目の採否とその理由について半構造化インタビューを行い、Delphi 法により合意形成された 47 の項目をリストアップした。リストアップされたそれぞれの項目について、その重要性に関する評価を行い、「消費者が商品パッケージを見た時にすぐにわかるよう、箱や製品ページに表示すべき情報（一次層）」と、「もっと詳しく知りたいときに QR コードや URL リンクから調べることができる情報（二次層）」

図 5 イ．IoT 機器向けプライバシーラベルの例



（資料）Emami-Naeini *et al.* 2020 の Fig. 3

[26] 参加者は米国の産官学 NGO から選抜され、少なくとも「プライバシー／セキュリティ分野 10 年以上の実務／研究」「著名書籍の著者」「標準化活動の実績」「IoT 製品チームのリーダー」「大学教員」のいずれかを満たす。

の 2 つに分類した。分析はテーマティックアナリシスにより行い、理論的飽和は 20 名時点で確認した。

つづいて、IoT 機器の購入経験のある消費者 15 名[27]に対する半構造化インタビューを実施した。ラベルの理解しやすさや実際の理解度、意思決定への影響を確認したほか、二層設計の賛否と利点欠点に関する意見や、購買時の行動イメージについて掘り下げた。主な評価としては、例えば、一次層の情報は全参加者に概ね理解され、特にセキュリティパッチの更新提供期限、IoT 機器が収集するデータの種類や収集の目的、収集したデータの共有先や販売先、収集したデータの保存場所が、購買判断に強く影響することが示された。一方、二次層に掲載される専門的な項目に対する理解度にはばらつきがあるものの、「詳細は載っていてほしい。」「必要なら自分で検索する。」「専門家が監視できる状態が望ましい。」という意見が主であった。

その結果を踏まえ、一次層には上記をはじめとする 12 項目を中心に配置し、二次層には収集したデータの保持期間や規格準拠や法令適合の状況など、セキュリティに関する詳細事項を配置する案が提案された。最終的に提案されたラベル項目の妥当性を検証したところ、非常に高い妥当性を有することが確認された[28]。

本研究は、専門家の意見を具体的なラベル項目へと落とし込んだうえで、実装を念頭においた二層構造を提案するなど、実用性の高い研究と評価できる。さらに、消費者インタビューを通じて設計に消費者の意見を反映させた点にも価値があるといえる。例えば、専門家は「データ共有頻度」を二次層に掲載することを提案したが、その後の消費者へのインタビューにより「収集したデータが誰に共有・販売されるか」への関心が非常に高いことが明らかとなったことを踏まえ、最終的な案では、より消費者が認知しやすい一次層に掲載することが提案されている。また、二次層の項目については各企業が必要に応じて情報を開示する仕組みとなっており、これによってメーカー間の透明性向上に向けた競争を促進する効果が期待される。

将来的には、消費者向けの情報開示にとどまらず、規制当局等によるセキュリティ検証といった第三者評価や、認証制度における評価項目としても活用されることが期待される。

ロ. 高齢者によるパスワードマネージャーの利用に関する研究

ウェブサイト上で用いるパスワードの保存・記憶や、安全で固有のパスワード

---

27 参加者は大学のオンライン掲示板上の募集サイトで募集した。応募者から、18 歳以上かつスマートホームまたはウェアラブルの購入経験者を条件にスクリーニングを行った。

28 具体的には、Cohen のカッパ係数により評価した。

の自動生成を可能とするツールは、パスワードマネージャーと呼ばれる。Ray らは、高齢者によるパスワードマネージャーの利用に関して、半構造化インタビューを用いた質的研究を行った（Ray, Wolf, and Kuber 2021）。調査票とコードブックは、先行研究（Pearman *et al.* 2019）と同一のものを使用しつつ、対象者を高齢者に限定して行われた。実験は、対面募集とスノーボール・サンプリングを併用して集めた米国の高齢者 26 名（平均年齢 70.4 歳）を対象として行った。これらの参加者は、パスワードマネージャーを利用していないグループ（10 名）と、ブラウザまたは OS に内蔵されたパスワードマネージャーを利用しているグループ（9 名）、パスワードマネージャー専用機器を使用しているグループ（7 名）に分類して比較した。

検証の結果、高齢者はパスワードのクラウド保存とデバイス間同期に対する不信感が強い一方で、家族からのアドバイスやパスワード管理に関するレクチャーを受けた場合には、パスワードの自動生成機能を活用するなど、相対的により安全な行動を選択する傾向を有することが確認された。また、自分の情報が悪用されるなどの脅威を認識している点や、一つのシステムに頼りすぎることが危険であるという懸念、自分の情報を自らの意思で統制することができることを重視する点は若年層と共通である一方で、新たな技術を採用するまでに、高齢者は若年層よりも多くの動機付けを要することや、ひとたび採用した技術については、若年層よりも継続的に使い続けることが、既存研究との比較により明らかとなった。このように、先行研究と同じ方法で対象者のみを変えた調査を行うことにより、既存研究の再現性を実証するとともに、結果を比較することにより高齢者特有の示唆を得られた点が有益である。

５．研究の透明性

（１）なぜ透明性が必要か

ここまで、量的研究および質的研究の方法論と具体的な事例を整理してきた。しかし、これらの研究成果を適切に評価し、知見として蓄積していくためには、研究の透明性を確保することが不可欠である。研究の透明性とは、研究の設計、実施内容、分析手法および結果について、第三者が内容を追跡・再評価できるだけの情報を体系的かつ再利用可能な形で提示することを指す。研究プロセスで生成された日々の記録やデータ、コード、分析手順、環境設定等の成果物といったアーティファクトを開示することによって、第三者は再分析や再現実験を通じて研究の妥当性を検証できるだけでなく、異なるデータを用いた実用性の検証研究（複製研究）にも繋がる。

このような透明性確保の動きは、コンピュータサイエンス分野で先行して見られた。当分野では、アーティファクトに対して入手可能性（Available）と機能

性（Functional）、再現性（Reproduced）などを段階的に審査してバッジを与える仕組みが普及しており、研究成果の再現性を高めるための標準化が進んできた（ACM 2020）。サイバーセキュリティ分野の主要な国際会議の一つであるUSENIX Securityにおいても、採択論文の著者に対しアーティファクトの共有を原則とし、第三者の利用可能性についての検証を必須化している。さらに希望者に対しては、機能性や再現性の詳細審査を実施する制度も開始された（USENIX 2024）。

もっとも、ユーザブルセキュリティ研究においては、研究対象の多様性や実験設計の複雑さから、これまで研究過程や分析手法の詳細な開示が必ずしも十分には意識されてこなかった。近年、国際会議等を中心にユーザブルセキュリティ研究の透明性確保に向けた取り組みが進展し、コミュニティ規範として整備されつつある。このような動向を踏まえ、本節では、その具体的な取り組みと課題について整理する。

（２）透明性向上に向けた取り組み

ユーザブルセキュリティ研究の分野で重要な国際会議の一つである SOUPS（Symposium on Usable Privacy and Security）は、研究成果を誰でも無償で自由に閲覧できるオープンアクセスの推進を早くから掲げ、会議で発表された論文集（プロシーディング）や発表資料の無償公開を行ってきた。加えて近年は、研究の透明性ガイドラインを共同で策定するためのワークショップが企画されるなど、研究情報の提供と第三者による再現性の確保をコミュニティとしての明確な目標に据えている（SOUPS 2025）。具体的には、①研究の目的と仮説、②データの前処理手順、③集計対象から除外するデータの判断基準、④主要指標と副次指標の区別、⑤サンプルサイズの根拠、⑥相関関係の分析や回帰分析に用いた設定条件、そして⑦実験を再現するために必要となるプログラムやデータの取得方法等について、第三者が再利用できる粒度での提示を求める方向で検討が進められている。

同様の動きとしては、研究コミュニティに対する大規模な調査に基づき、透明性確保のための包括的な推奨事項を提示した Klemmer らの研究などもある（Klemmer *et al.* 2025）。これらは、トップカンファレンスにおける標準的な報告様式として制度化することが視野に入れられている。

（３）再現性向上に向けた取り組み

複製研究は、既存研究の知見を、独立したデータまたは設定で再検証するものである。先行研究の結果が単なる偶然（偽陽性）や特定のデータセットに依存したものでないか（堅牢性）を担保する「直接複製」と、異なる対象者や環境にお

いても同様の理論が成立するか（汎用性）の担保と適用限界を明らかにする「概念複製」があり、その双方の重要性が高まっている。

従来、複製研究は新規性の高い研究に比べて学術的価値が劣後するとの認識が一般的であったが、SOUPS は比較的早い段階から複製論文を採択しており、知識の累積を重視する姿勢を表明してきた。2017 年の Call for Papers においては、複製研究の著者に対して、複製研究の意義、複製の類型、方法差分の明示、貢献の要約を求め、複製研究を受け容れることを明示的に示している。

しかし、ユーザブルセキュリティ研究をはじめとする人間とコンピュータの相互作用（Human-Computer Interaction : HCI）を扱う研究領域では、依然として、研究手法に関する重要な情報[29]やアンケート内容が非公開とされる研究が、他の研究領域と比較して多いことが指摘されている。複製研究をより推進していくためには、透明性のガイドラインとアーティファクト公開の制度化によって、研究基盤を整備していくことが重要である（Suray *et al.* 2024）。

2024 年に開催された USENIX Security において、Hasegawa らは、ユーザブルセキュリティ研究の分野では他の HCI 分野の研究と比較して実験参加者の地理的偏りが大きく、例えば年齢や性別、人種、教育の程度、社会的な背景など、参加者の特徴や属性に関する情報（デモグラフィック）が報告されないことにより複製可能性が損なわれていることを指摘した（WEIRD 問題[30]、Hasegawa, Inoue, and Akiyama 2024）。また、研究方法や参加者の募集方法は、地理的・言語的な障壁の影響を受けやすく、特定の地域の参加者に依存した研究設計に繋がっており、特に、非西洋圏に固有の問題を取りこぼす可能性が高い。今後は、デモグラフィックの詳細な情報を求めるとともに、非西洋圏のデータを用いた複製研究を促進することが、知見の一般化において重要な点となることを強調した。

６．ユーザブルセキュリティ研究の今後の展望

今後のユーザブルセキュリティ研究は、研究活動そのものの信頼性を支える基盤的課題と、新たに顕在化しつつある認知的脅威への対応、そして従来から議論されてきた社会実装上の課題の深化という、複数の層が重なり合う形で展開されていくと考えられる。

まず基盤として重要となるのが、研究プロセスの透明性の制度化である。前節で述べたように、ユーザブルセキュリティ研究は人間を対象とする実証研究に

---

29 具体的には、ある介入が持つ実際の影響の大きさを数値で示す効果量や、その効果量がどの範囲に収まっているのかを統計的に示す信頼区間など。

30 欧米を中心とした「西洋の（Western）、教育水準の高い（Educated）、工業化された（Industrialized）、裕福な（Rich）、民主主義の（Democratic）」人々の心理的特徴を指す頭字語。

強く依存する分野であり、その信頼性は手法や分析過程の開示に大きく左右される。Klemmer らが示したガイドラインに基づき、データの前処理手順、分析コード、事前登録といった要素の公開が標準的な実践として定着していくと考えられる。

次に、生成 AI の普及に伴い、情報操作や誘導に対する人間側の脆弱性が、ユーザブルセキュリティ研究における新たな中心課題として浮上しているように、生成 AI がもたらす新たな認知的な脅威をより深く議論することになるだろう。ここでは、誤情報を技術的に検知・除去するアプローチだけでなく、人間の認知や判断プロセスそのものをどのように支援あるいは保護するかが問われると考えられる。すでに、誤情報の典型的な操作手法を事前に理解させることで耐性を高める心理的接種の考え方は、個人レベルにとどまらない社会的対策として注目されている（Roozenbeek *et al.* 2022）。また、生成 AI に対する過度な信頼やリスクの過小評価といった問題に対しては、利用者の特性や利用文脈を踏まえた UI 設計を通じて、利用者が生成 AI の限界やリスクを正しく理解し、それに基づいて適切に活用できるような枠組みを構築することが重要であると指摘されている（Tolsdorf *et al.* 2025）。

一方で、パスキーに代表される新たな認証技術の普及は、ユーザブルセキュリティ研究が長年向き合ってきた社会実装上の課題を、改めて顕在化させている。技術的に安全であっても、組織への導入や長期運用が困難であれば、その効果は限定的となる。リカバリー設計やレガシーシステムとの共存といった運用面の課題は、認証技術が進歩した現在においてもなお解消されておらず、今後も重要な研究対象であり続けるだろう（Lassak *et al.* 2024）。加えて近年は、利便性や支援を目的として設計された機能が、対人関係の中で悪用される可能性にも注目が集まっている。例えばパスキーのような認証情報やデバイスが、家族やパートナーなどの本人と密接な関係のある人に本人が望まないかたちで共有・委譲させられることによって、過度な監視や支配に繋がるケースがありうることが指摘されている。こうした観点は、主として外部の攻撃者を想定してきた従来のセキュリティ設計においては、考慮されてこなかった。そのため、認証情報やデバイスの設計段階から、それらの共有機能が悪用される可能性を念頭において評価・検討する枠組みを確立することは、親密な人をも脅威として捉え直す新たな視点であり、人間中心のセキュリティ設計を次の段階へ進めるための重要な課題となると考えられる（Daffalla *et al.* 2025）。

7．まとめ

本稿では、ユーザブルセキュリティ研究を狭義と広義の 2 つの観点から整理するとともに、主要な研究手法として量的研究と質的研究を概説し、研究の透明

性を確保することの重要性を論じた。

狭義のユーザブルセキュリティ研究では、認証や警告表示などのセキュリティ機構を対象として、利用者の誤操作を防止し、安全性と使いやすさを両立するための研究が進められてきた。一方、広義のユーザブルセキュリティ研究では、利用者の行動や意思決定に着目し、偽情報への対応やセキュリティ教育の支援など、セキュリティに関わる社会的・人的側面を含む課題を研究対象としてきた。

ユーザブルセキュリティ研究では、量的研究による利用者の傾向や要因間の関係の把握と、質的研究による行動の背景や意思決定の文脈の理解が、それぞれ重要な役割を果たしている。一方で、これらの研究手法を用いて得られる知見を蓄積し、比較・検証可能な形で発展させていくためには、量的研究における測定尺度の開発や評価方法の整備、質的研究における分析手法や研究手続きの透明化など、研究方法に関するさらなる検討が必要である。

今後のユーザブルセキュリティ研究では、単一の研究手法に依存するのではなく、量的・質的アプローチを相補的に活用することで、利用者のセキュリティに関する行動や認知をより包括的に理解し、実効性の高いセキュリティ機構や支援方法の設計につなげていくことが重要である。また、研究によって得られた知見を、多様な利用者層や文化的背景を考慮しながら、実際のシステム設計や教育・啓発活動へ適用するとともに、その実践を通じて得られた新たな課題や知見を再び研究へ還元する循環を形成することが期待される。そのためには、研究手法や分析過程を明確に示し、その透明性や複製可能性を確保する取り組みを継続することで、信頼性の高い研究成果を蓄積させていくことが重要である。こうした取り組みにより、人間中心のセキュリティ実現に向けて研究成果を社会実装へ展開可能な知見として発展させていくことが期待される。

以　上

参考文献

Google、「実行時の権限」、『Android オープンソース プロジェクト』、Google、2026 年（https://source.android.com/docs/core/permissions/runtime_perms、2026 年 8 月 17 日）

Abels, Laura Marie, Matthew Smith, and Anna-Marie Ortloff. “I Never Reuse Passwords! Development and Validation of a Security and Privacy Social Desirability Scale (SP-SDS) for End Users without a Background in Computer Science.” *Proceedings of the Twenty-First USENIX Conference on Usable Privacy and Security*, no. 23 (2025): 415-434. https://www.usenix.org/system/files/soups2025-abels.pdf.

ACM. 2020. “Artifact Review and Badging Version 1.1.” Artifact Review and Badging – Current. Accessed August 17, 2026. https://www.acm.org/publications/policies/artifact-review-and-badging-current.

Akhawe, Devdatta and Adrienne Porter Felt. “Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness.” *Proceedings of the 22nd USENIX conference on Security* (2013): 257-272. https://www.usenix.org/system/files/conference/usenixsecurity13/sec13-paper_akhawe.pdf.

Apple. “Transparency is the best policy.” Apple. Accessed August 17, 2026. https://www.apple.com/privacy/labels/.

Basol, Melisa, Jon Roozenbeek, and Sander van der Linden. “Good News about Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News.” *Journal of cognition* 3, no. 1 (2020): 2. https://doi.org/10.5334/joc.91.

Braun, Virginia and Victoria Clarke. “Using Thematic Analysis in Psychology.” *Qualitative Reesarch in Psychology* 3, no. 2 (2006): 77-101. https://doi.org/10.1191/1478088706qp063oa.

Brooke, John. “SUS-A Quick and Dirty Usability Scale.” *Usability Evaluation in Industry* (1996): 189-194.

Daffalla, Alaa, Arkaprabha Bhattacharya, Jacob Wilder, Rahul Chatterjee, Nicola Dell, Rosanna Bellini, and Thomas Ristenpart. “A Framework for Abusability Analysis: TheCase of Passkeys in Interpersonal Threat Models.” *Proceedings of the 34th USENIX Conference on Security Symposium*, no. 401 (2025): 7819-7838. https://www.usenix.org/system/files/usenixsecurity25-daffalla.pdf.

Dhamija, Rachna, J. D. Tygar, and Marti Hearst. “Why Phishing Works.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2006): 581-590. https://doi.org/10.1145/1124772.1124861.

Egelman, Serge and Eyal Peer. “Scaling the Security Wall: Developing a Security Behavior Intentions Scale (SeBIS).” *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015): 2873-2882. https://doi.org/10.1145/2702123.2702249.

Emami-Naeini, Pardis, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. “Ask the Experts: What Should Be on an IoT Privacy and Security Label?” *Proceedings of 2020 IEEE Symposium on Security and Privacy* (2020): 447-464. https://ieeexplore.ieee.org/document/9152770.

Faklaris, Cori, Laura Dabbish, and Jason I. Hong. “A Self-Report Measure of End-User Security Attitudes (SA-6).” *Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security* (2019): 61-77. https://www.usenix.org/system/files/soups2019-faklaris.pdf.

Felt, Adrienne Porter, Alex Ainslie, Robert W. Reeder, Sunny Consolvo, Somas Thyagaraja, Alan Bettes, Helen Harris, and Jeff Grimes. “Improving SSL Warnings: Comprehension and Adherence.” *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015): 2893-2902. https://dl.acm.org/doi/10.1145/2702123.2702442.

————, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. “Android Permissions: User Attention, Comprehension, and Behavior.” *Proceedings of the Eighth Symposium on Usable Privacy and Security*, no. 3 (2012): 1-14. https://doi.org/10.1145/2335356.2335360.

————, Erika Chin, Steve Hanna, Dawn Song, and David Wagner. “Android Permissions Demystified.” *Proceedings of the 18th ACM conference on Computer and communications security* (2011): 627-638. https://doi.org/10.1145/2046707.2046779.

Frey, Suzanne. 2022. “Get More Information about Your Apps in Google Play.” Google. Accessed August 17, 2026. https://blog.google/products/google-play/data-safety/.

Gray,Colin M., Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. “The Dark (Patterns) Side of UX Design.” *Proceedings of the 2018 CHI Conference on*

Human Factors in Computing Systems, no. 534 (2018): 1-14. https://doi.org/10.1145/3173574.3174108.

Hasegawa, Ayako A., Daisuke Inoue, and Mitsuaki Akiyama. “How WEIRD is Usable Privacy and Security Research?” *Proceedings of the 33rd USENIX Conference on Security Symposium*, no. 182 (2024): 3241-3258. https://www.usenix.org/system/files/usenixsecurity24-hasegawa.pdf.

Kelley, Patrick Gage, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. “A “Nutrition Label” for Privacy.” *Proceedings of the 5th Symposium on Usable Privacy and Security*, no. 4 (2009): 1-12. https://doi.org/10.1145/1572532.1572538.

Klemmer, Jan H., Juliane Schmuser, Byron M. Lowens, Fabian Fischer, Lea Schmuser, and Florian Schaub. “Transparency in Usable Privacy and Security Research: Scholars’ Perspectives, Practices, and Recommendations.” *Proceedings of 2025 IEEE Symposium on Security and Privacy* (2025): 2658-2677.

Kovacs, Eduard. 2015. “New Chrome SSL Warning Improved Adherence Rates: Google.” SecurityWeek. Accessed August 17, 2026. https://www.securityweek.com/new-chrome-ssl-warning-improved-adherence-rates-google/.

Lassak, Leona, Elleen Pan, Blase Ur, and Maximilian Golla. “Why Aren't We Using Passkeys? Obstacles Companies Face Deploying FIDO2 Passwordless Authentication.” *Proceedings of the 33rd USENIX Conference on Security Symposium*, no. 404 (2024): 7231-7248. https://www.usenix.org/system/files/usenixsecurity24-lassak.pdf.

Lyastani, Sanam Ghorbani, Michael Schilling, Michaela Neumayr, Michael Backes, and Sven Bugiel. “Is FIDO2 the Kingslayer of User Authentication? A Comparative Usability Study of FIDO2 Passwordless Authentication.” *Proceedings of 2020 IEEE Symposium on Security and Privacy* (2020): 268-285. https://ieeexplore.ieee.org/document/9152694.

Mathur, Arunesh, Gunes Acar, Michael J. Friedman, Eli Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. “Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites.” *Proceedings of the ACM on Human-Computer Interaction* 3, no. 81 (2019): 1-32. https://doi.org/10.1145/3359183.

Matte, Celesti, Nataliia Bielova, and Cristiana Santos. “Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe’s Transparency

and Consent Framework." *Proceedings of 2020 IEEE Symposium on Security and Privacy* (2020): 791-809. https://ieeexplore.ieee.org/document/9152617.

Pearman, Sarah, Shikun Aerin Zhang, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. "Why People (Don't) Use Password Managers Effectively." *Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security* (2019): 319-338. https://www.usenix.org/system/files/soups2019-pearman.pdf.

Ray, Hirak, Flynn Wolf, and Ravi Kuber. "Why Older Adults (Don't) Use Password Managers." *Proceedings of the 30th USENIX Security Symposium* (2021): 73-90. https://www.usenix.org/system/files/sec21-ray.pdf.

Roozenbeek, Jon, Sander van der Linden, Beth Goldberg, Steve Rathje, and Stephan Lewandowsky. "Psychological Inoculation Improves Resilience against Misinformation on Social Media." *Science advances* 8, no. 34 (2022). https://www.science.org/doi/10.1126/sciadv.abo6254.

Sawaya, Yukiko, Sarah Lu, Takamasa Isohara, and Mahmood Sharif. "A High Coverage Cybersecurity Scale Predictive of User Behavior." *Proceedings of the 33rd USENIX Security Symposium* (2024). https://www.usenix.org/system/files/usenixsecurity24-sawaya.pdf.

Schechter, Emily. 2017. "Next Steps toward More Connection Security." Chromium Blog. Accessed August 17, 2026. https://blog.chromium.org/2017/04/next-steps-toward-more-connection.html.

SOUPS. 2025. "Workshop on Developing Research Transparency Guidelines in Human-centered Privacy and Security (Transparency@SOUPS'25)." Transparency @ SOUPS. Accessed August 17, 2026. https://transparency-at-soups.teamusec.de/.

Sunshine, Joshua, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. "Crying Wolf: An Empirical Study of SSL WarningEffectiveness." *Proceedings of the 18th conference on USENIX security symposium* (2009): 399-416. https://www.usenix.org/legacy/event/sec09/tech/full_papers/sunshine.pdf.

Suray, Jacques, Jan H. Klemmer, Juliane Schmuser, and Sascha Fahl. "How the Future Works at SOUPS: Analyzing Future Work Statements and Their Impact on Usable Security and Privacy Research." arXiv: 2405.20785 (2024). https://arxiv.org/abs/2405.20785.

Tolsdorf, Jan, Alan F. Luo, Monica Kodwani, Junho Eum, Mahmood Sharif, Michelle L. Mazurek, and Adam J. Aviv. “Safety Perceptions of Generative AI Conversational Agents: Uncovering Perceptual Differences in Trust, Risk, and Fairness.” *Proceedings of the Twenty-First USENIX Conference on Usable Privacy and Security*, no. 6 (2025): 93-112. https://www.usenix.org/system/files/soups2025-tolsdorf.pdf.

Unger, Nik, Sergej Dechand, Joseph Bonneau, Sascha Fahl, Henning Perl, Ian Goldberg, and Matthew Smith. “SoK: Secure Messaging.” *Proceedings of the 2015 IEEE Symposium on Security and Privacy* (2015): 232-249. https://doi.org/10.1109/SP.2015.22.

USENIX. 2024. “USENIX Security ’25 Call for Artifacts.” 34th USENIX Security Symposium. Accessed August 17, 2026. https://www.usenix.org/conference/usenixsecurity25/call-for-artifacts.

Van der Laan, Jinke D., Adriaan Heino, and Dick De Waard. “A Simple Procedure for the Assessment of Acceptance of Advanced Transport Telematics.” *Transportation Research Part C: Emerging Technologies* 5, no. 1(1997): 1-10. https://doi.org/10.1016/S0968-090X(96)00025-3.

Votipka, Daniel, Desiree Abrokwa, and Michelle L. Mazurek. “Building and Validating a Scale for Secure Software Development Self-Efficacy.” *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020): 1-20. https://doi.org/10.1145/3313831.3376754.

Zurko, Mary Ellen, and Richard T. Simon. “User-Centered Security.” *Proceedings of the 1996 workshop on New security paradigms* (1996): 27-33. https://www.nspw.org/2009/proceedings/1996/nspw1996-zurko.pdf.

別表

別表 1 SUS（System Usability Scale）の質問項目

項番	質問項目（原文）	質問項目（和訳）
1	I think that I would like to use this system frequently	このシステムを頻繁に使いたいと思う
2	I found the system unnecessarily complex	このシステムは不必要に複雑だと分かった
3	I thought the system was easy to use	このシステムは使いやすいと思った
4	I think that I would need the support of a technical person to be able to use this system	このシステムを利用するには、技術者のサポートが必要になると思う
5	I found the various functions in this system were well integrated	このシステムの様々な機能がうまく統合されていることが分かった
6	I thought there was too much inconsistency in this system	このシステムには矛盾が多すぎると思った
7	I would imagine that most people would learn to use this system very quickly	おそらくほとんどの人は、このシステムの使用方法を非常に早く習得するだろう
8	I found the system very cumbersome to use	そのシステムは非常に使いづらいと分かった
9	I felt very confident using the system	このシステムを使うことに非常に自信を持った
10	I needed to learn a lot of things before I could get going with this system	このシステムを使い始める前に、多くのことを学ぶ必要があった

別表 2 Van der Laan Acceptance Scale の質問項目

項番	質問項目（原文）	質問項目（和訳）
	I find such a system / the (...) system	システムを使った私の印象は…
1	Useful _____ Useless	役立つ _____ 役立たない
2	Pleasant _____ Unpleasant	快適である _____ 不快である
3	Bad _____ Good	悪い _____ 良い
4	Nice _____ Annoying	素晴らしい _____ 煩わしい
5	Effective _____ Superfluous	効果がある _____ 余計である
6	Irritating _____ Likeable	しゃくにさわる _____ 好ましい
7	Assisting _____ Worthless	助けとなる _____ 助けにならない
8	Undesirable _____ Desirable	望ましくない _____ 望ましい
9	Raising Alertness _____ Sleep-inducing	注意力を高める _____ 眠気を誘う

別表 3 SeBIS (Security Behavior Intentions Scale) の質問項目

項番	質問項目（原文）	質問項目（和訳）
1	I set my computer screen to automatically lock if I don't use it for a prolonged period of time.	長期間使用しない場合、コンピュータの画面が自動的にロックされるように設定した。
2	I use a password/passcode to unlock my laptop or tablet.	ノートパソコンやタブレットのロックを解除するためにパスワード/パスコードを使用している。
3	I manually lock my computer screen when I step away from it.	席を離れるときは、手動でコンピュータの画面をロックしている。
4	I use a PIN or passcode to unlock my mobile phone.	私は携帯電話のロックを解除するのに PIN またはパスコードを使用している。
5	I do not change my passwords, unless I have to.	必要がない限り、パスワードは変更しない。
6	I use different passwords for different accounts that I have.	持っているアカウントごとに異なるパスワードを使用している。
7	When I create a new online account, I try to use a password that goes beyond the site's minimum requirements.	新しいオンラインアカウントを作成する際には、サイトの最低要件を超えるパスワードを使用するように心がけている。
8	I do not include special characters in my password if it's not required.	パスワードに特殊文字は必要がない限り含めない。
9	When someone sends me a link, I open it without first verifying where it goes.	誰かがリンクを送ってくると、私はそれがどこに飛ぶのか確認せずに開いてしまう。
10	I know what website I'm visiting based on its look and feel, rather than by looking at the URL bar.	URL バーを見るのではなく、見た目や雰囲気からどのウェブサイトを訪問しているか分かる。
11	I submit information to websites without first verifying that it will be sent securely (e.g., SSL, "https://", a lock icon).	情報が安全に送信されることを事前に確認せずに（例：SSL、https://、鍵アイコン）、ウェブサイトに情報を送信する。
12	When browsing websites, I mouseover links to see where they go, before clicking them.	ウェブサイトを閲覧する際、リンクをクリックする前に、マウスオーバーしてリンク先を確認する。
13	If I discover a security problem, I continue what I was doing because I assume someone else will fix it.	セキュリティ上の問題を発見しても、私はそのまま作業を続ける。誰かが修正してくれるだろうと考えるから。
14	When I'm prompted about a software update, I install it right away.	ソフトウェアの更新を促されたら、すぐにインストールする。
15	I try to make sure that the programs I use are up-to-date.	私は使用するプログラムが最新の状態であることを確認するように努めている
16	I verify that my anti-virus software	自分のアンチウイルスソフトウェアが定

	has been regularly updating itself.	期的に更新されていることを確認している。

別表 4 SA-6 (A Self Report Measure of End-User Security Attitudes) の質問項目

項番	質問項目（原文）	質問項目（和訳）
1	I seek out opportunities to learn about security measures that are relevant to me.	私は、自分に関連するセキュリティ対策について学ぶ機会を積極的に探している。
2	I am extremely motivated to take all the steps needed to keep my online data and accounts safe.	オンライン上のデータとアカウントを安全に保つために必要なあらゆる措置を講じることに、私は非常に意欲的だ。
3	Generally, I diligently follow a routine about security practices.	一般的に、私はセキュリティ対策に関するルーティンを忠実に守っている。
4	I often am interested in articles about security threats.	私はセキュリティ脅威に関する記事によく興味を持つ。
5	I always pay attention to experts' advice about the steps I need to take to keep my online data and accounts safe.	私はオンライン上のデータやアカウントを安全に保つために取るべき手順について、専門家の助言には常に注意を払っている。
6	I am extremely knowledgeable about all the steps needed to keep my online data and accounts safe.	私はオンライン上のデータやアカウントを安全に保つために必要なすべての手順について、非常に詳しい知識を持っている。

別表 5 ESBS (Extended Security Behavior Scale) の質問項目

項番	質問項目（原文）	質問項目（和訳）
Factor 1: Data Securement（データの保全）		
1.1	I encrypt my email contents when sending sensitive information (e.g., banking and health information or social security number)	機微な情報（銀行情報や健康情報、社会保障番号など）を送信する際には、メールの内容を暗号化している。
1.2	I validate the digital certificates on the websites I visit	訪問するウェブサイトのデジタル証明書を確認している。
1.3	I review the validity of my root certificates	自分のルート証明書の有効性を確認している。
1.4	I validate the digital signatures files before opening them	ファイルを開く前に、そのデジタル署名を確認している。
1.5	I physically destroy drives I am done using and wish to erase	使い終えて消去したいドライブは、物理的に破壊している。
1.6	I encrypt my devices' disks to keep my data confidential	自分のデータを機密に保つため、デバイスのディスクを暗号化している。
1.7	I safely store my private key for email encryption	メール暗号化用の秘密鍵を安全に保管している。
Factor 2: Proactive Awareness（プロアクティブ意識）		
2.1	I verify whom I communicate with online (via email or online messaging apps) is really the person I intend to	オンライン（メールやメッセージアプリ）でやり取りする相手が、本当に意図した相手であるかを確認している。
2.2	I verify links (e.g., in the URL bar or by mouseover) to ensure that I am accessing intended websites	意図したウェブサイトにアクセスしていることを確かめるため、リンク（URL バーやマウスオーバーなど）を確認している。
2.3	I check the extensions (e.g., .exe, .pdf) of files I download	ダウンロードするファイルの拡張子（.exe や.pdf など）を確認している。
2.4	I turn on download notifications in my browsers	ブラウザのダウンロード通知を有効にしている。
2.5	When possible, I use two- or multi-factor authentication	可能な場合は、二要素認証または多要素認証を使用している。
2.6	I disable auto-run to prevent potentially malicious downloaded programs from running	ダウンロードした悪意のあるプログラムが実行されるのを防ぐため、自動実行を無効にしている。

項番	質問項目（原文）	質問項目（和訳）
Factor 3: Anti-virus（アンチウイルス）		
3.1	I scan attachments for viruses before downloading or opening them	添付ファイルは、ダウンロードや開封の前にウイルススキャンをしている。
3.2	I verify that my anti-virus software is up-to-date	自分のアンチウイルスソフトウェアが最新の状態であることを確認している。
3.3	I install anti-virus software when setting up my devices	デバイスを設定する際に、アンチウイルスソフトウェアをインストールしている。
Factor 4: Updates（更新）		
4.1	I turn on automatic updates for devices and applications upon installation	デバイスやアプリケーションのインストール時に、自動更新を有効にしている。
4.2	When I am prompted about a device or software update, I immediately install it	デバイスやソフトウェアの更新を促されたら、すぐにインストールしている。
Factor 5: Passwords（パスワード）		
5.1	I select hard-to-guess passwords (with multiple character types, without dictionary words, etc.)	推測されにくいパスワード（複数の文字種を含み、辞書に載っている単語を使わないものなど）を選んでいる。
5.2	I select different passwords for different accounts and devices	アカウントやデバイスごとに異なるパスワードを選んでいる。

別表 6 SSD-SES (Secure Software Development Self-Efficacy Scale) の質問項目

項番	質問項目（原文）	質問項目（和訳）
Vulnerability Identification and Mitigation（脆弱性の特定と緩和）		
C3	I can perform a threat risk analysis (e.g., likelihood of vulnerability, impact of exploitation, etc.)	脅威のリスク分析（脆弱性が生じる可能性や、悪用された場合の影響など）を行うことができる。
C4	I can identify potential security threats to the system	システムに対する潜在的なセキュリティ脅威を特定することができる。
C6	I can identify the common attack techniques used by attackers	攻撃者が用いる一般的な攻撃手法を特定することができる。
C11	I can identify potential attack vectors in the environment the system interacts with (e.g., hardware, libraries, etc.)	システムが関わる環境（ハードウェアやライブラリーなど）における潜在的な攻撃経路を特定することができる。
C15	I can identify common vulnerabilities of a programming language	プログラミング言語に共通してみられる脆弱性を特定することができる。
C39	I can design software to quarantine an attacker if a vulnerability is exploited	脆弱性が悪用された場合に攻撃者を隔離できるよう、ソフトウェアを設計することができる。
C44	I can mimic potential threats to the system	システムに対する潜在的な脅威を模擬することができる。
C47	I can evaluate security controls on the system's interfaces/interactions with other software systems	他のソフトウェアシステムとのインターフェースや相互作用に関するセキュリティ対策を評価することができる。
C48	I can evaluate security controls on the system's interfaces/interactions with hardware systems	ハードウェアシステムとのインターフェースや相互作用に関するセキュリティ対策を評価することができる。
Security Communication（セキュリティに関するコミュニケーション）		
C50	I can communicate security assumptions and requirements to other developers on the team to ensure vulnerabilities are not introduced due to misunderstandings	誤解によって脆弱性が生じないよう、セキュリティ上の前提や要件をチームの他の開発者に伝えることができる。
C51	I can communicate system details with other developers to ensure a thorough security review of the	コードのセキュリティレビューを徹底するため、システムの詳細を他の開発者に伝えることができる。

項番	質問項目（原文）	質問項目（和訳）
	code	
C53	I can discuss lessons learned from internal and external security incidents to ensure all development team members are aware of potential threats	開発チームの全員が潜在的な脅威を認識できるよう、社内外のセキュリティ事案から得られた教訓を共有し議論することができる。
C55	I can effectively communicate to company leadership identified security issues and the cost/risk trade-off associated with deciding whether or not to fix the problem	特定したセキュリティ上の問題と、その対処の要否に伴うコストとリスクのトレードオフを、経営層に効果的に伝えることができる。
C57	I can communicate functionality needs to security experts to get recommendations for secure solutions (e.g., secure libraries, languages, design patterns, and platforms)	安全な解決策（安全なライブラリーや言語、設計パターン、プラットフォームなど）の助言を得るため、必要な機能をセキュリティ専門家に伝えることができる。
C58	I know the appropriate point of contact/response team in my organization to contact if a vulnerability in production code is identified	本番コードに脆弱性が見つかった場合に連絡すべき、組織内の適切な窓口や対応チームを把握している。

（注）開発においては、セキュア開発フレームワークのレビューや専門家調査から 58 の候補を抽出し、開発者らによる複数ラウンドの評価を経て、信頼性・妥当性を確認したうえで最終的に 15 項目に絞られた（Votipka, Abrokwa, and Mazurek 2020）。

別表 7 SP-SDS (Security & Privacy Social Desirability Scale) の質問項目

略称（原文）	質問項目（原文）	質問項目（和訳）
data collected	I am always aware of what personal data is collected by the platforms I use and how it is used	自分が利用するプラットフォームによって、どのような個人データが収集され、どのように使われているかを常に把握している。
different passwords	I use different passwords for all my accounts	すべてのアカウントで異なるパスワードを使用している。
ignore update	I never ignore software update reminders	ソフトウェアの更新通知を無視することは決してない。
ignore warnings	I never ignore security warnings on my computer	コンピュータのセキュリティ警告を無視することは決してない。
illegal movies*	If I knew I wouldn't get caught, I would watch movies illegally	見つからないと分かっていれば、違法に映画を視聴するだろう。
pirated software*	If I knew I wouldn't get caught, I would use pirated software	見つからないと分かっていれば、海賊版ソフトウェアを使用するだろう。
policy access	I always read the privacy policy before giving an app/application access to my personal data	アプリケーションに自分の個人データへのアクセスを許可する前には、必ずプライバシーポリシーを読んでいる。
polite online	I am always just as polite online as I am in the real world	オンライン上でも、現実世界と同じように常に礼儀正しくふるまっている。
read policy	I always read the privacy policy completely before I agree to them	同意する前に、プライバシーポリシーを必ず最後まで読んでいる。
read terms	I always read the terms and conditions completely before I agree to them	同意する前に、利用規約を必ず最後まで読んでいる。
reuse passwords	I never reuse passwords	パスワードを使い回すことは決してない。
secure passwords	I only use secure passwords (passwords that are long and complex)	安全なパスワード（長く複雑なパスワード）のみを使用している。
troll comment	I have never considered posting a troll-comment	荒らしコメントを投稿しようと考えたことは一度もない。

（注）アスタリスク（*）を付した 2 項目（illegal movies、pirated software）は逆転項目であり、分析前に、点数を反転させる処理を行う。

補論．ユーザブルセキュリティ研究に用いる統計手法

（１）検定手法

ユーザブルセキュリティ研究における量的研究の実証実験では、それぞれ異なるグループや条件における検証結果に有意な違いがあるかについて、状況に応じた適切な統計手法を用いて評価する（表 A-1 を参照）。

独立した 2 つのグループ（2 群）の結果を比較する場合には、それぞれのグループから得られたデータの分布に応じて適切な検定手法を選択する。従来は正規分布かつ等分散を仮定できる場合に「t検定[31]」、分散が等しくない場合に「Welch のt検定[32]」を用いるとされてきた。しかし、近年では、分散の等質性を事前に検定して手法を分岐することにより統計的なバイアスが生じる可能性が指摘されており、実務上は分散の等質性を仮定しない「Welch のt検定」を用

表 A-1 検定手法の比較

比較対象	パラメトリック（正規分布あり）	ノンパラメトリック（順位尺度など）	概要
2 群（独立）	Welch のt検定	Mann-Whitney のU検定	独立した 2 群間で代表値に差があるかをみる手法
2 群（対応あり）	対応のあるt検定	Wilcoxon の符号順位検定	対応のある 2 群間で代表値に差があるかをみる手法
3 群以上（独立）	一元配置 ANOVA（等分散） Welch-ANOVA（不等分散）	Kruskal-Wallis 検定	少なくとも一つの群が他と異なる可能性があるかを判定する手法
多重比較補正	Bonferroni 補正、Holm 法、Benjamini-Hochberg 法		検定を繰り返すことによる偽陽性を防ぐ手法

[31] t検定は、2 群の平均値の差を、データのばらつきである標準誤差に基づいて評価する統計的仮説検定手法。帰無仮説を「両群の平均値に差がない」と設定し、データの平均値の差が標準誤差の何倍に当たるかを示す指標であるt値を算出する。さらに、算出されたt値に基づき、観察された結果と同じかそれ以上に極端な結果が偶然に得られる確率であるp値（p-value）を算出する。なお、統計的仮説検定では、一般にp値を用いて統計的有意性を評価する。通常、p< 0.05（5%未満）の場合には「統計的に有意」と判断し、得られた結果が単なる偶然の産物であるとは考えにくいと判断する。

[32] 従来のt検定は 2 群の分散が等しいことを仮定するのに対し、この仮定を前提としないt検定の拡張手法。各群の分散とサンプルサイズを考慮し、統計分析においてデータの独立性を反映した値である自由度を補正することで、より汎用的に 2 群の平均値の有意差を判定する。

いることが推奨される。また、データが正規性を満たさない場合には、「Mann–Whitney のU検定[33]」を用いることが一般的である。

同じグループに対して異なる 2 条件の検証結果を比較する場合であって、データに正規性があるケースでは「対応のあるt検定[34]」、正規性を有しないケースでは「Wilcoxon の符号順位検定[35]」が使われている。独立した 3 グループ（群）以上の比較では、「一元配置 ANOVA[36]（analysis of variance, 分散分析）」または「Welch-ANOVA[37]」を用いる。不等分散や非正規性が顕著な場合には「Kruskal–Wallis 検定[38]」が選択肢となる。

なお、複数のグループを比較するときの留意事項として、比較する回数が増えるほど、偶然グループ間に有意な差があるようにみえる結果（偽陽性）を検出するリスクが高まる。対策として、偽陽性の防御と検出力のバランスがよいとされる Holm 法[39]や Benjamini–Hochberg 法[40]が有効である[41]。

---

33 2 群の独立したサンプル間において、母集団の分布に差があるかを検討する手法。データを直接の値ではなく順位に変換して計算するため、正規分布を仮定できない場合や、外れ値を含むデータ、または順序尺度の分析に適している。

34 同一の調査対象から得られたペア（対応）のある 2 つのデータ群について、その差の平均値が 0 であるかを判定する手法。個体差によるばらつきを排除して比較できるため、通常のt検定（独立サンプル）に比べて検出力が高まる特性を持つ。

35 対応のある 2 群間において、差の大きさを順位に変換して統計的有意性を評価する手法。正規分布を前提としないため、「対応のあるt検定」の代替として用いられる。データの順序関係だけでなく、差の程度（強度）も考慮できる点が特徴である。

36 3 群以上の独立したサンプル間における平均値の差を検定する手法。各群内のばらつき（群内変動）と群間のばらつき（群間変動）の比（F値）を算出し、全体の平均から各群の平均が統計的に有意に離れているかを評価する。

37 各群の母分散が等しいことを前提としない分散分析の手法。標準的な一元配置 ANOVA は等分散性を仮定するが、実データにおいてこの仮定が満たされない場合、第一種の過誤（誤検定）が増大する。本手法は、自由度を補正することで、等分散性が保たれていない状況下でも信頼性の高い検定を可能にする。

38 3 群以上の独立したサンプル間における分布の差を検定する手法。データを順位に変換して評価するため、正規分布を仮定できない場合や外れ値を含むデータ、順序尺度の分析に適している。

39 複数の検定を行う際に第一種の過誤を制御する、段階的棄却法。各検定のp値を小さい順に並べ、有意水準を「α / (検定回数-順位+ 1)」と段階的に緩和しながら比較する。

40 全ての検定において第一種の過誤を厳格に抑える Bonferroni 補正や Holm 法とは異なり、有意と判定された項目のうち実際に差がないものが含まれる割合を許容範囲内に収める手法。

41 Bonferroni 補正を用いると偽陽性をほぼ確実に防ぐことが可能となるが、一方で本来検出すべき差までも見逃してしまう可能性がある。Bonferroni 補正は、同一データに対して複数の検定を繰り返す際、本来は差がないのに有意と判定する誤りである第一種の過誤の率が上昇する問題を回避するための補正手法。全体の有意水準を検定回数で除した値を、各検定の新たな有意水準として用いる。

得られた検定結果を解釈する際には、異なるグループや条件間の有意差の有無だけでなく、差の具体的な大きさを示す推定値とその推定の不確実性を表す信頼区間に着目することが重要である。有意差は「差が偶然か否か」を示すにとどまり、その差が実務上で意味のある大きさ（効果量）であるかは「推定値」から判断する必要がある。また、信頼区間は、サンプルサイズの影響を受けやすいp値に対し「真の値がどの程度の範囲に収まりそうか」という情報の精度を明示するため、より多角的な意思決定の判断材料を提供できる。

（２）回帰分析

回帰分析は、ある要因が結果に対してどの程度どのような方向に影響を与えているかという変数間の関係性を明らかにする。

回帰分析では、分析したい結果（従属変数[42]）のデータ形式によって、用いる数式（モデル）を使い分ける。例えば、従属変数が連続変数の場合には線形回帰が、二値変数の場合にはロジスティック回帰が、カウントデータの場合にはポアソン回帰や負の二項回帰が広く用いられる（表 A-2 を参照）。

参加者個人の個性やデータが収集された場所・環境の違いなど、複雑な要因が絡み合って結果に影響を及ぼし得るケースでは、線形混合モデル（linear mixed model : LMM）[43]や一般化線形混合モデル（generalized linear mixed model : GLMM）[44]が用いられる。

また、分析の正確性を保つため、説明変数同士が似すぎていないか（多重共線性）を確認し、データの分散が一定ではない場合には、統計的な補正（ロバスト

表 A-2 回帰分析手法の比較

調査したい結果の形式	推奨される回帰モデル	具体例
連続変数	線形回帰	利用時間、スコア
二値変数（二択、Yes/No）	ロジスティック回帰	攻撃にあったか否か、承諾したか否か
カウントデータ	ポアソン回帰、負の二項回帰	エラー発生回数、クリック数

42 分析において予測や説明の対象となる変数のこと。対して、その変動を説明するために用いる変数を独立変数または説明変数と呼ぶ。

43 データ全体の傾向だけでなく、実験参加者個人のクセや、測定場所ごとの違いなど、無視できない「個別のばらつき」を同時に計算に組み込む手法。

44 一般化線形モデルと混合モデルを組み合わせることにより、連続値や二値データなどさまざまなデータ型に対応可能な分析手法。一般化線形モデルではデータと説明変数の関係性をモデル化し、混合モデルでは全体に共通する要因やグループ間・個人間で異なる要因を計算に組み込み、データの構造的特性を反映させる。

標準誤差[45]）を適用して結論の信頼性を高める。

（３）再標本化

ユーザブルセキュリティ研究で用いられる再標本化の代表的な手法として、信頼区間の推定に用いられるブートストラップ法や、群間差の有意性を判定する置換検定（ランダマイゼーション検定）が挙げられる。

ブーストラップ法は、観測された標本データから、重複を許して繰り返し再抽出を行うことにより同じサイズの新しいデータを複数作成し、統計量の標本分布を近似的に構成する手法である。母集団（標本データの抽出元となる全データ）の分布形状を問わず、信頼区間の推定や標準誤差の算出を頑健に行うことができるため、小標本かつ正規性が疑わしいデータに対して有効である。

置換検定は、群間のラベルをランダムに並べ替えて「差がない」という帰無仮説に基づく分布をコンピュータ上で生成し、観測された統計量がその分布のどの位置にあるかを算出することでp値を求める手法である。

45 データのばらつき（分散）が一定であるという統計学的な仮定が満たされない場合に、分析結果の信頼性を確保するために用いられる補正手法。