インフォームド・コンセントに基づく Telegram 上の人的情報収集の実施可能性の検証

久保 颯汰[1,a)]　長山 侑央[2]　インミンパパ[3]　森 辰則[4]　吉岡 克成[4]

概要：
Telegram などの SNS を用いた人的情報収集 (HUMINT) に関する研究は, 脅威インテリジェンスの構築において極めて有益であるが, 対象が人間であることから人権を尊重しつつ適切に実施する必要がある. サイバーセキュリティの研究倫理に関する指針を示したメンロレポートでは, 人間を対象とする研究においてインフォームド・コンセントを通じて研究参加者にリスクや権利を通知する必要があることが示されている. 一方, 実験の性質上, 被験者への事前説明が困難な場合には, 事後的な説明など被験者への配慮を行った上でインフォームド・コンセントなしに実験を実施する場合や, 事後的な説明すら行わずに研究を実施する場合がある.
HUMINT においては, 会話の目的を明かすことにより脅威アクターの警戒心を高め, 情報収集自体が困難となる恐れがあるため, 被験者への説明を含めた研究の実施方法について慎重な検討が求められる. 本研究では, その検討の第一歩として, 会話の目的を明かした上で脅威アクターから情報収集を試みることにより, インフォームド・コンセントに基づく HUMINT 研究の実施可能性を検証する.

キーワード：インフォームド・コンセント, 研究倫理, サイバーセキュリティ研究

Examining the Feasibility of Conducting Human Intelligence (HUMINT) on Telegram Based on Informed Consent

Sota Kubo[1,a)]　Yukihiro Nagayama[2]　Yin Minn Pa Pa[3]　Tatsunori Mori[4]
Katsunari Yoshioka[4]

Abstract: Human intelligence (HUMINT) conducted via social networking services such as Telegram can be extremely valuable in building threat intelligence. However, as it involves interaction with human subjects, it must be conducted appropriately with due respect for human rights. The Menlo Report emphasizes the necessity of informing research participants of potential risks and their rights through informed consent when conducting research involving humans. However, due to the nature of certain experiments, prior explanation to subjects may be difficult. In such cases, research may be conducted with post-experiment explanations or, in some cases, without any explanation at all.
In the context of HUMINT, revealing the purpose of a conversation may increase the target threat actor's awareness and caution, thereby making information gathering itself more difficult. Thus, careful consideration is required when determining how to conduct research, including whether and how to inform participants. As a first step in this consideration, this study attempts to collect information from threat actors while clearly disclosing the purpose of the conversation, in order to examine the feasibility of conducting HUMINT research based on informed consent.

Keywords: Informed Consent, Research Ethics, Cybersecurity Research

1. はじめに

Telegram 等のソーシャルネットワーキングサービス（SNS）を用いた人的情報収集（HUMINT）は，脅威インテリジェンス構築において有益な手段として広く活用されている [1], [2], [3], [4], [5]．しかしながら，HUMINT に関する研究は対象が人間であることから，人権を尊重し，倫理的かつ適切な方法で実施されることが不可欠である．特に，メンロレポート [6] によれば人間を対象とする研究においては, 研究参加者からのインフォームド・コンセントの取得が原則として求められる．これは, 研究の目的, 手順, 潜在的なリスク, そして参加者の権利を事前に明確に説明し, その上で自由意思に基づく同意を得るという, 倫理的な研究遂行の基盤をなす概念である．

一方で, HUMINT の特性上, このインフォームド・コンセントの取得が困難となる場合が存在する．例えば, Telegram のような不特定多数のユーザーが互いに対話する SNS において, インフォームド・コンセントの概念に忠実に作成した長文の同意書に同意してもらうことが, その趣旨を理解した上で同意する事を意味するとは断定できない．また, 情報収集の目的を脅威アクターに事前に開示することが, 彼らの警戒心を高め, 結果として情報収集自体を不可能にする恐れがある．このような状況下において, 倫理的要件と研究の実効性との間で, どのようにバランスを取り, 研究を実施すべきかという点が重要な課題となっている．

人間を対象とする研究における倫理的指針については, メンロレポートにおいて詳細に議論されている．同レポートでは, 研究参加者に対して研究のリスクや権利をインフォームド・コンセントを通じて通知することの重要性が強調されている．しかし, 研究の性質上, 事前説明が困難な特定の状況においては, 事後的な説明の可能性や, 限定的な状況下での説明なしの研究実施についても言及されている．ただし, HUMINT におけるインフォームド・コンセントの具体的な適用事例や, 脅威アクターに対する情報収集における倫理的側面に特化した先行研究は限定的である．

本研究は, 上記の課題意識から，インフォームド・コンセントを前提とした HUMINT 研究の実施可能性を検証することを目的とする．具体的には, 以下のリサーチクエスチョン（RQ）を設定した：

RQ: 会話の目的を明かした上で脅威アクターから情報収集を試みた場合, 脅威アクターは実験への参加に同意し, 有効な情報提供を行うか？

本研究では, 上記の RQ に答えるため, 会話の目的を明かした上で脅威アクターから情報収集を試みるアプローチを採用した．具体的には, Telegram 上において, インフォームド・コンセントに相当する, 実験趣旨の説明, 研究者としての身分開示，および商品購入意思がないことを明記した同意書を対象者に送信した．通知内容が長文であることを考慮し，形式的な同意だけでなく内容の理解度を確認するため，追加の質問により理解度を検証した．研究者側に購入意思がないことを理解し，かつ成人であることを申告した対象者に限り，先行研究 [7] で用いられた HUMINT 支援用 LLM（Large Language Model）の出力を会話に使用した．LLM を使用する際は，その出力内容に倫理的な問題がないことを, 複数人による判定に基づき保証するプロセスを導入した．

本研究では，会話の目的を明かすことで脅威アクターの警戒心が高まるという仮説に基づき，ほとんどの対象者から返信が得られないものと予想していた．

2025 年 8 月 17 日時点での予備的な結果として, 複数のカテゴリから対象者を 30 人選定し, 前述のインフォームド・コンセントを送信したところ, 9 人から返信を得ることができた．そのうち, 研究意図を理解し, LLM による会話の段階に進んだ被験者は 0 人であった．インフォームド・コンセントを送信した被験者のうち，21 人からは返信がなかった．

本研究は, インフォームド・コンセントに基づく HUMINT 研究の実施可能性を検証する上での第一歩となるものである．この検証を通じて得られる知見は, 人権を尊重しつつ, より倫理的な脅威インテリジェンス構築のための HUMINT 研究のあり方について検討を進めるための基礎的なデータを提供する．これにより, 将来的に, 倫理的配慮と実効性を両立させた HUMINT 研究の確立に貢献できるものと考える．

2. 関連研究

2.1 HUMINT とサイバー脅威インテリジェンス

Bank Security [1], The Hacker News [2], Cyberint [3] などの報告は, Telegram をはじめとする SNS を利用した人的情報収集（HUMINT）が, サイバー犯罪コミュニティ内部の知見を得るために重要な手段であることを強調している．これらは HUMINT がサイバー脅威の早期発見や攻撃手法の把握に有効であることを具体的に示している．しかし, これらの研究や報告の焦点はあくまで運用上の有効性や実務的側面にあり, 研究倫理やインフォームド・コンセントの適用といった学術的な倫理的基盤の議論は欠落して

[1] 横浜国立大学理工学部
Faculty of Science and Engineering,Yokohama National University
[2] 横浜国立大学環境情報学府 Graduate School of Environment and Information Sciences,Yokohama National University
[3] 横浜国立大学大学院先端科学高等研究院
Institute of Advanced Sciences, Yokohama National University
[4] 横浜国立大学大学院環境情報研究院/先端科学高等研究院
Faculty of Environment and Information Sciences, Yokohama National University / Institute of Advanced Sciences, Yokohama National University
[a)] kubo-sota-dj@ynu.jp

いる．したがって, HUMINT を学術研究として確立するためには, 実用的視点に加えて倫理的枠組みの構築が必要である．

2.2 LLM を活用した HUMINT の自動化

鈴木ら [7] は, LLM を活用した HUMINT フレームワークを提案しており,「擬似攻撃者 LLM」と「HUMINT エージェント LLM」の対話を通じて調査者の知識不足や言語的ギャップを克服することを目指している．さらに,「法倫理監視 LLM」により対話の倫理性を監視する仕組みも導入されている． しかし, 倫理的懸念から実際の攻撃者ではなく擬似攻撃者 LLM とのロールプレイに留まっており, 提案手法が現実の攻撃者を対象とする場面で有効に機能するかは未検証である．このため, 実際の攻撃者との対話が伴う HUMINT 研究にどのように適用できるかは未だ大きな課題として残されている．

2.3 サイバー犯罪研究におけるインタビューとその課題

Hutchings[8] は, サイバー犯罪研究において定性的インタビューを実施する際のベストプラクティスを調査している．研究者はサイバー犯罪者に対して対面・電子メール・チャット・ビデオ通話など多様な方法を用いてきたが, 匿名性や信頼性の確保, データの真実性をめぐる課題が指摘されている．この研究は, HUMINT が抱える特有の困難性を整理した点で重要であるが, 倫理的要件との関係性は限定的にしか触れられていない．

2.4 インフォームド・コンセントと覆面調査の倫理

Healy [9] は, 犯罪学研究におけるインフォームド・コンセントの質を重視し, 書面や署名の形式よりも同意の実質的な理解が重要であると論じている．一方, Calvey [10], [11] は, 覆面民族誌調査を通じて, 研究者の身分を明かさずに調査を行うことが倫理的観点から避けられる一方で, インフォームド・コンセント自体が研究実施を困難にし得ることを示した．また, 医学系モデルに基づく従来の倫理審査が社会科学研究の現実に必ずしも適合しない点も指摘されている．これらの議論は, HUMINT 研究におけるインフォームド・コンセントの難しさを理解する上で有益であるが, サイバー空間における攻撃者を対象とした事例研究は乏しい．

2.5 研究ギャップと本研究の位置づけ

以上の先行研究から, HUMINT はサイバー脅威インテリジェンスにおいて有効な手段であること, また研究倫理におけるインフォームド・コンセントの重要性やその困難性が議論されてきたことが分かる．しかし, **実際の脅威アクターを対象とし, インフォームド・コンセントを前提に HUMINT 研究の実施可能性を検証した事例は存在しない**．本研究はこのギャップを埋めることを目的とし, 実際のサイバー攻撃者に対して会話の目的を開示し, インフォームド・コンセントを試みることで, 倫理的配慮と研究実効性の両立が可能かを初めて実証的に検討する．

3. 前提知識

3.1 インフォームド・コンセント

メンロレポートには. インフォームド・コンセントが, 研究者が研究プロジェクトとそれに伴うリスクを被験者に正確に説明し. 被験者がリスクを受け入れ. 参加に同意するか拒否するかを決定するプロセスであると述べられている. このプロセスは. 「通知, 理解, 自発性」の 3 つの要素から構成される．

メンロレポートの立場に基づけば. 研究者はインフォームド・コンセントを通じて. 研究の意図, 被験者が被るリスク, および被験者の権利など, 多岐にわたる項目を詳細に説明し, 同意を得ることが求められる．しかしながら, 本研究で対象とする攻撃者は不特定多数のユーザーが利用する SNS である Telegram を利用しており, 被験者が他の多数の相手とやり取りを行っていることが想定される．そのような環境において, 長文の同意文を一方的に送付し同意する旨の返信が得られたとしても, 被験者が研究の意図を十分に理解した上で協力していると断定することは困難である．

以上の状況を踏まえ, 本研究では Telegram 上の攻撃者に対して HUMINT を実施するにあたり，インフォームド・コンセントで確認すべき内容を簡潔に要約した文書を事前に送付する手法を採用した. 返信があった際には. 特に確認すべき 2 つの項目

- 研究者に商品, 情報の購入意思はないこと
- 未成年は実験に参加できないこと

について追加の質問を行うことでインフォームド・コンセントの取得を図った. ここで, 被験者が研究目的であることを理解しているかを確認することも重要であるが, その内容は同意文の冒頭に記載があり, 他の項目に比べ見逃しが起きづらいと考えられることと, 商品を購入しないことが理解されている場合, 研究目的の会話であることも理解されていることが妥当であると判断したため, この内容を直接聞く項目は省略した．

以上のプロセスを踏むことで, Telegram のダイレクトメッセージ (DM) 上では詳細な説明が困難であると考えられる研究内容を説明せずとも, 参加者がその意図を理解した上で研究に協力する, 質の高い同意が得られると考えた．

3.2 HUMINT（Human Intelligence）

HUMINT（Human Intelligence）は，人間を情報源とするインテリジェンス収集手法であり，諜報活動のみならずサイバー脅威インテリジェンス（CTI）の分野においても

広く活用されている．近年では，Telegram 等の SNS やクローズドなオンラインフォーラムが攻撃者コミュニティの主要な活動基盤となっており，これらの場における観察や潜入を通じた人的情報収集が重要性を増している [2], [4], [5].

HUMINT によって得られる情報は多岐にわたる．たとえば，攻撃者が取引するアクセス権や利用するツール，標的となる組織や産業分野，流通している認証情報やゼロデイ脆弱性に関する知見などが含まれる [1], [3]．さらに，攻撃者同士の評判や協力関係，動機といった非技術的側面も把握可能であり，技術的データのみでは理解しにくい攻撃者の意図や行動様式の解明に寄与する [1].

具体的に Bank Security によれば，HUMINT の利点には「閉鎖フォーラムに投稿される前の機密情報の取得」，「未公表の被害対象の把握」，「新たな脆弱性やマルウェアの早期発見」，「新規に出現する監視対象源（Telegram チャネルやプライベートグループ等）の特定」，「脅威アクターの背景理解」などが含まれ，自動化ツールでは代替できない価値を持つとされている [1].

このように，HUMINT は OSINT（Open Source Intelligence）や自動化ツールによる収集と補完的に活用されることで，より包括的な脅威理解と先制的な防御戦略の構築を可能にする．

4. 調査手法

本研究では，インフォームド・コンセントに基づく HUMINT の実施可能性を検証するため，以下の手順で研究を実施する．この手法は，会話の目的を事前に明示することで，倫理的配慮と研究目的の達成との両立を試みるものである．

4.1 調査対象の選定

本研究では，Telegram 上で不正なツールや情報を売買していると考えられる脅威アクターのうち，Telegram の検索機能を用いて投稿者に直接メッセージを送信できることを確認した 30 件を調査対象とした（表 1）．

このうち，No.1〜No.10 および No.14 の計 11 件は，青砥ら [12] が収集した Telegram チャンネルのデータを分析し，悪意のある投稿として選定したものである．一方，残りの 19 件については，別途 Telegram 上で手動検索を行い，不正ツールの売買を示唆する広告が頻繁に投稿されている公開グループを特定し，その投稿者の投稿を対象として選定した．各投稿の概要を表 1 に示す．選定方法を 2 種類に分けた理由は，手動検索のみに依拠すると対象となる投稿が偏る可能性があると考えたためである．将来的に HUMINT をさまざまな投稿を対象に実施する事を視野に入れ，多様な投稿に対して調査を行うため，自動化された収集システムによる選定も併用した．実際，この方法により，詐欺の疑いがある投稿 (No.1, No.3, No.6, No.8)，試験代行に関する投稿 (No.7)，ソフトウェアライセンス違反の疑いがある投稿など，手動検索と傾向の異なる対象を調査に含めることができた．

選定されたユーザーに対しては，その投稿内容から，研究対象としての妥当性を慎重に検討した．詐欺の疑いがある投稿については投稿内容のみから違法性（詐欺行為を実際に行なっているか）を確かめることはできないが，ここで選定した投稿はその内容が商品の具体的な情報を出さず，早めの購入を迫るなどといった詐欺の特徴に合致していると人手で判断したものであり，違法性の有無に関わらず HUMINT の対象になると判断したため，今回調査を行うものとした．

表 1 調査対象となった Telegram 投稿の一覧

No.	ポスト内容
1	仮想通貨に関連する詐欺の疑い $_1$
2	個人情報，企業情報の販売
3	仮想通貨に関連する詐欺の疑い $_2$
4	仮想通貨に関連する詐欺の疑い $_3$
5	仮想通貨に関連する詐欺の疑い $_4$
6	ソフトウェアのライセンス販売に関わる詐欺の疑い
7	資格試験の代行受験や不正行為
8	投資詐欺の疑い
9	無登録での違法な金融商品の取引勧誘
10	ソフトウェアの「クラック版（不正な改変版）」の無償配布
11	複数のゲームアカウントの販売
12	迷惑メール送信ツール，詐欺サイトのテンプレート，個人情報販売 $_1$
13	暗号資産に関する詐欺，迷惑メール送信，ワンタイムパスワードの不正取得
14	SMS を用いた詐欺ツールの販売
15	なりすまし電話用ツール
16	システムのハッキング，スパイ行為
17	迷惑メール送信ツール，詐欺に悪用されるツール，ハッキング手法の販売
18	迷惑メール送信ツール，詐欺サイトのテンプレート，個人情報販売 $_2$
19	銀行口座ログイン情報，偽装カードの販売 $_1$
20	銀行口座ログイン情報，偽装カードの販売 $_2$
21	なりすまし電話用ツール販売 $_1$
22	なりすまし電話用ツール販売 $_2$
23	なりすまし電話用ツール，個人情報販売
24	迷惑メール送信ツール，詐欺サイトのテンプレート，個人情報販売
25	迷惑メール送信ツール，個人情報販売
26	DDOS サービスの販売
27	詐欺ツール販売
28	迷惑メール送信サービス，フィッシングツールの販売
29	迷惑メール送信サービス
30	なりすまし電話用ツール販売 $_3$

4.2 インフォームド・コンセントに基づく対話の実施

選定された調査対象に対し, 以下のような対話プロトコルに従って接触を試みる.

(1) **初期接触と目的の開示**: 対象ユーザーに対し, 自身が日本の研究者であることと, 研究目的を明確に伝える. 具体的には, 図 1 で示す内容を送信する.

> "I'm a university researcher in Japan studying how markets operate on Telegram. If you're willing, I'd like to briefly chat with you here on Telegram (just a few minutes) about how you use this platform. Participation is voluntary — you can stop anytime. Your identity will remain completely anonymous. Findings may be used in an academic paper. This is a survey, so I will NOT be purchasing any of your products. Also, if you are a minor, you CANNOT participate in this experiment. If you agree to these terms, I would appreciate your response."

図 1 初期接触時に送信した同意文の原文

メンロレポートに基づけばこの同意文で研究内容や被験者のリスク, 研究者の利益に詳細に説明することが期待されるが, Telegram の不特定多数のユーザが互いに対話を行うという性質上，長いメッセージは読み飛ばされてしまう恐れがあるため, 必要な情報に絞って端的に説明を行った. 本来説明すべき項目について被験者から質問を受けた場合, その都度説明を行うものとする.

(2) **理解度の確認**: 相手からの返信を確認後, 図 2 で示す 2 つの質問を順番に行い, 同意文の内容を理解していることを確認する. これらの質問は同意文の内容を完全に包含するものではないが, これら 2 つの項目について齟齬がなければ本実験は致命的な誤解を生じさせることなく遂行できると考え，2 点に絞って理解度の確認を行った. なお, 2 つ目の質問では未成年でないことの確認を行うが, これは自己申告に基づく確認であり未成年を被験者としないことをこのプロセスにより保証することはできない. これらの同意文の送信から, 理解度の確認までを今回の実験におけるインフォームド・コンセントとする.

> (a) "Do you understand that I will not be purchasing any products from you?"
> (b) "Do you confirm that you are not a minor?"

図 2 同意文の理解度を確認するための質問

(3) **リマインド**: 図 1 の同意文を送信後, 1 週間以上返信がない対象に対しては, 図 3 で示すリマインドメッセージを送信する.

(4) **対話の開始**: 同意文および理解度を確認する 2 つの質問すべてに対し肯定的な返信が得られた場合にのみインフォームド・コンセントを取得できたとみなし, 情報収集のための対話を開始する. ここからの対話は先行研究 [7] で用いられている HUMINT エージェント LLM の出力を利用する. なお, HUMINT エージェント LLM が倫理的に問題のある発言がないかを確認する目的でメンロレポート内容を理解した学生 2 名の監視のもと対話を行うものとする. 収集する情報は脅威アクターが投稿内で開示していない商品等に関する情報である. 具体的にどのような情報を収集するかは, インフォームド・コンセントを取得できた場合に限り, 該当ポストを再度確認し, そのポストからは収集できない情報を人手で判別することで決定する.

> "Just a quick follow-up on my request for your help with my experiment. Please let me know if you're interested."

図 3 リマインドメッセージの原文

(5) **対話の終了**: 被験者から拒否するようなメッセージがきた場合や, 被験者によって会話が削除された場合は, 必要に応じて謝罪と感謝を述べた後, 対話を終了した. 未成年であることが判明した場合は, 未成年は実験に参加できないことを伝え, 協力に感謝を述べた上で, 対話を終了した. また, 2025 年 8 月 17 日中に返信がない場合も対話を終了する.

4.3 実施可能性の評価

本手法の実施可能性は, 以下の複数の指標を用いて定量的に評価する.

- **応答率**: 同意文またはリマインドメッセージに対して, 脅威アクターからの応答があった割合.
- **同意率**: 応答があったユーザーのうち, インフォームド・コンセントに同意し, 対話に参加した割合.

対話は全て記録し，返信が届いたにも関わらず, 同意が得られなかった場合は, それぞれのケースでその原因を分析した.

5. 調査結果と考察

本研究では, 提案手法の有効性を検証するため, 以下の項目について実験を行った. 実験は 2025 年 7 月 27 日から 8 月 17 日にかけて実施した.

5.1 インフォームド・コンセントの反応

対象とした 30 件の投稿者に対し, 同意文を送信した結果, 以下の反応が得られた.

- 応答率: 30%（30 件中 9 件）
- 同意率: 0%（30 件中 0 件）

応答した 9 件には, 同意文や質問文に返答のない 20 件に対して行ったリマインド後の返答 4 件も含まれる. なお, 応答があったにも関わらず同意に至らなかった 9 件について, 以下のような反応があった.

- BOT のような対応をされた場合: 2 件 (No.5, No.14)
- 図 2 で示す質問により, 相手が未成年であることが発覚した場合: 1 件 (No.9)
- 一度は承諾したが, 図 2 の質問で改めて商品を買わないことを知らせた場合に返信が返ってこなくなった場合: 2 件 (No.11, No15)
- 時間がないことを理由に断られた場合: 1 件 (No.12)
- 図 2 で示す質問により, 未成年であるかどうかを尋ねる質問に無回答であった場合: 1 件 (No.16)
- 強い侮辱による拒否: 1 件 (No.26)
- メリットがないことの確認: 1 件 (No.30)

5.2 反応の分析

本実験では, 脅威アクターからの同意が得られなかったため, 脅威インテリジェンスの構築を目的とした本格的な対話は実施できなかった. しかし, 同意に至らなかった 9 件の投稿者との間で行われたやり取りでは, 脅威アクターの多様な振る舞いが確認できた. No.1 から No.30 までの各対話における同意文, リマインド, 質問文 (a), 質問文 (b) の送信時刻とそれに対する返信時刻を表 2 にまとめた. 送信時刻, および返信時刻の「-」はそれぞれ送信しなかったこと, 返信がなかったことを示す. また, 結果の「同意文既読のみ」は Tegram の機能により脅威アクターが同意文を確認した（既読になった）ことがわかっているが, それ以外の応答が得られなかったことを示す. また, 「未読」は Telegram 上で同意文すら既読にならなかったことを示す. 「DM 削除済み」は脅威アクターの操作により Telegram 上で DM が確認できなくなっていたことを示す. これら以外の場合は個別に結果を記している.

5.2.1 同意に至らなかった原因の詳細分析

同意に至らなかった 9 件のケースについて, その原因をより詳細に分析した.

- **BOT のような対応をされた場合 (No.5, No.14)**: 2 件の投稿者から返信があり, 1 件は, 多額の利益が出ると見られる仮想通貨のマイニングツールを売る投稿に対してメッセージを送った結果, 1 分以内にサブスクリプションへの加入を促す返信が返ってきたケースである. もう 1 件は, 複数の国に SMS を一括送信するサービスを宣伝する投稿で, 同意文を送っても返信はなかったが, 10 日後にリマインドを送ると BOT から連絡を取るように催促する返信が即座に返ってきた. 後者においては後日確認するとアカウントが削除されていた.
- **未成年であることが発覚した場合 (No.9)**: 投資家を名乗る投稿で特定の通貨の価値が暴騰する旨を示唆する投稿をしていた投稿者に連絡を取った. 同意文送信後, 3 分以内に肯定的な返事が得られたものの, 年齢確認の質問に対し未成年であるとの申告を受けたため, 実験には参加できないことの説明を行い対話を終了した.
- **商品購入の否定による返信途絶 (No.11, No15)**: 2 件の投稿者から返信があり, 1 件はゲーム機のアカウントを販売していることを示唆する投稿であり, もう 1 件は詐欺行為目的のサービスを展開していると見られる投稿であった. 1 回目の同意文の送信に対する反応は見られなかったが, リマインド文に対しては前者は 4 分以内に, 後者は 8 分以内に肯定的な反応を示した. しかし, 商品を購入しない旨を明確にした途端に返信が途絶えた, 前者に関しては DM が削除されていた.
- **時間がないことを理由に拒否 (No.12)**: 大量の迷惑メールや迷惑 SMS の送信を促進するサービスの販売を示唆する脅威アクターに対して連絡を取ったところ, 3 分以内に同意文の送信に対して時間がないことを理由に実験への参加を断られた.
- **年齢確認の質問に無回答 (No.16)**: 他人のアカウントに不正にアクセスするサービスを宣伝していた脅威アクターに対して連絡を取ったところ, 同意文に対しては 237 分以内に, 理解度を確認する 1 つ目の質問に対しては 68 分以内に肯定的な質問を得ることができた. その後 39 時間以内に送信した年齢確認の質問に対しては無回答であった.
- **強い侮辱による拒否 (No.26)**: DDOS サービスを宣伝する投稿をする脅威アクターに対して同意文を送ったところ, 返信がなかった. 後日リマインドのメッセージを送信したところ, 7 分以内に強い侮辱の言葉で参加を拒否したため, 返信に対するお礼と謝罪を述べ, 対話を終了した.
- **メリットがないことの確認 (No.30)**: 電話のなりすまし機能を有するサービスを宣伝する脅威アクターに対して同意文を送ったところ, 9 分以内に実験に参加するメリットの説明を求める返信を得た. 実験に対する利益はなく, 自発的な参加に限られることを説明したところ返信はなかったため, 対話を終了した.

5.3 考察

本実験では, 投稿者から同意を得ることは叶わなかったが, これらの対話記録から Telegram 上における彼らの活動実態や今後の HUMINT 活動における課題が見えてきた. アカウント名に Bot と名がついていないアカウントでも自動応答を利用するケースが確認された. このような相手に HUMINT 活動を行うことはできないが, 今回の実験対象 30 件のうち自動応答と見られるケースは 2 件のみであるため HUMINT への影響は少ないと考えられる. 同意文の後の理解度を問う質問で未成年であるとの申告を受けたため実験を終了したケースがあったため, 今後 HUMINT を行う際には未成年である可能性を考慮した実験デザインを考える必要がある. 商品購入の意思がないことを明確にす

表 2 実験時刻と調査結果

No.	同意文		リマインド		質問文 (a)		質問文 (b)		結果
	送信時刻	返信時刻	送信時刻	返信時刻	送信時刻	返信時刻	送信時刻	返信時刻	
1	7/27 5:48	-	8/10 13:41	-	-	-	-	-	同意文既読のみ
2	7/27 5:51	-	8/10 13:41	-	-	-	-	-	未読
3	7/27 5:54	-	8/10 13:40	-	-	-	-	-	同意文既読のみ
4	7/27 5:58	-	8/10 13:40	-	-	-	-	-	同意文既読のみ
5	7/27 18:40	7/27 18:40	-	-	7/27 18:41	-	-	-	BOT と見られる
6	7/27 18:45	-	8/10 13:38	-	-	-	-	-	同意文既読のみ
7	7/27 18:47	-	8/8 19:37	-	-	-	-	-	同意文既読のみ
8	7/28 18:46	-	-	-	-	-	-	-	DM 削除済み
9	7/28 10:47	7/28 10:49	-	-	7/28 10:50	7/28 10:50	7/28 10:50	7/28 10:53	未成年であるとの申告を得て終了
10	7/28 10:49	-	8/8 19:36	-	-	-	-	-	同意文既読のみ
11	7/30 0:40	-	8/8 19:35	8/8 19:38	8/8 7:33	-	-	-	同意文同意後, DM 削除
12	7/30 0:46	7/30 0:48	-	-	-	-	-	-	時間がない事を理由に拒否
13	7/30 1:00	-	-	-	-	-	-	-	DM 削除済み
14	7/30 2:49	-	8/9 7:28	8/9 7:28	-	-	-	-	BOT と見られる, DM 削除済み
15	7/30 2:57	-	8/8 19:14	8/8 19:21	8/8 19:24	-	-	-	質問文 (a) に未回答
16	7/30 3:03	7/30 6:59	8/8 19:34	-	7/30 7:57	7/30 9:04	7/31 23:32	-	質問文 (b) に未回答
17	7/30 3:05	-	8/8 19:14	-	-	-	-	-	同意文既読のみ
18	7/30 3:18	-	8/8 19:14	-	-	-	-	-	同意文既読のみ
19	7/30 3:21	-	8/8 19:14	-	-	-	-	-	同意文既読のみ
20	7/30 3:21	-	8/8 19:14	-	-	-	-	-	未読
21	7/30 3:27	-	-	-	-	-	-	-	DM 削除済み
22	7/30 3:30	-	-	-	-	-	-	-	DM 削除済み
23	7/30 3:38	-	-	-	-	-	-	-	DM 削除済み
24	7/30 3:43	-	-	-	-	-	-	-	DM 削除済み
25	7/30 3:56	-	8/8 19:30	-	-	-	-	-	DM 削除済み
26	7/30 4:14	-	8/8 19:25	8/8 20:31	-	-	-	-	強い侮辱の言葉で拒否
27	7/30 4:55	-	8/8 19:25	-	-	-	-	-	同意文既読のみ
28	7/30 5:00	-	8/8 19:25	-	-	-	-	-	DM 削除済み
29	7/30 5:14	-	8/8 19:30	-	-	-	-	-	DM 削除済み
30	7/30 5:16	7/30 5:24	-	-	-	-	-	-	被験者の利益の説明を求められた. DM 削除済み

ると返信が途絶える事例から, 脅威アクターは同意文を十分に読んでいない可能性がある. これは, インフォームド・コンセントを取得するためには形式的な同意文を送るだけでなく, 対話を通じて被験者が実験趣旨を正しく理解しているかどうかを丁寧に確認する必要性を示唆している. また, 未成年でないことを確認する段階で協力を拒否されるケースも確認された. これは 2 つ目の質問文の送信が返信を受けてから 39 時間以内とそれまでの送信ペースに比べ, 時間がかかってしまったことや, 複数の質問を重ねることによる被験者に事務的な負担が生じることにより, 実験への参加意欲が削がれてしまったと考えられる. 一方で, メリットがないことを確認するケースからは条件次第では実験への協力に応じる脅威アクターが存在する可能性も示唆された.

以上を踏まえると, Telegram 上の脅威アクターに対し, 研究目的である事を明かした上で会話を行った場合, 脅威アクターは様々な要因により, 情報提供を行わなかった. この結果を踏まえると, HUMINT を行う上で脅威アクターにインフォームド・コンセントを取得することは現実的でないと言える. そのため, 学術的研究として HUMINT を行うためには, 脅威アクターに対して欺瞞を含む実験方法を検討する必要がある. ただし, 今回の調査に対しても否定的な反応を示した脅威アクターは多数存在したため, 報復などのリスクの大小だけでなく, 人権を尊重するという観点からも, 被験者への配慮は必要になってくるだろう. そのため, 今後の HUMINT の研究は, メンロレポートや過去の犯罪学, 民族誌における秘匿型研究などを参考に, 脅威インテリジェンス構築のための情報収集と被験者への配慮のバランスをとった研究となることが期待される.

6. 倫理的配慮

本研究は, 研究対象者である Telegram ユーザー, および研究者自身をステークホルダーとして捉え, それぞれの人権尊重と潜在的リスクへの対策を講じた.

研究対象者への配慮

Telegram ユーザーに対しては, 会話の目的を明確に伝え, その意図を十分に理解していることを確認した上で情報収集を行った. インフォームド, コンセントの取得後, 社会全体の被害防止に有益な情報が得られた場合, その通知の要否は都度議論により決定する. また，対象者の特定に繋がる具体的な情報は掲載せず，研究データは安全に保管し，外部に提供しないことで対象者を保護し，その旨も対象者に伝えた.

研究者自身への配慮

悪意ある脅威アクターからの報復の可能性を考慮し, 本実験は仮想環境上で実行することで, 研究者所有機器への損害リスクを排除した. また, 身元特定を防ぐため, 研究者は偽名を用いて Telegram サービスを利用し, 会話中に個人的な詳細情報が特定されることがないよう細心の注意を払った.

7. まとめと今後の課題

本研究は, メンロレポートに準拠したインフォームド・コンセントに基づく HUMINT の実施可能性を検証することを目的とした. 実験の結果, 脅威アクターとの接触は可能であるものの, 同意率が極めて低いことが明らかになった. このことから, HUMINT の実施においてはインフォームド・コンセントを取得する既存の手法には限界があることがわかった. したがって, 今後 HUMINT に関する実験を遂行するためには, 欺瞞を含む実験を行うことも検討する必要がある. ただし, 今回の研究内容に対しても強い不快感を感じる脅威アクターがいることを確認しているため, 欺瞞を伴う研究には被験者に対する配慮を欠かさず, 慎重に行うことが求められるだろう.

謝辞 本研究の一部は, ＮＥＤＯ（国立研究開発法人新エネルギー・産業技術総合開発機構）の委託事業「経済安全保障重要技術育成プログラム／先進的サイバー防御機能・分析能力強化」（JPNP24003）によるものである. また, 松村 尚典氏, 青砥 陸氏, 金子 翔威氏のご協力に感謝申し上げる.

参考文献

[1] Bank Security. Cyber intelligence: Humint operations. https://bank-security.medium.com/cyber-intelligence-humint-operations-2d3d526e4007, 2021. Accessed: 2025-08-17.

[2] The Hacker News. Humint: Diving deep into the dark web. https://thehackernews.com/2024/07/humint-diving-deep-into-dark-web.html, July 2024. Accessed: 2025-08-17.

[3] Tomas Domine and Yehonatan Wiesel. Why and how to perform telegram monitoring to protect your business. https://cyberint.com/blog/dark-web/why-and-how-to-perform-telegram-monitoring-to-protect-your-business/, July 2024. Accessed: 2025-08-17.

[4] Information technology Promotion Agency (IPA). 令和 5 年度「高度 it 人材育成のための実践的演習プログラム」成果報告書. https://www.ipa.go.jp/jinzai/ics/core_human_resource/final_project/2024/f55m8k0000003510-att/f55m8k000000358r.pdf, 2024. Accessed: 2025-08-17.

[5] Group-IB. Vulnerability intelligence (vulint). https://www.group-ib.com/resources/knowledge-hub/vulnerability-intelligence/, 2025. Accessed: 2025-08-17.

[6] The menlo report: Ethical principles guiding information and communication technology research. https://www.dhs.gov/sites/default/files/publications/CSD-MenloPrinciplesCORE-20120803_1.pdf, August 2012. Accessed: 2025-08-17.

[7] 鈴木涼介, 川口大和, インミンパパ, 山岡裕明, 吉岡克成. サイバー攻撃者とのテキストベース対話による情報収集フレームワーク ～法と研究倫理への配慮と llm 活用～. 電子情報通信学会技術研究報告, Vol. 124, No. 422, pp. 391–398, mar 2025.

[8] Alice Hutchings and Thomas J. Holt. Interviewing cybercrime offenders. JQCJC, Vol. 7, No. 1, pp. 75–94, fall 2018.

[9] Deirdre Healy. Ethics and criminological research: Charting a way forward. IRISH PROBATION JOURNAL, Vol. 6, pp. 171–179, September 2009.

[10] David Calvey. The art and politics of covert research: Doing 'situated ethics' in the field. Sociology, Vol. 42, No. 5, pp. 905–918, october 2008.

[11] David Calvey. Covert ethnography in criminology: A submerged yet creative tradition. Current Issues in Criminal Justice, Vol. 25, No. 1, pp. 541–550, July 2013.

[12] 青砥陸, インミンパパ, 吉岡克成. 類似チャンネル提示機能を活用した telegram におけるサイバー犯罪関連チャンネルの発見と分析. 電子情報通信学会技術研究報告, Vol. 124, No. 422, pp. 399–406, mar 2025.