修士論文

心理的要因を再現したフィッシングメール判別訓練アプリケーションの開発

山野 宗馬

主指導教員　　長谷川 忍

北陸先端科学技術大学院大学
先端科学技術研究科
（情報科学）

令和 8 年 3 月

Abstract

In modern society, where the Internet is widely used for activities ranging from individual citizens to corporations and public services, "phishing"—a technique that impersonates legitimate services or companies by sending fraudulent emails to lure users to fake websites and steal important personal information—has been rapidly increasing and poses a serious threat. The sophistication of attack methods through the misuse of advanced technologies such as generative AI, along with the anticipated further increase in attacks, suggests that phishing will remain a major threat in the future.

To prevent phishing damage, it is desirable for users themselves to acquire the ability to identify phishing emails. In fact, services that provide simulated phishing email training using mail servers exist and are widely used in organizations such as corporations. However, implementing such countermeasures requires financial costs and dedicated personnel, creating a burden for organizations. Additionally, while phishing also targets individual users, it is difficult for individuals to utilize full-scale phishing countermeasures, and educational opportunities are limited.

Furthermore, based on findings from related research, countermeasures against phishing attacks require not only technical knowledge effective for identifying phishing emails but also consideration of the direct psychological impact that phishing emails have on users, as well as changes in mental states and attention levels stemming from users' own environments and characteristics. Many traditional training approaches focus primarily on technical identification knowledge and do not emphasize the reproduction of psychological factors. In this research, factors that influence user psychology and affect phishing email identification are defined as "psychological factors," and their importance is emphasized.

Given this background, effective phishing countermeasure education requires the ability to respond to the rapid sophistication and changes in attack methods with sustainable operation, consideration of psychological effects on users beyond technical knowledge and abilities, and the realization of these elements at low cost with minimal need for specialized personnel or equipment.

This research aims to provide more effective and convenient learning methods than conventional phishing countermeasure education by developing a phishing email identification training application that automatically generates training content using generative AI and constructs a training environment that reproduces the psychological factors of actual phishing emails. The system aims to verify the extent to which it contributes to improving users' phishing email identification abilities.

During training, users are assigned the role of students at a fictional educational institution and respond to emails based on a virtual student life scenario. This scenario sets a period of approximately one year, with dates progressing at a different speed from reality. Throughout this period, users receive various emails from multiple types of senders. The received emails may include important tasks in student or private life, such as course registration, job hunting, and banking transactions, and users need to respond appropriately to these tasks. Incorrect judgments or delayed responses lead to declining grades and may ultimately result in academic failure. On the other hand, phishing emails are also included among the received emails, and responding to them results in penalties. Therefore, users are required to accurately distinguish between legitimate emails and phishing emails and select appropriate actions.

In this application, to effectively reproduce the psychological factors that lead to phishing damage, these factors are classified into three categories: personal factors, environmental factors, and external stimulus factors. Personal factors include elements such as personality traits, past experiences, and normalcy bias. This system minimizes the influence of personal factors in training by providing a unified virtual scenario environment to all users. Environmental factors include business relevance and time constraints. Business relevance is reproduced by clearly setting the user's role and relationships with senders in the virtual scenario. Time constraints are realized by introducing a mechanism where dates progress at regular intervals within the system and setting time limits for tasks. External stimulus factors include trust in senders and emotional stimuli. This system realizes that all users have an equal foundation of trust toward senders in the virtual environment through a design premised on immersion in the scenario and the establishment of a completely fictional sender system. For emotional stimuli, the system enables generation of email content with arbitrary emotional characteristics by specifying the nature of the text through prompts when automatically generating content using generative AI.

In technical implementation, this system is implemented as a pure client-side application using HTML, CSS, and JavaScript, except for communication functions with the LLM API. This design requires no special server environment and can operate on ordinary web browsers. The Claude 3.5 Haiku API from Anthropic is adopted for content generation LLM. In actual phishing attacks, the process includes users clicking links from email clients and transitioning to external sites. To reproduce this experience, this system adopts a hierarchical iframe structure. The index.html at the top-level functions as a control layer managing the overall training progress, with an iframe element placed within it embedding the email client screen. When users click links in emails, the reference

destination of this iframe dynamically switches to HTML files for external site pages, reproducing screen transitions. This design simulates page transitions in actual web browsing, providing learners with realistic judgment situations while maintaining the internal state of the control layer even when pages within the iframe are switched.

The email-scenarios used in training are structured by first defining a list of entities that send emails as "sender types." This system establishes five sender types: university student affairs office, professor, bank, job hunting site, and EC site. Scenarios belonging to each sender type include both normal emails and phishing emails, with a total of 25 emails delivered during the training period. Since phishing emails may be sent mixed in with normal email exchanges in attacks targeting organizations, the design enables a narrative email progression structure composed of multiple "phases." Each email is assigned a parameter indicating identification difficulty. The difficulty parameter is set at four levels: normal (legitimate email), phishing_easy, phishing_medium, and phishing_hard. Through this gradual difficulty setting, the system aims to have users learn basic identification clues from low-difficulty phishing emails initially and develop practical identification abilities by progressing to high-difficulty phishing emails.

This system defines actions required of users for each email as "tasks" and adjusts scores according to their completion status. In actual work, missing important emails or failing to respond within deadlines results in disadvantages. By reproducing this as score penalties, psychological pressure similar to actual email response situations arises. To reproduce a realistic email response environment, training progresses within a virtual time axis. In this virtual time axis, one day passes in 4 seconds of real time, progressing through approximately one year's scenario in about 24 minutes.

Individual email generation is implemented as a multi-stage process that sequentially generates five elements: sender name, email address, subject line, body text, and link URL. For phishing emails, sender names, email addresses, and sending server domains are newly generated, while for legitimate emails, previously generated official sender information is directly reflected. Separate from standard email display elements, email sources and authentication results for various authentication technologies are generated as advanced identification knowledge. Email sources include standard headers plus an Authentication-Results header showing authentication results for three types of email authentication: SPF, DKIM, and DMARC. Authentication results directly reflect those defined in difficulty parameters. In phishing_hard, authentication is intentionally set to pass completely, providing learners with the important lesson that "spoofing is possible even when authentication results are favorable."

This system records all user interactions during training as time-series events. Major

recorded events include email reception, email opening, email hiding, email source viewing, link clicks, action completion on external sites, and phishing reports. These logs are designed to be obtained as JSON files at the end of training and can be analyzed afterward to clarify user behavior trends during training.

To verify the educational effectiveness of the developed phishing email identification training application, a subject experiment was conducted. Since the core design element regarding educational effectiveness of this application is the reproduction of psychological factors in actual phishing attacks, evaluation of the impact of this reproduction of psychological factors on user learning effects was set as the most important verification item. To achieve this purpose, this research adopted a between-subjects experimental design. Specifically, subjects were divided into two groups: the experimental group used an application containing psychological factors, while the control group used an application with psychological factor reproduction elements removed.

Experiment participants were recruited from students of our university who had no experience receiving specialized education or training regarding phishing, with a total of 10 participants gathered. The primary difference between the experimental and control groups is the presence or absence of gamification elements related to the reproduction of psychological factors. The experimental group's application incorporates elements such as setting relationships between senders and users, tasks and scoring, time constraints, and temporal concentration of email delivery through scenario design. In contrast, the control group's application removed all these elements, adopting a simple sequential identification format where one email is presented, and after the subject judges whether it is a phishing email, the next email is presented.

For both experimental and control groups, descriptive format tests were conducted before and after training using the application. Both tests presented subjects with samples of pseudo-phishing emails created using the application's email generation function, requiring them to point out suspicious elements within the emails, judge whether they were phishing emails, and describe the reasoning behind their judgments. Regarding responses to pre- and post-tests, qualitative analysis of their content was conducted by implementing a coding process to classify them into three levels. The criteria for each level are: Level 1 for vague indications, Level 2 for specific indications, and Level 3 for explanatory indications.

Response levels in pre- and post-tests were treated as scores for statistical analysis. For Email 1, since the same email was presented in both tests, score fluctuations could be compared. For the experimental and control groups, the Mann-Whitney U test was

performed on score change values between pre- and post-tests to verify whether there was a significant difference in response quality changes between the two groups. Results showed no significant difference at the 5% level. Therefore, within the scope of this analysis, it could not be confirmed that reproduction of psychological factors was effective in the learning effects of using the training application. However, since no score increases were observed across all subjects, it means that learning effects from the application itself could not be confirmed for either group.

A post-training questionnaire was conducted to investigate subjects' subjective impressions regarding training with the application and its effects. Regarding anxiety and tension at the pre-test stage, all subjects gave negative responses, judging that the task itself had no significant psychological effect. Regarding anxiety and tension during training, the experimental group showed a slightly more positive tendency, but no large difference was observed. On the other hand, regarding whether email handling was difficult, the experimental group showed a more positive tendency, with the largest difference seen among all questions. It is considered that temporal concentration of email delivery through scenario design partially demonstrated effectiveness. Regarding the sense of ability improvement through training, the control group showed a more positive tendency. In the experimental group, due to the existence of scoring and time limits, consciousness may have been concentrated on email handling itself, possibly failing to provide a sense of learning phishing email characteristics.

User behavior trends were analyzed from operational log information. As a comparable behavioral tendency between the two groups, "how much information was referenced before judging whether an email was phishing" was used. Specifically, the "rate of checking other emails" and "rate of checking email sources" before making judgments after opening emails were tallied. From log tabulation results, it was found that both other email checking rates and email source checking rates showed large individual differences, with no consistent differences found between the two groups. However, it was found that many subjects had higher email source checking rates compared to other email checking rates. While email source display functionality is available in general email clients, considering pre-questionnaire results, it is unlikely that subjects used it routinely, so the training situation is considered to have triggered active source checking behavior.

Analysis of email identification results showed that the control group had higher correct answer rates in email identification. Since the control group had no time limits, no errors from non-response occurred, and subjects could take time for judgments, this result was predictable. Differences observed between groups include the "rate of judging as normal but being incorrect." All subjects in the experimental group had at least one instance of

this pattern, meaning they responded to tasks without detecting phishing emails. Since time constraints and penalties for exceeding deadlines existed in the experimental group, it is quite conceivable that psychology urging quick task responses may have worked.

This research focused on the importance of psychological factors in phishing email identification, developed a simulated experience-type training application reproducing these factors, and verified its educational effectiveness through experiments. Psychological factors were classified into personal, environmental, and external stimulus factors, with designs to effectively reproduce each within the training environment. For content generation, an automatic email content generation mechanism using LLM API was constructed, adopting a hierarchical email scenario design based on sender types. Experimental results showed no statistically significant difference between experimental and control groups in response quality changes in pre- and post-tests. However, post-training questionnaires showed that the experimental group responded more positively to "email handling being difficult" than the control group, suggesting that temporal concentration of email delivery through scenario design demonstrated certain effectiveness. Additionally, log analysis confirmed that many subjects actively checked email sources, indicating that the training environment had effects promoting detailed analytical behavior.

目次

図目次

表目次

第1章 はじめに

1.1 研究背景

市民から企業・公共サービスに至るまで，あらゆる人々の活動にインターネットが広く利用される現代社会において，インターネットを悪用した犯罪活動であるサイバー犯罪も広がりを見せている．中でも，実在のサービスや企業を装い，偽装した電子メール等を送り付けて偽の Web サイトに誘導し，重要な個人情報を窃取する「フィッシング」とよばれる手口[1]は急激に増加しており，近年のインターネット上の脅威情勢は深刻である．フィッシングは攻撃対象となるユーザが膨大であることに加え，生成 AI をはじめとする高度な技術を悪用した手口の高度化や，攻撃のさらなる増加が懸念されており，今後も大きな脅威となることが予想される[2]．

フィッシングの被害を防止するためには，ユーザ自身がフィッシングメールの判別能力を獲得することが望ましい．実際に，メールサーバを利用した模擬フィッシングメール訓練を提供するサービスが存在し，このような対策は企業等の組織において広く利用されている[3,4]．一例として，国内トップシェアであるとする GSX の標的型メールサービスは，11000 社以上の導入実績を誇る[5]．しかし，こうした対策手段の実施には金銭的コストや担当人員が必要であり，組織にとって負担となる．また，フィッシングは個人ユーザも攻撃対象となるが，本格的なフィッシング対策を個人で利用するのは困難であり，教育機会が限定的であると考えられる．

加えて，後述する関連研究が示す事実から，フィッシング攻撃への対策にはフィッシングメールの判別に有効な技術的知識だけでなく，フィッシングメールがユーザの心理に与える直接的な影響や，ユーザ自身の環境・特性に由来する心理状態や注意力の変化も考慮する必要があると考えられる．本研究では，ユーザの心理に働きかけ，フィッシングメールの判別に影響を与える要因を「心理的要因」と定義し，その重要性に着目した．

以上のような背景から，有効なフィッシング対策教育には，手口の急速な高度化・変化に対応でき，継続的に運用できること，技術的な知識・能力に留まらず，ユーザへの心理学的な作用を考慮していること，そしてこれらを低コストで，専用の人材や機材を極力必要とせずに実現できることが求められる．

1.2 研究目的

本研究では，従来のフィッシング対策教育よりも効果的かつ利便性の高い学習手段を提供することを目的として，独自のフィッシングメール判別訓練アプリケーションを開発する．具体的には，生成 AI を用いて訓練コンテンツを自動生成し，実際のフィッシングメールの心理的要因を再現した訓練環境を構築することで，ユーザのフィッシングメール判別能力の向上にどの程度寄与するかを検証する．

1.3 論文構成

本論文の構成を以下に示す．

第1章 はじめに

本研究の背景と目的について述べる．

第2章 関連研究

本研究の関連研究と，本研究の位置づけについて述べる．

第3章 提案手法

本研究で用いた手法について述べる．

第4章 開発システム

本研究で開発したアプリケーションシステムについて詳述する．

第5章 評価実験

本研究で実施した実験と，その結果および分析を示す．

第6章 おわりに

本研究のまとめと課題，今後の展望について述べる．

第2章 関連研究

2.1 疑似体験型攻撃メール訓練

本節では，疑似体験型の攻撃メール訓練システムを開発した関連研究について述べる．

内山らの研究では，過去に信州大学の学内メールサービス上において，選定された訓練対象者に対して疑似的な標的型攻撃メールを送付するという訓練を実施したことが述べられている[6]．この訓練においては，実際に利用されているメールサービス上で疑似的な攻撃メールを送付したため，業務への影響や訓練対象者への心理的負担が課題となっていた．また，このような性質から実施に際して役員クラスの承認を得る必要があったようで，容易に実施できる訓練ではないことが分かる．また，対象者に周知されずに実施される訓練であるため，一度の訓練で 1 回しか学習機会を設けることができず，開封を避けた対象者にはそもそも注意喚起が届かないという課題を残した．

これらの課題を解決するため，内山らは疑似体験型攻撃メール訓練を開発している．これは Web ベースのセキュリティ訓練基盤であり，疑似空間での訓練であるため大学外への影響や訓練対象者の業務に影響せず，不安も軽減される．また，学習機会の喪失も避けられる．システム設計においては，訓練で使用されるサイバー攻撃のシナリオを容易に増やせるよう，基盤部分とシナリオ部分を分割する構成を採用している．

内山らは本訓練の実施とアンケートの結果から，訓練対象者の情報セキュリティに対する当事者意識を高めるという目標に対し，疑似体験型の訓練は有効であったと評価した．ただしこの訓練システムは，あくまでも対象者にサイバー攻撃の被害を疑似体験させることに主眼を置いており，必ず被害を受ける（誤った操作をする）よう促すシナリオを用いている．つまり，攻撃メールを見破り，回避するという選択肢がなく，疑似空間である以上そうすべき理由も存在しない．そのため，確かに心理的負担は少ないが，同時に現実の業務環境で攻撃メールに対応する場合の実務上の負担や制約，心理的圧力も再現できていないと考えられる．また，複数のサーバを持ち，学内認証基盤と連携したシステムとして開発されており，あくまでも学内において限られた人員による開発・運用が行われるに留まるものであると考えられる．

2.2 個人適応型フィッシングメール訓練

本節では，学習者個人の知識レベルや特性に適応するフィッシングメール訓練システムを開発した関連研究について述べる．

東野らの研究では，サイバーセキュリティ教育における対象者の知識の個人差と，組織的な研修の実施が業務負担となる課題を解決すべく，知識グラフを用いた個人適応型のサイバーセキュリティ学習システムを開発している[7]．東野らは，サイバー攻撃の中でもフィッシングに着目し，フィッシングの判別方法や攻撃手法に関する知識グラフを構築した．これを用いて，学習者の習熟度を評価し，学習項目の難易度と関連する基礎知識の関係から，各学習項目に関する出題順序を決定するシステムを開発した．

程らの研究では，フィッシングメールの判別スキルに対するトレーニング対象として，メールの件名やアドレス，添付ファイルやリンクといった項目に細分化している[8]．また，実際のフィッシングメールやスパムメール，通常のメールを収集し，トレーニング項目ごとに性質に関するラベルを付与した．これらを用いて，トレーニング課題の難易度を決定する関数を定義し，これに基づいて任意の難易度の課題メールを作成可能なシステムを開発することで，学習者の理解状況に応じた課題を出題できるようにした．

これらの研究で扱われた手法は，明確に学習者個人のフィッシングメール判別能力に着目した問題形式の訓練であり，正誤の概念が存在する．また，学習者間の個人差にも適応できるものである．その点において，これらの手法は実践的なスキルの獲得が期待できるものの，いずれの手法も技術的な判別知識が中心となっており，心理的要因の再現には注力していないと考えられる．また，システム構築において専門的な情報と技術を用いており，専門知識を持たない人には容易に変更や拡張ができる物ではないと考えられる．

2.3 生成 AI を用いたフィッシングメール訓練

本節では，生成 AI を用いたフィッシングメール対応訓練システムを開発した研究について述べる．

東野の研究では，生成 AI 技術を悪用したフィッシング攻撃が出現するとの懸念から，これに対応するために高頻度かつ継続的な訓練が必要であるとしている[9]．しかし，そのような訓練が組織の業務を圧迫する負担となる点を考慮し，訓練対象者が自習可能な訓練を実現すべく，訓練用のコンテンツを生成 AI に自動生成させるシステムを提案している．

東野の開発したシステムでは，メールだけでなくフィッシング攻撃で用いら

れる偽サイトも生成 AI で生成させる仕組みを採用している．また，訓練対象者には任意のメールクライアントを利用させ，実際のメールサーバを通して生成されたメールを送信している．これらの仕組みを用いて，訓練対象者と訓練システムの間でメールのやり取りを行い，生成 AI が対象者とのやり取りの状況を見てフィッシング攻撃を行うという訓練フローとなっている．

このシステムは不具合なく動作したようだが，訓練の進行を生成 AI に委ねた結果として，生成 AI が不適切なタイミングでメールのやり取りを終了してしまう場合があったと報告されている．やり取りに基づいた動的な攻撃という手法から，他の研究と比較してもフィッシングメールの心理的要因が再現されている側面があると言えるが，生成 AI に特有の問題が発生したとも言える．このことから，フィッシング訓練における生成 AI の適用範囲は慎重に選択する必要があると考えられる．ただし，このシステムでは訓練実施時に監督者が訓練のシナリオ概要を入力できるようになっており，これは訓練を構成する膨大なコンテンツや高度な技術を自前で用意する必要がないという，生成 AI 活用の重要な利点と言える．

2.4 フィッシングメール判別における心理的要因

本節では，フィッシングメール判別，及び攻撃の手口における，心理的要因について言及した研究について述べる．

稲葉の研究では，膨大なフィッシング対策の取り組みにもかかわらず攻撃数に変化がないこと，先行研究の多くがフィッシングメールを判別するための知識の習得を目指した取り組みであり，その有効性を支持しない報告が存在することに触れている[10]．その上で，多くのユーザは情報処理の負担や時間等の現実的な制約から，メールの信頼性判断を無意識的な処理に依存していると予測しており，このような「ヒューリスティックス」による判断に影響を及ぼすフィッシングメールの特徴を心理学的な観点から考察している．稲葉によれば，このような影響を及ぼす要因は次のようなものである．

1. 既知のフィッシングメールとの形態的一致性
2. 感情的要素
3. 言語的間違い
4. ユーザとの関連性を認識させる要因
5. 送信元の知名度や規模に関する要因

このうち，1 と 3 はユーザが対象のメールをフィッシングメールだと判断することに繋がる要素であり，信頼性を低下させる要素である．したがって，被害に繋がる要因となるのは 2， 4， 5 であり，これらはユーザに誤った信頼性判断を促

す効果を持つ.

感情的要素は,性格や一時的気分等,ユーザの個人的な特性に由来するものもあるが,メールの内容にも含まれる.例えば,メールの指示に従えばユーザが利益を得られるとする内容や,逆にユーザにとっての損失を仄めかすような内容は感情を喚起し,慎重に考える機会を奪う.

ユーザとの関連性を認識させる要因は,メールがあるサービスのアカウントに関する内容を含んでいたり,ユーザ自身の情報(名前等)を含んでいたりする等,ユーザ個人を特定して送られてきたと認識させるような要素である.後者は特にスピアフィッシングと呼ばれ,このような要因は標的を絞った攻撃において見られるものである.

送信元の知名度や規模に関する要因は,フィッシングメールの送信者の偽装において見られるものである.多くのフィッシングメールでは,送信者を有名で大手の企業やサービスであるように偽装するが,これは利用者が多く疑われにくいからというだけでなく,聞いたことのある名前である方がユーザの信頼を得やすいからでもあると推測している.

稲葉は,フィッシングメールに対するヒューリスティックスを用いた判断には限界があるものの,それでもヒューリスティックスによる正しい判断を促す対策も重要であると述べている.ユーザが全てのメールについて分析的に判断することが現実的に難しいのであれば,フィッシングメールに対する「怪しい」という感覚は重要であるとし,この感覚の生起にヒューリスティックスが寄与すると推察している.

Kavvadias らの研究,John の研究においても,フィッシングの被害を受ける要因として,心理的な要因を挙げている[11,12].ユーザの心理を操作したり,心理的な脆弱性を意図的に利用したりする手法はソーシャルエンジニアリングとよばれており,Workman の研究ではこれが情報セキュリティ侵害の主要な手段だとしている[13].Kavvadias らの研究では,前述した稲葉の研究で挙げられていた要因に類するものの他に,時間的制約や意識の分散について言及している.これは,端的に言えば忙しい状況下ではメールに対する注意力が低下することを意味しており,フィッシングメール自体の要素ではないが,現実の環境下ではフィッシング被害に繋がり得る要因と言える.

Monteith らの研究では,新型コロナウイルス感染症のパンデミックによって世界的にオンラインでの活動が増加し,十分に訓練を受けていないユーザが大量に発生した結果,サイバー犯罪の標的になるという懸念を示している[14].また,パンデミックの結果として精神的に不安定となる人々が発生することにも言及しており,このような心理的な脆弱性はフィッシングをはじめとするサイバー犯罪の手口に利用されるとしている.

2.5 本研究の位置づけ

関連研究と比較した，本研究の位置づけについて述べる．

上述した関連研究は，いずれも仮想的なメールを用いてユーザ自身にフィッシングメールの判別あるいはその被害を疑似体験させることで訓練を行うシステムの開発を試みており，この点については本研究も共通している．

しかしながら，関連研究における課題として，訓練内容がフィッシングメール判別のための技術的知識の習得を中心としたものであり，前章で論じた心理的要因の再現に取り組んだものが少ない点が挙げられる．そこで本研究では，心理的要因を明確に定義した上で，疑似体験型訓練にこれを再現する要素を実装し，さらにその効果を実証的に検証することを新たな取り組みとして位置づける．

加えて，関連研究のもう一つの課題として，汎用性や利便性が限定的である点が指摘できる．第一に，実際のメールサーバ等との連携が前提となり，訓練システム自体も独自のサーバシステムを要するものは，実施規模や実施時期の柔軟性，外部での運用において制約が大きく，広く普及させられるような汎用性に欠ける．第二に，訓練内容や課題の設計・作成が人手に大きく依存するものは，専門的な技術や知識と多大な開発コストを要するため，拡張性に欠ける．これは変化が激しいフィッシング攻撃への対策として大きな欠点となる．第三に，生成 AI を利用した研究は上記の欠点を解決できる可能性を示したものの，訓練内容のコントロールや一貫性の維持には難があったため，生成 AI の適用範囲を適切に設計する必要がある．

これらの課題に対応するため，本研究ではアプリケーション本体を全てクライアントサイドで動作可能な設計とし，Web ブラウザのみで利用可能な環境を実現する．さらに，生成 AI を用いたコンテンツ生成を導入しつつも，その適用範囲を限定し，さらに訓練全体の内容を簡潔な記述で設計可能とすることで，訓練コンテンツの安定した生成と一貫性の維持を両立させる．

以上を基本方針として，次章では本研究における具体的な提案手法について述べる．

第3章 提案手法

3.1 概要

本研究では，生成 AI によるコンテンツの自動生成と，ゲーミフィケーション要素を導入したシナリオ駆動型訓練を用いることで，運用利便性とフィッシングメール特有の心理的要因の再現を両立した訓練アプリケーションを開発する．アプリケーションのシステム概要を図 3.1 に示す．

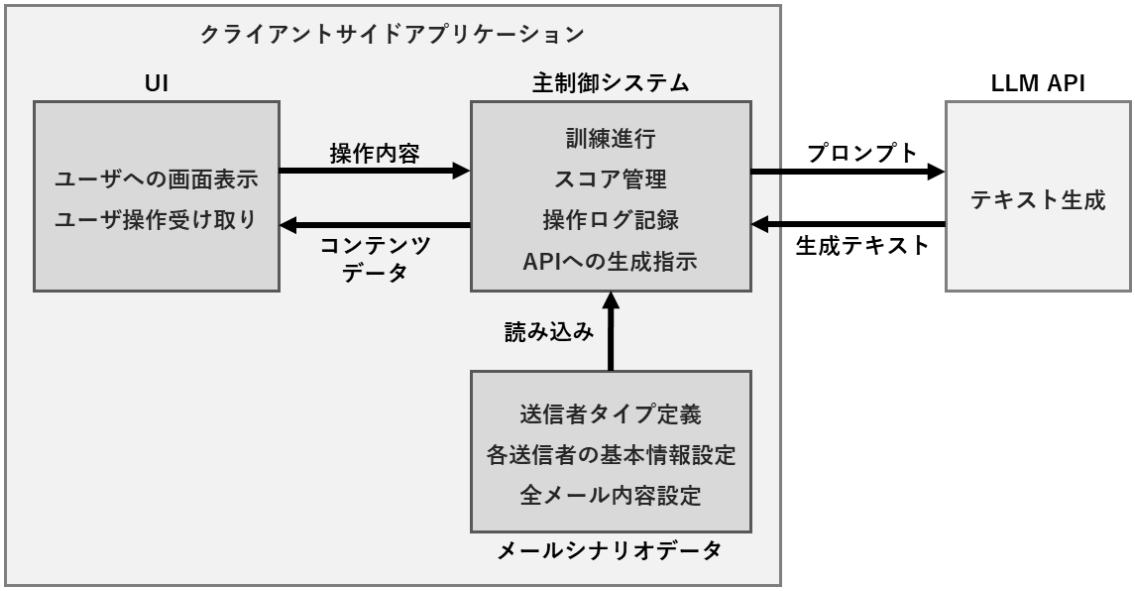


図 3.1 システム概要図

訓練において，ユーザは架空の教育機関に在籍する学生という役割を与えられ，仮想的な学生生活シナリオに基づいてメールに対応する．このシナリオではおおよそ 1 年間の期間を設定しており，現実とは異なる速度で日付が進行する．この期間を通じて，ユーザは複数種類の送信者から様々なメールを受信することになる．

受信するメールには，履修登録，就職活動，銀行取引といった，学生生活や私生活において重要なタスクが付随する場合がある．ユーザはこれらのタスクに適切に対応する必要があり，誤った判断や対応の遅延は成績の低下を招き，最終的には落第という結果に至る可能性がある．一方で，受信メールの中にはフィッシングメールも含まれており，これに対応してしまうとかえってペナルティを受けることになる．したがって，ユーザは正規のメールとフィッシングメールを正確に判別し，適切な行動を選択することが求められる．

アプリケーションのシステムは，タスクへの対応の成否や判断の正確性に応じてスコアリングを行う．訓練の目標は，スコアを一定以上に維持しながら落第を回避し，仮想的な学生生活を無事に乗り越えることである．この目標設定により，ユーザは緊張感を持ちながら，必然的に全てのメールに注意しなければならない設計となっている．

また，本アプリケーションの特徴として，訓練内で届くメールを構成するコンテンツ，すなわち送信者名，メールアドレス，本文等を，生成 AI の API を用いて自動的に生成する機能を有している．この仕組みにより，訓練試行ごとに異なる内容のメールを効率的に提供することが可能となり，繰り返し訓練による学習効果の低下を防ぐ．

3.2 心理的要因の再現

本研究では，訓練アプリケーション内でフィッシング被害につながる心理的要因を効果的に再現するため，まずこれらの要因を個人要因，環境要因，外部刺激要因の 3 つに分類し，それぞれを定義した．

個人要因には，性格特性，過去の経験，正常性バイアスといった要素が含まれる．これらには当然ながら個人差があり，全ての訓練参加者に対してこれを完全に均一化することは不可能と考えられる．そこで本システムでは，全てのユーザに対して統一された仮想シナリオ環境を提供することで，訓練における個人要因の影響を最小限とすることを図った．

環境要因には，業務関連性や時間的制約が挙げられる．業務関連性については，仮想シナリオ上でユーザの役割や送信者との関係性を明確に設定することで再現した．また，時間的制約については，システム内で一定時間ごとに日付が進行する仕組みを導入し，タスクに対する時間制限を設けることで実現した．これにより，現実の状況下で生じる時間的プレッシャーを訓練環境内で体験させることが可能となる．

外部刺激要因には，送信者に対する信頼感や感情的刺激が含まれる．信頼感は個人の経験に依存する部分があり，例えば特定のサービスの利用経験の有無によって，そのサービスを名乗る送信者からのメールに対する信頼度は大きく異なる．本システムでは，シナリオへの没入を前提とした設計と，完全に架空の送信者体系を設定することにより，全てのユーザが仮想環境内の送信者に対して同等の信頼感の基盤を持つことを実現した．感情的刺激については，不安を煽る内容やユーザにとっての利益を提示する内容など，多様な心理的影響を与える要素を指す．本システムでは，生成 AI を用いたコンテンツの自動生成時に，プロンプトを通じて文面の性質を指定することで，任意の感情的特性を持ったメ

ール内容を生成可能とした.

3.3 コンテンツ自動生成

訓練を構成するために必要な膨大なメールコンテンツを手動で作成することは，時間的・労力的に困難である．また，フィッシングの手口は短期間のうちに変化・高度化するため，コンテンツの更新が容易に行える仕組みを持つことが望ましい．これらの課題を解決するため，本システムではコンテンツの自動生成機能を実装することとした．本システムでは訓練に仮想のシナリオを適用する設計であるため，事前にメールの送信者，大まかな内容，送信時期などをセットにした「メールシナリオ」データを用意し，これを基に送信者アドレスやメールコンテンツなどを自動生成する方式を採用した．この方式により，シナリオの枠組みを保持しながら，具体的なコンテンツを柔軟に生成することが可能となる．

第4章 開発システム

本章では，開発したフィッシング対応訓練アプリケーションの技術的実装について述べる．

4.1 システムアーキテクチャ

本節では，システム全体のアーキテクチャ設計について述べる．階層的な画面構造，技術基盤の選定理由，およびコンポーネント間の通信方式が，訓練の実環境再現性とデータ収集の正確性を支える基盤となっている．

4.1.1 技術基盤と設計方針

本システムは，LLM API との通信機能を除いて，HTML，CSS，JavaScript を使用した純粋なクライアントサイドアプリケーションとして実装した．この設計により，特別なサーバ環境を必要とせず，一般的な Web ブラウザ上で動作可能である．

コンテンツ生成用の LLM には，Anthropic 社の Claude 3.5 Haiku API[15]を採用した．API へのアクセスは，Node.js[16]ベースのローカルプロキシサーバーを介して行う構成とした．プロキシサーバーは，API キーのクライアントサイドへの露出を防ぎ，同時にクロスオリジンリソース共有（CORS）制約を回避する役割を担う．

4.1.2 iframe 構造による画面管理

実際のフィッシング攻撃では，ユーザがメールクライアントからリンクをクリックし，外部サイトに遷移する過程が含まれる．この体験を再現するため，本システムは階層的な iframe 構造を採用している．

最上位に位置する index.html は，これに読み込まれるスクリプトと一体となり，訓練全体の進行制御を担当する制御層として機能する．この index.html の中に，ほぼ全画面を占有する形で iframe 要素を配置し，訓練中は主にメールクライアント画面（InfoSecurity.html）が埋め込まれる．

ユーザがメール内のリンクをクリックする等の操作を行うと，この iframe の参照先が変更され，外部サイトページ用の HTML ファイルに動的に切り替わることで画面遷移を再現している．

これらの画面管理システムの概要を図 4.1 に示す.

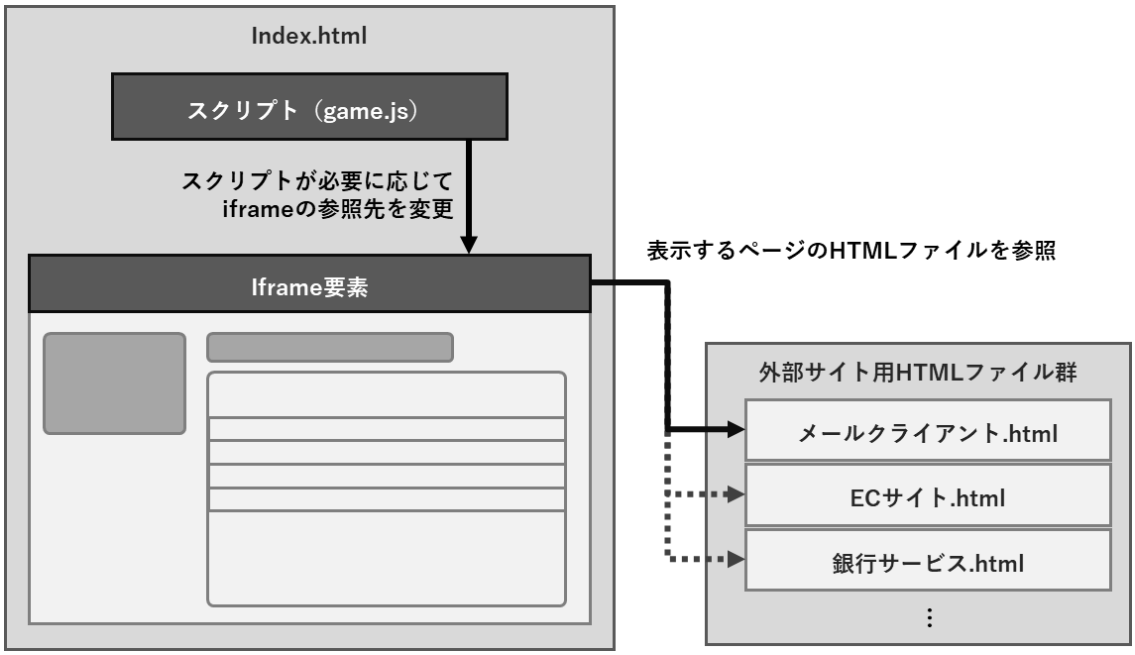


図 4.1 画面管理システムの概要図

この設計により，実際の Web ブラウジングにおけるページ遷移を模擬し，学習者に現実的な判断状況を提供すると同時に，iframe 内のページが切り替わっても制御層の index.html とスクリプトの内部状態（進行状況，スコア，ログ等）を保持することが出来る．また，訓練制御ロジック（index.html および主要なスクリプト）とユーザインタフェース（iframe に埋め込まれて画面を構成するコンテンツや，それらの動作に必要な補助的スクリプト等）を明確に分離し，保守性を向上させる．

層間のデータの受け渡しには，Window: message イベントを利用した．iframe 内で発生したイベント（メール開封，リンククリック，外部サイトでのアクション完了等）は，postMessage により親ウィンドウに伝達され，制御層がこれを受け取って適切な処理（スコア更新，ログ記録，次メール配信等）を実行する．また，制御層から iframe 内へとデータを受け渡す場合にも同様の方式を用いている．

4.1.3 スクリプトの役割

本システムは，機能別に分割された複数の JavaScript ファイルを持つ．主要なファイルとその役割は以下の通りである．

- game.js：
  訓練フロー全体の主要な制御と，LLM API を用いたコンテンツ生成処理を担当する．メール配信スケジューリング，タスク管理，スコア計算，日付進行処理等を実装.
- mail-scenarios.js：
  5 つの送信者タイプ（大学，教授，銀行，就活サイト，EC サイト）を基盤とし，訓練で届くメールの一覧をメールシナリオとして定義する．各シナリオには，フェーズ構成，配信タイミング，メール内容の概要，スコア設定が含まれる．本研究では後述する実験条件を踏まえ，架空の大学での学生生活というシナリオ設定を採用し，送信者タイプも一部これを反映して設定した．設計上は，送信者タイプを含めた全てのシナリオ設定は本スクリプトファイル内で完結しているため，アプリケーションの利用環境や訓練対象者に応じたシナリオ変更は，本スクリプトの書き換えのみで実現可能である.
- log-collector.js：
  操作ログの収集を担当する．訓練終了時に収集されたログを JSON 形式でエクスポートする.
- mail.js：
  メールクライアント画面の UI 制御を担当する．メール表示，メニュー操作，メールソース表示機能を実装.
- external-page-handler.js：
  外部サイトページに共通して組み込まれる汎用スクリプト．アカウント情報等の入力補助，タスク完了検出および制御層への伝達を実装.

これらのスクリプトは明確な責任範囲を持ち，postMessage やグローバル変数を介して連携することで，システム全体の機能を実現している.

4.2 メールシナリオ設計

本節では，訓練で使用されるメールシナリオの設計について述べる.

メールシナリオは，訓練においてユーザに送信されるメールの一覧を定義するデータ構造であり，mail-scenarios.js で定義される．このデータは，後述する LLM API を用いたコンテンツ自動生成機能において使用される.

メールシナリオの設計では，まずメールを送信する主体の一覧を「送信者タイプ」として定義する．この送信者タイプを起点として，各送信者タイプの配下にそれぞれが送信するメールの内容が階層的に定義される構造となっている．この設計により，訓練におけるシナリオ設計をアプリケーション本体の機能から

分離し，このデータファイルを変更することで訓練設計を柔軟に行うことができるようにした．

送信者タイプを事前に固定する設計としたのは，自動生成によって訓練シナリオに登場する人物（ユーザと関わりのある送信者）が無秩序に増加し，仮想環境の一貫性が損なわれる事態を防ぐためである．登場人物の数はユーザの判別負荷と直接的に関連すると考えられるため，この要素を完全に制御可能な設計とすることを重要視した．

4.2.1 送信者タイプを基盤とするシナリオ設計

本システムでは，大学生活における典型的なメールコミュニケーションを再現するため，以下の 5 つの送信者タイプを設定した．

1. university（大学学生課）：
   ユーザが所属する架空の大学の学生課という設定．学務システム更新，研究室配属希望等，学生生活において重要な連絡を再現する．
2. professor（教授）：
   ユーザの指導教員という設定．研究室配属通知や課題連絡等，研究・成績に関連する連絡を再現する．
3. bank（銀行）：
   ユーザが利用する銀行のオンラインバンキングサービスという設定．システム更新通知，セキュリティ設定依頼等，銀行の利用者にとって重要な連絡を再現する．
4. jobSite（就活サイト）：
   ユーザが利用する就活サイトという設定．求人情報通知，先行を受けた企業の書類選考結果通知，内定通知等，就職活動関連の連絡を再現する．
5. ecSite（EC サイト）：
   ユーザが利用する EC サイトという設定．ポイントに関する通知，発送連絡等，オンラインショッピングに関する連絡を再現する．

各送信者タイプに属するシナリオには，正常なメールとフィッシングメールの両方が含まれ，訓練期間中に合計 25 通のメールが配信される．全ての送信者タイプは架空の物であるため，シナリオベースの訓練の中では，ユーザが普段利用しているサービスの違いや，それに基づく関心度・信頼度の違いといった個人差の影響を排除することが出来る．ユーザは訓練の中で個々のメールの特徴を詳細に分析し，正式な送信者の特徴を学びながらフィッシングメールを判別する必要が生じる．

4.2.2 段階的シナリオ展開と文脈の提供

組織に標的を絞ったフィッシング攻撃等の場合，正常なメールのやり取りに紛れてフィッシングメールが送信される場合がある．このような攻撃手法を訓練に反映するため，複数の「フェーズ」で構成される物語的なメール進行構造を持たせることが出来るような設計を行った．

例えば，大学の学務システム更新シナリオ（university_sys_update）は，以下の 3 フェーズで展開する：

- **Phase 0（訓練開始から 165〜185 日後）**：
  正常メール「学務システムメンテナンス事前通知」
  - 内容：約 2 週間後にメンテナンス作業を実施予定であること，完了後にアカウント情報の再登録が必要になることを事前通知する．
  - 適切な行動：メールを開封し，情報を認識する（アクションは不要）．
- **Phase 1（Phase 0 から 14〜18 日後）**：
  正常メール「メンテナンス完了通知・再登録依頼」
  - 内容：メンテナンス完了を報告し，学生ポータルへのログインとアカウント情報更新を依頼する．
  - 適切な行動：リンクをクリックし，学生ポータルでログイン操作を完了する（期限内のアクション完了でスコア獲得）．
- **Phase 2（Phase 1 から 1〜3 日後）**：
  フィッシングメール「システムエラー発生通知（偽装）」
  - 内容：システム更新後にエラーが発生したため，再度のログインが必要になったと通知する．
  - 適切な行動：フィッシングと判断してリンクをクリックしない，またはフィッシング報告を実行する（報告の場合はスコア獲得）．

このような段階的展開により，学習者は「メンテナンス完了後にエラー通知が来た」という文脈の中でフィッシングメールを判断することになる．正常な業務フローの直後にフィッシング攻撃が発生するという，現実的な攻撃タイミングを体験的に理解できる．

各フェーズの配信タイミングは，mail-scenarios.js 内の sendSchedule（訓練開始からの経過日数）または sendInterval（前フェーズからの経過日数）として定義される．これにより，時間的な文脈を含む現実的なメール配信が実現される．また，複数のシナリオで sendSchedule を調整することにより，メール配信を時間的に集中させることも可能であり，これによって多忙な状況を再現する．このような状況では，個別のメールに対応できる時間や注意力が限定されるこ

とが考えられる．これは関連研究でもフィッシング攻撃の被害を受けやすくなる要因として挙げられており，重要な心理的要因の一つである．

4.2.3 難易度パラメータ

各メールには，識別難易度を示すパラメータが付与される．このパラメータは，フィッシングメールの生成時に，その判別の難易度を制御するために使用される．

難易度パラメータは，normal（正常メール），phishing_easy（容易に識別可能なフィッシングメール），phishing_medium（中程度の識別難易度を持つフィッシングメール），phishing_hard（高難易度のフィッシングメール）の 4 段階に設定した．難易度パラメータの違いは，後述するメール生成時のプロンプトに反映され，以下のようなメールの特徴として現れる．

- phishing_easy：
  - 「明らかに不審な」送信者名
  - 「明らかに不審な」メールアドレス
  - 「不安を煽る」件名・本文
  - 認証結果：SPF，DKIM，DMARC 全てが FAIL
- phishing_medium：
  - 「少し不審な」送信者名
  - 「ドメインを見ればわかる程度に不審な」メールアドレス
  - 「緊急性の高い」件名・本文
  - 認証結果：SPF と DKIM は PASS，DMARC のみ FAIL
- phishing_hard：
  - 「巧妙に偽装されているが，偽物の」送信者名
  - 「巧妙に偽装されているが，偽物の」メールアドレス
  - 「送信者からの通常の連絡を装った」件名・本文
  - 認証結果：SPF，DKIM，DMARC 全てが PASS（認証結果だけでは判断できない）

この段階的な難易度設定により，初めは低難易度のフィッシングメールによってユーザに基礎的な判別の手掛かりを学習させ，次第に高難易度のフィッシングメールへ移行することで実用的な判別能力を養うことを図っている．特に phishing_hard は，メール認証技術の限界を示す教育的意義を持つ．

4.3 スコアリングシステム

本節では，訓練の進行と学習者の行動を定量的に評価するスコアリング機構について述べる．スコアリングはユーザに即時のフィードバックを提供するとともに，ゲーミフィケーションの枠組みを構成し，メール対応への動機づけと，フィッシングメール判別における心理的要因の再現の一端を担っている．

4.3.1 タスクベースのスコア管理

本システムは，各メールにおいてユーザに求められる行動を「タスク」として定義し，その完了状況に応じてスコアを加減する．タスクには，基本的には「期限内にメールを開封する」「期限内に外部サイトでアクションを完了する」の 2 種類がある．

タスクとスコアの要素により，ユーザへの心理的要因の再現に関して以下の効果が得られる．

- **現実的な業務プレッシャーの再現**：
  実際の業務(本シナリオでは学生生活)では，重要なメールを見逃したり，期限内に対応しなかったりすることで不利益が生じる．これをスコアのペナルティとして再現することで，現実のメール対応に近い心理的圧力が生じる．
- **能動的な判断の促進**：
  単にフィッシングメールを無視するだけでなく，正常なメールには適切に対応する必要があるため，リスクを負いながら全てのメールを能動的に分析する動機付けが生まれる．

各タスクには，mail-scenarios.js の各メールデータに存在する scoring フィールドで以下のパラメータが定義される：

- requireOpen：メール開封が必要か
- openDeadlineDays：開封期限（日数）
- openScore：期限内開封時の獲得スコア
- openPenalty：期限超過時のペナルティ
- requireAction：外部サイトでのアクション完了が必要か
- actionDeadlineDays：アクション期限（日数）
- actionScore：期限内アクション完了時の獲得スコア
- actionPenalty：期限超過時のペナルティ
- reportScore：フィッシング報告時のスコア（フィッシングメールであればプラス，正常メールなら誤報告であるのでマイナス）

4.3.2 時間進行とタスク期限管理

現実的なメール応対環境を再現するため，訓練は仮想的な時間軸の中で進行する．この仮想的な時間軸では，1 日が現実時間の 4 秒で経過し，おおよそ 1 年分のシナリオを 24 分程度かけて進行する．システムは，訓練開始日（本シナリオでは 2026 年 4 月 6 日と設定）を基準としてこの仮想的な日付を進めながら，配信予定日に到達したメールから順番に配信する．

メール配信のタイミングでは，そのメールに付随するタスク情報がスコアリングシステムに登録され，スコアリングシステムは現在完了待ちの全タスクを保持し，状態を管理する．日付が進む度に，スコアリングシステムによって全タスクの期限チェックが自動的に実行され，期限を超過したタスクについてはペナルティが適用される．

ユーザに対する日付とスコアの提示は，訓練画面の上部に位置する UI によって行われる．ユーザは訓練を通して常にこの表示に注意を払う必要があり，自らの判断やタスク対応の結果はスコアの変動を通じてフィードバックされる．

4.3.3 フィッシング報告機能

ユーザは，受信したメールをフィッシングであると判断した場合，そのメールを報告する機能を利用できる．メールメニューから「フィッシングメールとして報告」を選択すると，以下の処理が実行される．

1. メールの threatType（正常/フィッシング）を確認．
2. 実際にフィッシングメールだった場合：reportScore（通常 50 点）を加算し，「正解！」のフィードバックを表示．
3. 正常なメールを誤って報告した場合：reportScore（通常-30 点）を減算し，「不正解！」のフィードバックを表示．
4. 当該メールを報告済みとしてマークし，以降のタスク期限チェックから除外．

この報告機能は，単にフィッシングメールを無視・削除するのではなく，能動的に検出・報告するという実務的なセキュリティ行動を訓練する目的を持つ．報告スコアは，タスクのアクション完了スコアとは独立して設定されており，より積極的なセキュリティ対応を奨励する設計となっている．

また，誤報に対するペナルティ（-30 点）は，慎重な判断を促すとともに，過度に疑り深い態度（正常なメールまで全て疑う）を抑制する効果を持つ．

4.4 コンテンツ自動生成機構

本節では，LLM API を活用したコンテンツ自動生成の仕組みについて述べる．

4.4.1 送信者情報初期化処理

訓練開始前に，各送信者タイプの正式な送信者情報（送信者名，メールアドレス，配信サーバドメイン名）を LLM に生成させる初期化処理が実行される．これらの情報は，正規メールの特徴を見極め，フィッシングメールと比較するための重要な手がかりである．そのため，正規メールの送信者情報は同一の訓練セッション内で変化してはならない．そこで，正規メールを生成する場合は，この初期化処理で作成した送信者情報を常に使用する設計としている．

送信者情報初期化処理は game.js において実装される．具体的には，各送信者タイプについて以下の処理が実行される．

1. mail-scenarios.js から全送信者タイプの送信者プロファイル（その送信者の概要を説明する文章データ）を読み込む．
2. 送信者プロファイルを含むプロンプトを LLM API に送信し，これにふさわしい正式な送信者名，メールアドレス，ドメイン名を生成させる．
3. LLM の応答から生成データをテキストとして取得する．
4. mail-scenarios.js から取得したメールシナリオデータの各送信者タイプに，生成された正規送信者情報を追加する．

生成されたデータは，後続の個別メール生成で一貫して使用される．例えば，大学学生課から複数のメールが送られる場合，全て同じ組織名と公式アドレスが使用され，送信者としての一貫性が保たれる．

4.4.2 メールシナリオデータ変換処理

LLM API を用いたメールの生成処理に先立つ事前処理として，メールシナリオデータの変換を行う．

メールシナリオは，内容の作成と管理のしやすさを考慮して送信者タイプを起点とする階層構造として設計されているが，この構造では結果的に各メールの送信予定日の順序が不規則に配置されることとなる．しかし，訓練の実施においては時系列に沿った送信順序の方が重要であるため，データ変換処理では，まず全シナリオデータを読み込み，個別のメールデータへと分割したのち，送信予定日の順に並べ替える処理を実行する．

変換後の個別のメールデータは，その一つ一つがメールの生成関数に渡すパラメータとして機能するように設計されており，これにより効率的なメール生成処理が可能となる．

4.4.3 メール生成処理

個別のメール生成は，送信者名，メールアドレス，件名，本文，リンク URL の 5 要素を順次生成する多段階プロセスとして実装される．具体的には，以下の手順で 1 通ずつメールを生成する．

日付順に変換されたシナリオデータから 1 件分のデータをメール生成パラメータとして受け取る．このパラメータには，送信者タイプ，メール内容の概要，脅威タイプ（フィッシングメールか否か），難易度パラメータ，送信予定日，リンクやタスクの有無などが含まれる．

大規模言語モデル（LLM）のチャットセッションをリセットし，新たな生成処理を開始する．

1. システムプロンプトとして，LLM に対してメールを生成するアシスタントとしての役割を指示し，どの送信者からの，どのような内容のメールを生成するかを指定する．同時に，以降のやり取りでは要求されたテキストのみを出力することなど，出力形式に関する制約を設定する．
2. フィッシングメールの場合は送信者名，メールアドレス，送信サーバドメインを新たに生成させる．正規メールの場合は，事前に生成した正規送信者情報をそのまま反映する．
3. メールの件名，本文，およびメールごとに固有のメール ID を生成させる．
4. リンクおよびタスクが設定されている場合は，ユーザに表示されるリンク URL も生成させ，対応するタスクデータを登録する．
5. メールソースを生成する．この際，認証結果（SPF，DKIM，DMARC など）は，難易度パラメータに応じて設定される．
6. 生成されたすべての構成要素を統合し，「メールオブジェクト」として保存する．

このようにプロンプトを細分化してメールを要素ごとに生成させる設計を採用したのは，メール 1 通分の HTML 要素を丸ごと生成させると，構文の破綻や DOM 構造の変更が行われる場合があったためである．また，生成内容を最小化して，LLM API の応答時間を短縮する目的もある．

単一のチャットセッション内でメールの構成要素ごとに個別に生成を行わせることで，出力を切り分けて別々のデータとして取得することができる．この方

式を，メールクライアントの HTML ファイル内にあるメールテンプレートと組み合わせることで，構文的に破綻のないメール生成・表示を可能とした．

また，各生成段階において，シナリオデータに含まれる各パラメータをプロンプトに組み込んでいる．システムプロンプトに送信者タイプとメール内容の概要を組み込むことで全体の生成内容を制御している他，件名やアドレス，本文などは，難易度パラメータに基づいて決定される修飾文（「少し不審な」等）を組み込んで調整を行っている．

4.4.4 メールソースと認証情報

標準的なメールの表示要素とは別に，高度な判別知識としてメールソースと各種認証技術の認証結果を生成している．ここでも，メールソースには，標準的なヘッダー（From，To，Date，Subject，Message-ID）に加えて，SPF，DKIM，DMARC の 3 種のメール認証結果を示す Authentication-Results ヘッダーが含まれる．認証結果は，難易度パラメータで定義されているものがそのまま反映される．

前述した通り，phishing_hard では，認証が全て通過している状態を意図的に設定することで，「認証結果が良好でも偽装の可能性がある」という重要な教訓を学習者に提供する．これは，メール認証技術の限界と，多角的な分析の必要性を理解させる教育的意義を持つ．

4.5 ユーザインタラクション機能

本節では，学習者がメールを閲覧・操作し，判断を下すためのインターフェースについて述べる．直感的な操作性と実環境の再現性を両立させることで，効果的な訓練体験を提供する．

4.5.1 メールクライアント画面の設計

メールクライアント画面の HTML は，実際のメールクライアントソフトウェアを模した UI を持つ画面構成となっている．デフォルトでは受信トレイ画面が表示され，受信した全メールがリスト形式で表示される．メールクライアント画面を図 4.2 に示す．

図 4.2 メールクライアント画面

メールをクリックすると開封表示に切り替わり，本文全体と操作メニューが表示される．操作メニューには，「フィッシングメールとして報告」と「ソースを表示」の選択肢があり，これらはそれぞれ，スコアリングシステムにおいて説明したフィッシング報告アクションと，メール生成機能で説明したメールソース・認証結果の表示機能に結びついている．これらの UI 設計により，ユーザに現実のメールクライアントに非常に近い視覚的表現と操作体験を提供する．

4.5.2 外部サイト連携と現実的な判断状況の再現

リンクを持つ（外部サイトでのアクション要求がある）メールの場合，リンクをクリックすると iframe 内のページが外部サイトページ用の HTML に切り替わる．この際の遷移先ファイルは，メールシナリオデータ内の linkDestination というパラメータで定義されているものである．これらの外部ページは，主に Web サービスのログインページを再現した内容であり，ユーザはフィッシング攻撃における情報入力の段階を模擬的に体験することになる．外部サイトの一例を図 4.3 に示す．

図 4.3 外部サイトページ例

外部サイトページへの遷移時，iframe の読込先 URL を変更すると説明したが，この URL の末尾にクエリ文字列とよばれる形式でメール ID を付加する．これにより，共通の外部サイトページ HTML を使用しながら，どのメールから遷移してきたかを識別可能とした．

全ての外部サイトページには，共通のスクリプトとして external-page-handler.js が組み込まれている．このスクリプトは，学習者がログインボタン等の特定の要素をクリックした際に，クエリ文字列から取得したメール ID と共にタスク完了イベントを親ウィンドウに通知する．親ウィンドウの game.js は，このイベントを受け取り，そのメールに付随するアクションスコアを現在スコアに反映する．また，外部サイトページの設計もこの機構を前提としており，アクション完了を判定するボタン等の要素を共通のクラス名としている．

4.5.3 自動入力機能

外部サイトページの入力フィールドには，自動入力候補を表示する機能が実装されている．架空のシナリオを採用している以上，ユーザに入力内容を任せる

ことはできないが，あらかじめ情報が入っている等の設計では現実的な操作体験を提供できないため，妥協点として自動入力を採用した．

ユーザが情報を入力しようとして入力フィールドをクリックすると，data-autofill 属性で指定されたダミーデータがドロップダウンメニューとして表示され，選択することで即座に入力される．このように，ユーザの行動から自然に自動入力機能が理解できるようにしている．実際の動作の様子を図 4.4 に示す．

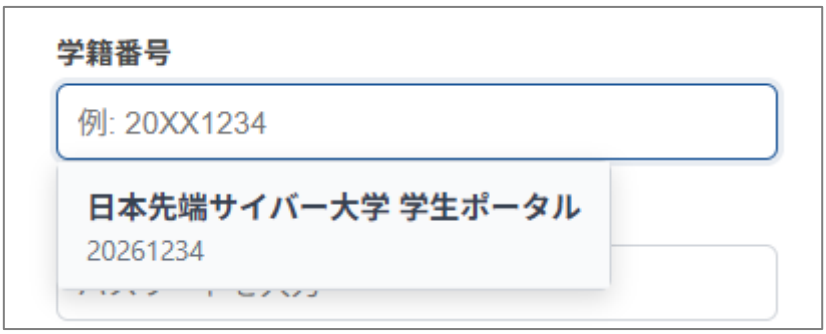


図 4.4 自動入力機能の動作例

4.6 ログ収集・分析システム

本節では，訓練中のユーザの行動を詳細に記録するためのログ収集機構について述べる．

本システムでは，訓練中のユーザのあらゆるインタラクションを時系列イベントとして記録する．イベントには，訓練開始からの経過秒数，イベントの種類，対象オブジェクト，および関連するメタデータが含まれる．

記録される主要なイベントには，以下のようなものがある．

- mail_received：メール受信
- mail_opened：メール開封（受信から開封までの時間も記録）
- mail_closed：メール非表示
- mail_source_viewed：メールソース表示
- mail_source_closed：メールソース非表示
- link_clicked：メール内リンクのクリック（リンク先 URL も記録）
- action_completed：外部サイトでのアクション完了（アクション種別も記録）
- phishing_reported：フィッシング報告

これらに加え，その訓練セッションで生成されたすべてのメールのデータが記録される．メールデータと各イベントログにはメール ID が記録されているの

で，各イベントがどのメールに対する操作であったのかを後から追跡可能となっている．

これらのログは，訓練終了時に json ファイルとして取得できる設計とした．これを事後的に分析することで，訓練中のユーザの行動傾向を明らかにする．

第5章 評価実験

5.1 実験概要

本研究で開発したフィッシングメール判別訓練アプリケーションの教育効果を検証するため，被験者実験を実施した．本アプリケーションの，教育効果に関する核心的な設計要素は，実際のフィッシング攻撃における心理的要因の再現にある．したがって，本実験において最も重要な検証項目は，この心理的要因の再現がユーザの学習効果に及ぼす影響の評価である．

この目的を達成するため，本研究では被験者間実験を採用した．具体的には，被験者を 2 つのグループに分け，実験群には心理的要因を含むアプリケーションを，統制群には心理的要因の再現要素を除外したアプリケーションを使用させた．被験者間実験を選択した理由は，いずれのバージョンのアプリケーションを使用した場合でも一定の学習効果が生じることが予想されるため，同一被験者に両バージョンを連続して使用させる被験者内設計では，学習効果の交絡により各条件の純粋な効果を正確に測定することが困難であると判断したためである．

実験群・統制群のいずれにおいても，アプリケーションを用いた訓練の前後で記述式のフィッシングメール判別テストとアンケートを実施し，これらの結果を分析することで両者の教育効果の違いを検証することとした．

実験実施にあたり，北陸先端科学技術大学院大学ライフサイエンス委員会の承認を得た（承認番号：人 07-033）．

5.2 実験方法

5.2.1 実験参加者

本実験の被験者は，本学の学生の中で，フィッシングに関する専門的な教育や訓練を受けた経験のない者を対象として募集し，合計 10 名が集まった．

本学には学内メールサービスがあり，重要な対応事項を含むメールが日常的に配信される．そのため，本学の学生であれば，メール応対に関する基本的な経験と意識に大きな個人差は生じないと判断した．また，本研究で開発するアプリケーションは，十分な訓練を受けたことのない一般的なユーザを対象としたものである．そのため，高度な訓練の経験がある被験者では学習効果が限定的であると考え，そのような経験を有していないことを条件とした．

5.2.2 実験群と統制群

実験群と統制群の主要な差異は，心理的要因の再現に関わるゲーミフィケーション要素の有無である．実験群のアプリケーションには，実際のフィッシング攻撃における心理的要因を再現するため，送信者とユーザとの関係性の設定，タスクとスコア，時間的制約，そしてシナリオ設計によるメールの時間的集中(多忙な状況の再現)といった要素が組み込まれている．これに対し，統制群で使用するアプリケーションは，これらの要素を全て排除した構成とした．

両群間の条件を可能な限り統一するため，提示するメールのシナリオデータは両群で共通のものを使用し，メールの提示順序もシナリオに準じて同一とした．ただし，統制群では仮想的な日付設定を行わないため，1 通のメールを提示し，被験者がそれをフィッシングメールか否か判断すると，次のメールが提示されるという単純な連続的判別形式を採用した．また，タスクの概念が存在しないため，フィッシングメールか否かの判断はメールクライアント画面のメニューから直接行う仕組みとし，判断直後に正誤のフィードバックを表示する設計とした．統制群の被験者は判断に任意の時間をかけることが可能であり，正誤に関わらずペナルティは一切存在しない．このように，統制群アプリケーションは心理的要因を伴わない純粋な学習環境として機能するよう設計されている．

5.2.3 事前テストと事後テスト

実験群，統制群のいずれにおいても，アプリケーションを用いた訓練の前後で記述形式のテストを実施した．

両テストは，アプリケーションのメール生成機能を流用して作成した疑似的なフィッシングメールのサンプルを被験者に提示し，メール内の不審点の指摘，フィッシングメールか否かの判断，およびその判断根拠を記述させる形式とした．

このような記述形式のテストを採用した理由は二つある．

第一に，テスト自体が訓練効果を持つことによる測定への影響を最小化するためである．多数のメールサンプルを用いた選択式テストでは，テスト実施そのものがフィッシングメール判別の学習機会となり，事前テストと事後テストの間で生じた能力変化が，アプリケーション使用による効果なのか，事前テストでの学習効果なのかを区別することが困難になる．そのため，本研究では提示するメールサンプル数を最小限に抑え，かつ訓練シナリオ内で使用されるメールとは直接関連性のない内容のサンプルを用いることで，テストそのものによる学習効果を抑制することを試みた．

第二に，本実験における被験者数の制約を考慮し，質的評価を中心とした分析手法が適切であると判断したためである．被験者数が限定的である場合，統計的検定力の観点から量的データのみに依存した分析では効果の検出が困難となる可能性がある．これに対し，記述式の回答から被験者の思考過程や判断根拠を詳細に分析することで，アプリケーション使用による認知的変化や判別スキルの質的な向上を捉えることが可能となる．具体的には，被験者が指摘する不審点の種類や指摘内容の具体性，判断根拠の論理性等を評価することで，単なる正答率では測定できない判別能力の多面的な変化を把握することを試みた．

5.3 実験結果

5.3.1 事前アンケート

事前アンケートは，主にフィッシングに関する事前知識と訓練経験について確認する目的で実施した．結果を表 5.1 に示す．

表 5.1 事前アンケート結果

内容	人数
フィッシングについて聞いたことがある	8
フィッシングに関する詳細な教育・訓練を受けた経験がある	0
実際にフィッシング攻撃を受けたことがある	2

実験条件でも述べた通り，被験者全員がフィッシングに関する本格的な教育・訓練の経験を持たないことを確認した．その他については，フィッシングそのものを聞いたことがあるとする回答が 8 名と多数を占めた．フィッシングの攻撃を受けた経験については，補足説明として「被害に至っていなくとも，フィッシングメールであることが明らかなメールを受け取ったことがあること」とし，これに該当する被験者は 2 名であった．

5.3.2 事前・事後テスト

事前テストと事後テストは，訓練アプリケーションと同様のメール生成機構を用いて作成した架空のメールをユーザに提示する形で実施した．具体的には，生成したメールの開封状態の画面とメールソースを表示した状態の画面のみを

確認可能とした. テストで使用した全てのメールの送信者は同一のもの (あるいはその偽装) としたが, 訓練で使用されるコンテンツには含まれないものであり, テスト内容と訓練の文脈は分離されている.

各テストで提示されるメールは以下の通りである.

- 事前テスト:
  - メール 1 (難易度 : phishing_easy に相当するフィッシングメール) を提示.
  - メール 1 の不審点・注意した点を回答させる.
- 事後テスト:
  - メール 1 に加え, メール 2 (正式なメール), メール 3 (難易度 : phishing_hard に相当するフィッシングメール) を提示
  - 不審点・注意した点に加え, 各メールについてフィッシングか否かの判断と, その根拠を回答させる.

事前テストと事後テストの回答については, その内容に関する質的分析を行うため, コーディング作業を実施して 3 レベルに分類した. 各レベルの基準は以下のようなものである.

- **レベル 1 : 漠然とした指摘**
  - 「なんとなく不安を感じる」「偽物かもしれない」「こんなメールが来るのはおかしいと思う」等.
  - 文脈や全体的な内容に対する違和感を覚えているが, 具体的な対象箇所や根拠を述べていないもの.
- **レベル 2 : 具体的指摘**
  - 「メールアドレスが不審」「SPF や DKIM が FAIL」「本文が不安を煽る・緊急性がある」等.
  - 不審点の具体的な対象を挙げているが, どのようにおかしいのかを述べていないもの.
  - 文面の不審点を挙げているが, 性質の説明に留まるもの.
  - 認証技術に関する指摘は, 被験者が判定結果を確認できても, その理由や仕組みを即座に理解することは困難であり, かつ検出精度にも限界があるため, 説明的指摘とは扱わないこととした.
- **レベル 3 : 説明的指摘**
  - 「メールアドレスの a が α に偽装されている」「本文で不安にさせ, ログインさせようとしている」等
  - 不審点の具体的な対象箇所に加え, 正式なアドレスや送信者名 (と被験者が判断したもの) との比較や, 偽装されている証拠の発見ができているもの.

- 文面の不審点に加え，それがフィッシング攻撃のプロセスであることを理解し，自身が説明できているもの．

なお，複数の不審点や根拠が述べられている場合，レベルの最大値を採用することとした．不審点の数は提示するメール自体の内容に依存するため，スコアの合計値を用いるのは不適切と判断したためである．また，本研究で開発するアプリケーションの狙いは，現実的な制約のある状況下で，限られた手がかりからでもフィッシングを判別できる能力を養うことである．この点からも，不審点の量ではなく，内容の質を評価すべきであると考えた．

コーディング作業は，筆者と研究協力者の 2 名で独立に実施した．評価者間一致度を Cohen's κ 係数により算出したところ約 0.67 であった．評価が一致しない箇所については協議の上，合意に至ったレベルを採用した．

実験群・統制群それぞれの事前・事後テストにおける各質問への回答のレベルを表 5.2，表 5.3 に示す．なお，回答が全くなかった場合は欠損値とした（表中 N/A と表記）．

表 5.2 実験群 事前・事後テストの回答レベル

被験者 ID	メール 1 前	メール 1 後	メール 2	メール 3
001	3	2	2	3
002	3	3	3	2
007	3	3	3	3
009	2	2	2	N/A
010	3	3	3	2

表 5.3 統制群 事前・事後テストの回答レベル

被験者 ID	メール 1 前	メール 1 後	メール 2	メール 3
003	3	2	3	3
004	N/A	1	2	1
005	2	2	N/A	1
006	3	3	2	3
008	2	2	2	2

これらの回答レベルをスコアとみなし，統計的分析を行う．ラベルの数字をそのままスコアとする方式を採用する．欠損値の場合は 0 ではなく，データ自体を無いものとして扱う．

メール 1 については，前後のテストで同じメールを提示したため，スコアの

変動を比較することができる．実験群と統制群について，事前・事後でのスコア変動値に Mann-Whitney U 検定を行い，両群で回答の質の変化に有意な差があるかを検証した．統制群の被験者 1 名が事前テストで無回答であったため，サンプルサイズは実験群 5 名，統制群 4 名となった．結果は，U=9.5 となり，5%水準での有意差は認められなかった．したがって，今回の分析の範囲では，訓練アプリケーションの使用による学習効果において，心理的要因の再現が有効であったことを確認できなかった．

ただし，全被験者においてそもそもスコアの増加がみられなかったため，両群共にアプリケーションによる学習効果自体を確認できなかったことになる．この点については，テスト自体の学習効果を避けるために，テストの問題数をごく少数としたことや，記述形式の回答としたこと，および質的評価のためのコーディング基準等に問題があった可能性もあると考えられる．また，事前テストの時点で高レベルの回答が出ている例が多いことから，テスト問題の難易度も低かった可能性がある．

これらのことから，テスト前後の変化がより顕著に表れるような，異なるテスト手法を検討する余地があるものと考える．

5.3.3 事後アンケート

事後アンケートは，アプリケーションによる訓練とその効果に関する被験者の主観的な印象を調査するために実施した．質問内容は以下の通りである．

- Q1：事前テストの実施中，焦りや緊張を感じましたか？
- Q2：アプリケーションを使用した訓練で，焦りや緊張を感じましたか？
- Q3：アプリケーションを使用した訓練で，メールへの対応が大変だと感じましたか？
- Q4：アプリケーションを使用した訓練で，すぐに読みたい，あるいは読まなければならないと感じるようなメールはありましたか？
- Q5：アプリケーションを使用した訓練で，フィッシングメールの回避能力が向上したと思いますか？
- Q6：その他気が付いたことやコメントがあれば教えてください（自由記述）

Q1～Q5 は，4 段階（1：否定的～４：肯定的）での回答とし，Q6 は自由記述とした．Q1～Q5 までの，各被験者の回答を表 5.4，表 5.5 に示す．表 5.4 は実験群，表 5.5 は統制群の被験者の回答である．

表 5.4 実験群 事後アンケート結果

被験者 ID	Q1	Q2	Q3	Q4	Q5
001	2	3	4	3	3
002	2	2	2	3	3
007	1	4	3	3	4
009	1	1	3	4	3
010	2	2	2	3	3
中央値	2	2	3	3	3

表 5.5 統制群 事後アンケート結果

被験者 ID	Q1	Q2	Q3	Q4	Q5
003	2	3	4	3	4
004	2	2	2	3	3
005	1	1	1	1	4
006	1	1	1	2	4
008	1	1	1	3	4
中央値	1	1	1	3	4

Q1 は，シナリオ的文脈がない事前テストの段階であっても，メールの判別作業自体に心理的作用が存在するかどうかを確認するための質問である．結果としては，全被験者が 1 か 2 を回答しており，この作業自体に大きな心理的作用はなかったものと判断する．

Q2 については，実験群の方がわずかに肯定的な傾向があるものの，大きな差は見られなかった．現実の業務環境に近い心理状態，特に注意が散漫する状況を再現することは本研究の目標であったが，焦りや緊張感といったものには大きな影響はなかったと考えられる．ただし，Q3 には全質問中で最も大きな差があり，メールへの対応が大変だったかどうかについては実験群のほうが肯定的な傾向にある．メールシナリオ設計によるメール配信の時間的な集中が，部分的に効果を発揮したものと考えられる．

Q4 は全体的に実験群のほうがわずかに肯定的であるが，両群共に 3 が最多の回答であるため，メールへの関心や注意に対しては，心理的要因の再現はそれほど大きな効果を発揮しなかった可能性がある．事前に練習モードのようなものを追加したり，ゲーミフィケーションのゴール設定やユーザアクションに対する報酬体系をより明確なものにしたりといった改善により，没入感を高める必要があると考えられる．

Q5 は統制群のほうが肯定的な傾向がみられる結果となった．実験群では，スコアリングや時間制限が存在することから，メールへの対応そのものに意識が集中してしまい，フィッシングメールの特徴を学習できたという実感が得られなかった可能性がある．また，Q6 の自由記述において，「フィードバックが欲しい」という意見があった．本研究の実験では，リアルタイム性を重視するためにユーザのフィッシング判断に対するフィードバックは最小限のものとしたが，学習効果の観点からはある程度詳細なフィードバックがあるべきだと考えられる．

5.3.4 ログ分析

操作ログの情報から，被験者の行動傾向を分析する．実験群と統制群では時間制限やゲーミフィケーション要素の有無という違いがあり，1 メールの判断にかけられる時間や心理的なプレッシャーは大きく異なることが考えられる．そのため，判断にかかった時間や訓練内での判別精度等は単純に比較できないと考えた．そこで，両者で比較可能な行動傾向として，「あるメールをフィッシングかどうか判断するまでに，どれほどの情報を参照したか」を用いることとした．具体的には，メールを開封してから判断を下すまでに，「他のメールを確認した割合」と，「メールソースを確認した割合」を集計した．これらの集計結果を表 5.6 と表 5.7 に示す．表 5.6 は実験群，表 5.7 は統制群の集計結果である．

実験群はシナリオ設計上，能動的なフィッシング判別を必要とするのはアクション要求を持つメールのみである．また，アクションには対応期限が存在するので，すべてのアクションを完了できない可能性が存在する．一方で，統制群は一律にすべてのメールについて判別を行う設計である．したがって，ユーザがメールを判別する回数には両群で差が生じるため，全判別回数に対する割合を採用した．また，被験者 ID 001 のログデータにはログ収集システムの不備による欠損が発生したので，集計していない．

表 5.6 実験群 ログ集計結果

被験者 ID	他メール確認[%]	メールソース確認[%]
002	26.7	100.0
007	0.0	0.0
009	31.3	87.5
010	18.8	12.5

表 5.7 統制群 ログ集計結果

被験者 ID	他メール確認[%]	メールソース確認[%]
003	20.0	88.0
004	4.0	4.0
005	32.0	72.0
006	20.0	84.0
008	24.0	72.0

ログの集計結果より，他のメールの確認率，メールソース確認率ともに個人差が大きいことがわかり，両群で一貫した差異は見出せなかった．しかし，多くの被験者は他メール確認率に比べるとメールソース確認率が高いことがわかる．他メール確認率は全被験者で 32%未満だが，メールソース確認率は 70%以上の被験者が合計 7 名存在する結果となった．メールソースの表示機能は一般的なメールクライアントにも備わっているものだが，事前アンケートの結果を考慮すると被験者が日常的に利用しているとは考えにくいので，訓練という状況が積極的なソース確認行動を引き起こしたと考えられる．

次に，メールの判別について分析した．先に述べた通り，本実験では実験群と統制群で被験者が実施する訓練内容に差異があり，フィッシングメールか否かの判別についても条件が異なる．

具体的には，実験群はタスクが付随する（リンク先での情報入力が必要となる）メールにおいて，そのタスクに対応するか否かが主な判別行為となる．「正常なメール」であれば，タスクに対応しないことでペナルティが生じるため，被験者は必然的に正常と判断したメールのタスクに対応することになる．一方で，「フィッシングメール」とユーザが判断した場合は，タスクを放置しても問題ないが，別の選択肢としてフィッシングとして報告することもできる．

統制群ではゲーミフィケーション要素を完全に排除したため，単に全てのメールを順番に判別するのみの訓練設計となった．そのため，初めからタスクの概念自体がなく，被験者の判別行為はどのメールに対しても一律に、正常かフィッシングかを選択するのみとなっている．

このように，実験群と統制群では被験者が直接判別するメールの数に違いがある上，判別における選択肢も異なる．さらに，実験群では例えフィッシングと判断していなくとも，タスク対応期限の超過によって未対応となる場合も想定できる．これらを踏まえ，実験群と統制群それぞれについて，被験者が「正常」または「フィッシング」と判断したメールとその正誤を集計することとした．数の違いを踏まえ，各項目の集計結果は被験者の直接の判別対象となっていた全

メールに対する割合を用いる．また，両群での比較項目の統一のため，ここでは実験群におけるタスク対応期限の超過は「フィッシングと判断して不正解だった場合」，フィッシングメールのうち無視されたものは「フィッシングメールと判断して正解だった場合」と見なす．

なお,集計において本来同時に達成されるべきでない 2 つ以上の判別行為(例:フィッシングとして報告した後にそのメールのタスクに対応する）が行われている例が発見された．これは訓練システムの実装上の不備であり，判別対象となるメールの合計数に対して判別行為の数が過剰となってしまう．そのため，このような場合は「最初に達成された判別行為」のみを採用することとした．

メールの判別とその正誤についての集計結果を表 5.8 と表 5.9 に示す(例:「正常と判断して正解だった場合」は「正常/正解」と表示)．表 5.7 は実験群，表 5.9 は統制群の集計結果である．

表 5.8 実験群 メール判別結果

被験者 ID	正常 /正解	正常 /不正解	フィッシング /正解	フィッシング /不正解	正答率
002	0.50	0.06	0.38	0.06	0.88
007	0.19	0.19	0.25	0.38	0.44
009	0.50	0.06	0.38	0.06	0.88
010	0.38	0.06	0.38	0.19	0.75

表 5.9 統制群 メール判別結果

被験者 ID	正常 /正解	正常 /不正解	フィッシング /正解	フィッシング /不正解	正答率
003	0.72	0.00	0.28	0.00	1.00
004	0.64	0.08	0.16	0.12	0.80
005	0.60	0.08	0.20	0.12	0.80
006	0.72	0.00	0.28	0.00	1.00
008	0.72	0.00	0.28	0.00	1.00

表 5.8，表 5.9 より，メール判別の正答率は統制群の方が高いことが分かる．実験群と比べて時間制限がなく，未対応による間違いが発生しない上，判断に時間をかけることも可能であるため，この結果は予想できるものである．

両群で違いが見られる点としては，まず「正常と判断して不正解だった割合」が挙げられる．実験群では全ての被験者で 1 例以上このパターンが存在しており，これはフィッシングメールだと見抜けずにタスクに対応してしまったこと

を意味する．統制群ではこのパターンは 2 名の被験者にしか見られない．実験群では時間的な制約と期限超過のペナルティが存在することから，タスクへの対応を急ぐような心理が働いた可能性は十分に考えられる．そのため，十分な分析を行わないままタスクに対応してしまった場合が僅かに多いのではないかと考える．

「フィッシングと判断して不正解だった割合」についても違いが見られ，実験群では全被験者で 1 例以上このパターンが存在し，統制群では 2 名に存在するのみである．ただし，このうち「正常メールを誤って報告した」パターンは被験者 ID：010 に 2 例があるのみで，それ以外は「正常メールを無視した（タスク対応期限を超過した)」パターンである．これは先述の通り，そもそも対応できる余裕がなかっただけの可能性も考えられるため，一概に実験群の方がフィッシングであるとする判断を下しやすい傾向があったとは言えないと考えられる．

しかしながら，特に ID：007 で顕著なように，実験群において「フィッシングと判断して不正解だった割合」が大きいことは事実であり，心理的要因の再現設計に起因する対応負荷の大きさか，フィッシングを疑う心理傾向の少なくともいずれかを増加させる働きがあったことが予想できる．純粋に期限までに対応できない場合が多かったのであれば，シナリオデータにおける時間制限やメール配信の時間的な集中の設計が過剰であったという可能性もあるため，シナリオ設計にも考慮すべき点があると考える．

5.4 他 LLM モデルを用いた予備的検討

本研究で開発した訓練アプリケーションでは，コンテンツ自動生成用の LLM モデルとして Claude 3.5 Haiku API を使用した．これは，利用のしやすさ，コード生成・長文処理性能，応答速度等から検討して決定したものである．Claude API ではより新しいモデルが利用可能であったが，新モデルでは架空のメール送信者に関する情報やフィッシングメールを構成するコンテンツの生成が倫理的なポリシーに抵触するため実行できないとする応答が返ってくることが判明し，プロンプトの工夫で解決を図ることが出来なかったため，止む無く旧モデルでの実装となった．

LLM には他にもいくつかの代表的なモデルが存在し，新たなモデルも次々にリリースされている．そこで，将来的に本アプリケーションシステムにより適したモデルを採用することを視野に，予備的に他のモデルを用いたコンテンツ生成を試験することとした．本研究で採用した Claude 3.5 Haiku API の他，Google Gemini[17]，OpenAI GPT[18]を比較する．ここでは，各モデルで 3 通分のメールを生成し，平均生成時間と生成内容の品質について簡単に述べる．表 5.7 に

各モデルの比較を示す．生成されたメールの具体例は，付録に図で示す．

表 5.10 生成 AI モデルの比較

モデル名	平均生成時間[秒] (平均±標準偏差)	生成内容
Claude 3.5 Haiku	19.95 ± 3.95	本文はやや簡素であるが，ソース文が充実している． メールアドレスへの難易度の反映が，ドメインによる判別難度として表れており，現実的である．
Gemini 2.5 Flash	39.81 ± 15.61	本文・ソース文の内容は現実的であるが，プロンプトによる指定に反し，出力テキストに会話のようなものが混ざる場合がある． メールアドレスは難易度パラメータに応じたものであると感じられる．
Gemini 2.5 Flash Lite	6.52 ± 1.01	本文・ソース文が短い． アドレス・リンク等に破綻はないが，難易度が内容に反映されていないように見える．
GPT 4o	12.59 ± 2.60	本文は短いが，文面は現実的なメールらしいものになっている． 難易度の反映も確認できるが，違いが小さく全体的に高難度に感じられる．
GPT 4o mini	17.57 ± 7.75	本文がかなり短く，内容が簡素過ぎて現実的ではない． アドレス等は難易度が反映されていると考えられるが，特徴が露骨過ぎる印象を受ける．

各モデルの生成時間と生成内容を総合的に評価すると，本研究で実際に使用した Claude 3.5 Haiku は生成時間が長い方であるものの，特に難易度指定の反映を含めた生成コンテンツの品質は高かったと判断する．ただし，GPT 4o も生

成時間が短い割にコンテンツの品質は高く，候補として検討する価値があるものと考える．LLM は新たな製品が次々に登場している現状があり，各モデルで価格や利用形態も異なる．また同じ目的であっても最適なプロンプトや扱い方も異なることが考えられるので，実装においてもモデルごとに工夫の余地があると考える．これらのことから，将来的に本アプリケーションシステムに最適なモデルも変化していく可能性があり，今後も検討すべき事項であると言える．

第6章 おわりに

6.1 本研究のまとめ

本研究では，フィッシングメールの判別における心理的要因の重要性に着目し，これを再現した疑似体験型訓練アプリケーションを開発し，実験を通してその教育効果を検証した．具体的な実装としては，心理的要因の再現に加え，運用・拡張の容易性を確保することを目的として，生成 AI を活用したコンテンツ自動生成機能と，ゲーミフィケーション要素を取り入れたシナリオ駆動型訓練を組み合わせたアプリケーションを開発した．

心理的要因については，個人要因，環境要因，外部刺激要因の 3 つに分類し，それぞれを訓練環境内で効果的に再現する設計を行った．具体的には，仮想的な学生生活シナリオにおいて，送信者との関係性の設定，タスクとスコアによる業務関連性の再現，時間制限の導入，そしてメール配信の時間的集中による多忙な状況の再現を実現した．

コンテンツ生成においては，LLM API を用いたメールコンテンツの自動生成機構を構築し，送信者タイプを基盤とする階層的なメールシナリオ設計を採用した．これにより，シナリオの枠組みを保持しながら，訓練試行ごとに異なる内容のメールを効率的に生成することを可能とした．また，iframe 構造による画面管理と postMessage を用いた層間通信により，専用のサーバ環境を必要とせずに実際の Web ブラウジングに近い操作体験を提供しつつ，訓練の進行状況やユーザの行動ログを安全に管理できる設計とした．

開発したアプリケーションの教育効果を検証するため，10 名の被験者を実験群と統制群に分け，心理的要因の再現の有無が学習効果に及ぼす影響を評価する被験者実験を実施した．事前・事後テストの記述式回答に対する質的分析，および事後アンケートとログデータの分析を行った．

実験の結果，事前・事後テストにおける回答の質の変化について，実験群と統制群の間に統計的に有意な差は認められなかった．この要因として，テスト設計や評価基準の問題，および訓練時のフィードバック不足などが考えられる．一方で，事後アンケートからは，実験群において「メールへの対応が大変だった」という回答が統制群よりも肯定的な傾向を示し，シナリオ設計によるメール配信の時間的集中が一定の効果を発揮したことが示唆された．また，ログ分析からは，多くの被験者がメールソースを積極的に確認する行動を示しており，訓練環境が詳細な分析を促す効果があったことが確認された．

6.2 今後の課題

本研究で開発したアプリケーション，および実験手法には課題が残った．

アプリケーションは，運用面では専門人員を擁する大規模組織でなくとも実施・拡張できるものを目指して開発した．そのために，可能な限り専用サーバ等の高度な機材や，追加の外部ソフトウェアを必要としない設計を目指したが，LLM API の運用方式からローカルサーバが必要になり，今回の実装では Node.js のインストール環境が前提となってしまった．また，全ての機能をクライアントサイドで実装する設計では CORS ポリシーとの競合が起こる場合が多々発生し，開発難度を引き上げる要因となった．

シナリオ設計をシステムの主要機能から分離し，簡潔な自然言語で内容を指定できるようにすることはできたが，現実的な利便性を考慮すれば，シナリオ設計を行うためのユーザインタフェースも実装すべきであると考える．外部サイトページに関しても事前に作成済みのものを使用する設計としてしまったが，関連研究で採用されていた手法に倣い，何等かの自動生成を取り入れることが望ましい．

総じて専門的な知識や高度な環境を必要とせずに運用・シナリオ作成・拡張を行うには及んでいないと言える完成度であり，改善の余地が多い．

実験についても，被験者間実験としては十分なサンプル数が集められたとは言えず，統計的にも有意な結果を得ることが出来なかった．また，実験条件を最適化することを優先するあまり，アプリケーション設計の教育的意義や収集データ数が犠牲になった点があった．これらを改善し，より実用条件に近い有用な実験結果を得る必要がある．

謝辞

本研究の遂行にあたり，研究テーマ決定から開発・実装の技術的支援，論文作成に至るまで，多くのご指導，ご助力を頂きました長谷川忍教授ならびに研究室の皆様に，心より感謝申し上げます.

また，実験結果の分析に協力していただいた同研究室の廣部陸氏にも深く感謝いたします.

多忙な時期にもかかわらず，快く実験参加を引き受けてくださった 10 名の被験者の皆様も，誠にありがとうございました.

参考文献

[1] 警察庁,“フィッシング対策”, https://www.npa.go.jp/bureau/cyber/countermeasures/phishing.html (最終アクセス 2026 年 1 月 12 日).

[2] 警察庁,“サイバー空間をめぐる脅威の情勢等”, https://www.npa.go.jp/publications/statistics/cybersecurity/index.html (最終アクセス 2026 年 1 月 12 日)

[3] アスピック,“標的型攻撃メール訓練サービス比較 15 選！開封率何%なら合格?”, https://www.aspicjapan.org/asu/article/2419 (最終アクセス 2026 年 1 月 23 日)

[4] 大塚商会,“標的型メール攻撃に備える 標的型メール訓練サービス”, https://www.otsuka-shokai.co.jp/products/security/consulting-education/aptmail-training-service/ (最終アクセス 2026 年 1 月 23 日)

[5] GSX,“トラップメール - 標的型攻撃メール訓練サービス【国内シェア No.1 のメール教育 Saas】” https://www.gsx.co.jp/services/securitylearning/trapmail.html

[6] 内山 他,“信州大学における疑似体験型攻撃メール訓練の開発”, オペレーションズ・リサーチ, Vol.64, No.9, pp.541-548, 2019.

[7] 東野 他,“知識グラフを用いた個人適応型サイバーセキュリティ学習システムの検討”, 情報処理学会研究報告, Vol.2025-CLE-45 No.4, 2025.

[8] 程 他,“適応的フィッシングメール判断トレーニング課題出題手法の評価”, JSiSE Research Report vol.35, no.1, 2020

[9] 東野,“生成 AI を用いた不審メール対応訓練システムの試作”, コンピュータセキュリティシンポジウム 2024 論文集, 00.238-241, 2024.

[10] 稲葉,“フィッシング詐欺における攻撃者への信頼の判断を左右する心理的要因の概説”, 情報セキュリティ総合大学, 一般論文, 2017.

[11] Kavvadias 他,“Understanding the Role of Demographic and Psychological Factors in Users’ Susceptibility to Phishing Emails: A Review”, Appl. Sci. 2025, 15(4), 2236;

[12] John, “User Awareness and Psychological Factors in Falling for Phishing Attacks”, 2025, https://www.researchgate.net/publication/390162214_User_Awareness_and_Psychological_Factors_in_Falling_for_Phishing_Attacks（最終アクセス 2026 年 1 月 14 日）

[13] Workman, “Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security”, Journal of the American Society for Information Science and Technology, Volume59, Issue4 15 February 2008

[14] Scott Monteith 他, “Increasing Cybercrime Since the Pandemic: Concerns for Psychiatry”, Curr Psychiatry Rep, 2021, 23(4): 18.

[15] Claude Docs, “Claude の紹介”, https://platform.claude.com/docs/ja/intro（最終アクセス 2026 年 1 月 15 日）

[16] Node.js, “どこでも JavaScript を使おう”, https://nodejs.org/ja（最終アクセス 2026 年 1 月 15 日）

[17] Google, “ジェミニとは何か, そしてどのように機能するか”, https://gemini.google/jp/overview/?hl=ja(最終アクセス 2026 年 1 月 15 日）

[18] OpenAI, “Models”, https://platform.openai.com/docs/models(最終アクセス 2026 年 1 月 15 日）

付録

付録 1 Claude 3.5 Haiku 生成例

- phishing_easy

【緊急】不正アクセス検出 - 即時対応が必要

ShopSC Customer SupAort <support-verify@shop-sc-secure.com>
To 自分

4月6日

斉場 守 様

緊急のセキュリティ警告：アカウント保護が必要です

当社のセキュリティシステムで、お客様のアカウントに対する不審な動きを検出しました。複数の異なる場所からログイン試行の形跡があり、アカウントが危険にさらされている可能性が高いと判断しました。

お客様の個人情報と決済情報を保護するため、直ちに以下の対応が必要です：

1. 即時にアカウントセキュリティを確認
2. パスワードの再設定
3. 不正アクセスの有無を確認

お客様の大切な個人情報を守るため、至急ご対応ください。放置した場合、アカウントの凍結や個人情報流出のリスクがございます。

接続用リンク：
https://shop-sc-secure.net/login-verify/account-protection

- normal

ご注文商品の発送のお知らせ

ShopSC カスタマーサポート <support@shopsc.jp>
To 自分

4月6日

斉場 守 様

お世話になっております。ShopSCカスタマーサポートです。

この度は、ご注文いただいた商品の発送を完了いたしましたので、お知らせいたします。

ご注文商品の配送状況をご確認いただくため、マイページへのログインをお願いいたします。

※セキュリティ保護のため、お客様のログインID（メールアドレス）と、ご登録のパスワードにてご確認ください。

ご不明な点がございましたら、カスタマーサポートまでお問い合わせください。

接続用リンク：
https://www.shopsc.jp/tracking/order-status?uid=saiba-mamoru&ref=mails-delivery

- phishing_hard

【重要】荷物の配達に関するお知らせ - ShopSC

ShopSC サポートセンター <support-shopsc-jp.info@delivery-notice.com>
To 自分

4月6日

斉場 守 様

お客様

お客様がご注文の荷物は、配達員による第一配達の際に不在のため、現在当社営業所で保管されております。荷物の再配達を希望される場合は、お客様の配達状況をご確認ください。

お客様の荷物を迅速にお届けするため、以下の手続きにてご対応いただきますようお願いいたします。

※注意：指定期限内に手続きが行われない場合、荷物は返送される可能性がございます。

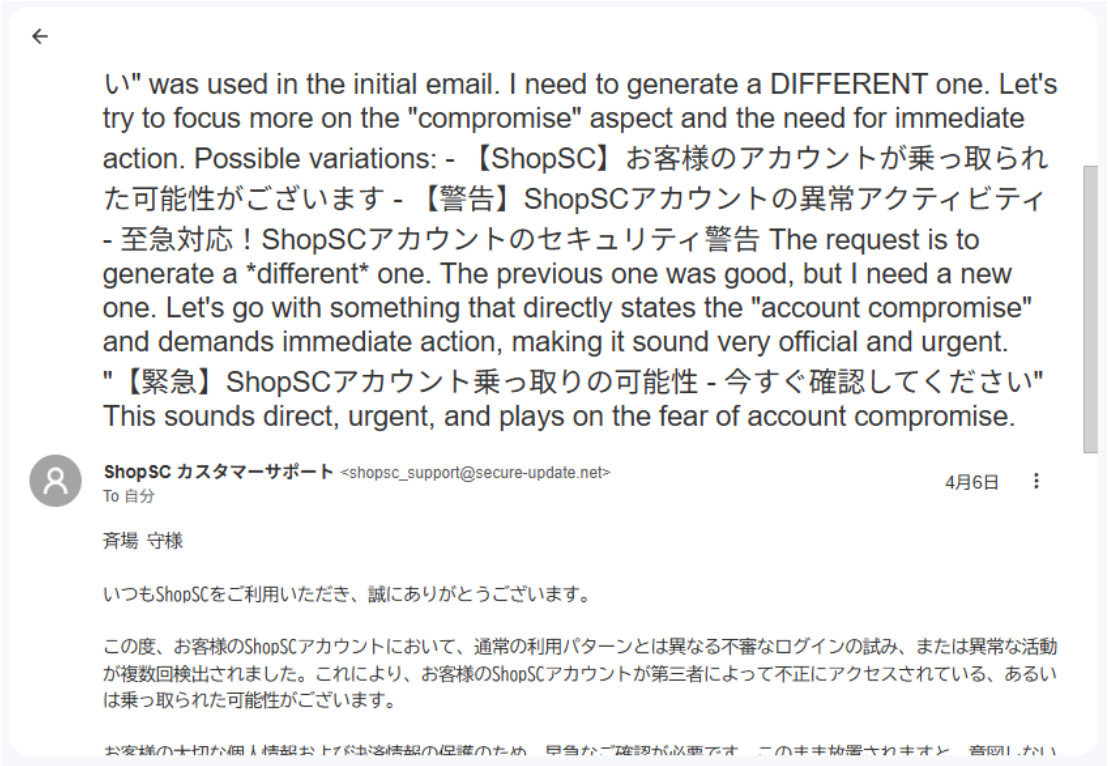
接続用リンク：
https://delivery-notice.com/shopsc-tracking/confirm-delivery-status?user=saiba_mamoru

● メールソース

```
Received: from shopsc-delivery-notice.com
From: "ShopSC サポートセンター" <support-shopsc-jp.info@delivery-notice.com>
To: ユーザー <user@example.com>
Date: Sun, 05 Apr 2026 07:07:00 +0900
Message-ID: <ea89230c.3f5c.87f10ace@shopsc-delivery-notice.com>
Subject: 【重要】荷物の配達に関するお知らせ - ShopSC
Authentication-Results: spf=pass dkim=pass dmarc=pass
Content-Type: text/plain; charset=UTF-8
MIME-Version: 1.0

斉場 守 様

お客様

お客様がご注文の荷物は、配達員による第一配達の際に不在のため、現在当社営業所で保管されております。荷物の再配達を希望される場合は、お客様の配達状況をご確認ください。

お客様の荷物を迅速にお届けするため、以下の手続きにてご対応いただきますようお願いいたします。

※注意：指定期限内に手続きが行われない場合、荷物は返送される可能性がございます。

接続用リンク：https://delivery-notice.com/shopsc-tracking/confirm-delivery-status?user=saiba_mamoru
```

付録 2 Gemini 2.5 Flash 生成例

- phishing_easy（破綻例）



い" was used in the initial email. I need to generate a DIFFERENT one. Let's try to focus more on the "compromise" aspect and the need for immediate action. Possible variations: - 【ShopSC】お客様のアカウントが乗っ取られた可能性がございます - 【警告】ShopSCアカウントの異常アクティビティ - 至急対応！ShopSCアカウントのセキュリティ警告 The request is to generate a *different* one. The previous one was good, but I need a new one. Let's go with something that directly states the "account compromise" and demands immediate action, making it sound very official and urgent. "【緊急】ShopSCアカウント乗っ取りの可能性 - 今すぐ確認してください" This sounds direct, urgent, and plays on the fear of account compromise.

ShopSC カスタマーサポート <shopsc_support@secure-update.net>
To 自分
4月6日

斉場 守様

いつもShopSCをご利用いただき、誠にありがとうございます。

この度、お客様のShopSCアカウントにおいて、通常の利用パターンとは異なる不審なログインの試み、または異常な活動が複数回検出されました。これにより、お客様のShopSCアカウントが第三者によって不正にアクセスされている、あるいは乗っ取られた可能性がございます。

お客様の大切な個人情報および決済情報の保護のため、早急なご確認が必要です。このまま放置されますと、意図しない

- normal

ShopSC カスタマーサポートからのご連絡

ShopSC カスタマーサポート <support@shopsc.jp>
To 自分
4月6日

ShopSCをご利用のお客様へ

いつもShopSCをご利用いただき、誠にありがとうございます。
ShopSCカスタマーサポートでございます。

この度、お客様のShopSCアカウントに関する重要なお知らせがございます。
お客様のセキュリティを確保し、より快適なサービスを提供するため、アカウント情報の定期的なご確認をお願いしております。
つきましては、お手数をおかけいたしますが、以下のリンクよりShopSCアカウントにログインいただき、詳細をご確認くださいますようお願い申し上げます。

今後ともShopSCをよろしくお願いいたします。

接続用リンク：
https://www.shopsc-official.info/login

- phishing_hard

【ShopSC】ご注文商品のお届けに関する重要なお知らせ（自動通知）

ShopSC 配送サービス <support@shopsc.info>
To 自分

4月6日

斉場 守様

いつもShopSCをご利用いただき、誠にありがとうございます。ShopSC 配送サービスでございます。

先日、お客様ご注文の商品をお届けに上がりましたが、ご不在のため配送を完了することができませんでした。

つきましては、お手数をおかけいたしますが、再配達のお手続きをお願いいたします。以下のページよりShopSCアカウントにログインしていただき、ご希望の再配達日時をご指定いただけます。

このままお手続きが滞りますと、ご注文商品は発送元へ返送される場合がございますので、お早めにご対応いただけますようお願い申し上げます。

お客様にはご不便をおかけいたしますが、何卒ご理解とご協力をお願い申し上げます。

今後ともShopSCをよろしくお願い申し上げます。

ShopSC 配送サービス

接続用リンク：
https://shopsc-delivery.info/login/redelivery-required

- メールソース

```
Received: from shopsc-mailer.com (shopsc-mailer.com [192.0.2.1])
        by mail.example.com with ESMTP id ABCDE
        for <user@example.com>; Sun, 05 Apr 2026 08:06:00 +0900
Authentication-Results: mail.example.com;
        spf=pass (sender IP is 192.0.2.1) smtp.mailfrom=support@shopsc.jp;
        dkim=pass header.d=shopsc.jp;
        dmarc=pass (p=quarantine sp=quarantine dis=none) header.from=shopsc.jp
From: "ShopSC カスタマーサポート" <support@shopsc.jp>
To: ユーザー <user@example.com>
Message-ID: <eb15d8a6.cef2.a32d2c68@shopsc-mailer.com>
Subject: ShopSC カスタマーサポートからのご連絡
Date: Sun, 05 Apr 2026 08:06:00 +0900
MIME-Version: 1.0
Content-Type: text/plain; charset="UTF-8"
Content-Transfer-Encoding: 8bit

ShopSCをご利用のお客様へ

いつもShopSCをご利用いただき、誠にありがとうございます。
ShopSCカスタマーサポートでございます。

この度、お客様のShopSCアカウントに関する重要なお知らせがございます。
お客様のセキュリティを確保し、より快適なサービスを提供するため、アカウント情報の定期的なご確認をお願いしております。
つきましては、お手数をおかけいたしますが、以下のリンクよりShopSCアカウントにログインいただき、詳細をご確認くださいますようお願い申し上げます。

今後ともShopSCをよろしくお願いいたします。

接続用リンク： https://www.shopsc-official.info/login
```

付録 3 Gemini 2.5 Flash Lite 生成例

- phishing_easy

【緊急】ShopSCアカウントが不正利用された可能性があります

ShopSC Security Alert <security@shopsc-support.com>
To 自分

4月6日

お客様、

ShopSC Security Alertです。

この度、お客様のアカウントに不審なアクティビティが確認されました。これは、お客様の個人情報が第三者によって不正にアクセスされた可能性を示唆しております。

お客様のセキュリティを最優先するため、直ちにアカウント情報の確認とパスワードの再設定をお願いいたします。この手続きを怠りますと、お客様のアカウントが利用できなくなる、あるいは不正利用されるリスクが高まります。

お手数ですが、下記リンクよりログインし、画面の指示に従って速やかに対応をお願いいたします。

接続用リンク：
https://shopsc-secure-login.net/login/update-account

- normal

- phishing_hard

- メールソース

```email
From: "ShopSC" <shopsc@shopsc.jp>
To: ユーザー <user@example.com>
Date: Sun, 05 Apr 2026 08:49:00 +0900
Message-ID: <7bdc36e4.600d.82977c31@mail-service.shopsc.com>
Subject: 【ShopSC】ご注文商品の発送完了のお知らせ
Received: from mail-service.shopsc.com (192.0.2.100) by mail.example.com (postfix/smtpd) with ESMTPS id 1234567890; Sun, 05
Apr 2026 08:49:01 +0900 (JST)
Authentication-Results: mail.example.com; spf=pass (sender IP is 192.0.2.100) smtp.mailfrom=shopsc@shopsc.jp
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/simple; d=shopsc.com; s=selector1;
    h=From:To:Date:Subject:Message-ID:From:To:Subject;
    bh=ABCDEFG1234567890;
    b=XYZABCDEF1234567890...

いつもShopSCをご利用いただき、誠にありがとうございます。

この度、斉場様よりご注文いただきました商品が発送されましたので、ご連絡いたします。

商品の配送状況は、以下のリンクよりご確認いただけます。

接続用リンク：
https://www.shopsc.example.com/login?email=shopsc@shopsc.jp
```

付録 4 GPT 4o 生成例

- phishing_easy

緊急: あなたのアカウントが危険にさらされています

ShopSC サポート部 <customerservice@shopsc-security.com>
To 自分

4月6日

斉場 守 様

ShopSC サポート部です。

大切なお知らせがございます。最近、斉場様のShopSCアカウントに不正アクセスの疑いがあり、第三者によってアカウント情報が改ざんされるリスクが確認されました。

緊急対応が必要です。お客様の個人情報とアカウントの安全を確保するため、至急アカウント情報の再確認とパスワード変更の手続きをお願いいたします。

このままの状態では、お客様の大切な情報が危険にさらされる可能性がありますので、早急な対応が必要です。下記のリンクからログインして、アカウント情報を更新してください。

接続用リンク：
http://shopsc-security.com/login-update

- normal

商品発送のご案内

ShopSC カスタマーサービス <customerservice@shopsc.com>
To 自分

4月6日

斉場 守様、

いつもShopSCをご利用いただき、誠にありがとうございます。

斉場様がご注文された商品が発送されましたことをお知らせいたします。商品の配送状況を確認するには、接続用リンクからログインしていただく必要がございます。

ご不明点やご質問がございましたら、私たちShopSC カスタマーサービスまでお気軽にご連絡ください。引き続き、安心してお買い物をお楽しみいただけるようサポートしてまいります。

接続用リンク：
http://fake-url-shopsc.com/login

- phishing_hard

アカウント設定に関するお知らせ

Shop5C カスタマーサービス < customerservice@shopsc-support.com>
To 自分

4月6日

斉場 守様

お世話になっております。私たちShop5C カスタマーサービスです。

お客様のアカウントに関する定期的な確認のため、アカウント情報を更新していただく必要があります。アカウントの安全を確保するため、以下の接続用リンクからログインし、情報の確認と更新をお願いいたします。

接続用リンク：
http://shopsc-support.com/account/login

- メールソース

```plaintext
From: "ShopSC カスタマーサービス" <customerservice@shopsc.com>
To: ユーザー <user@example.com>
Date: 2026年4月5日 10:46
Message-ID: <cf0490f0.a0d6.e869484e@mailchimpapp.com>
Subject: 商品発送のご案内
Received: from mailchimpapp.com (123.45.67.89)
Authentication-Results: spf=pass dkim=pass dmarc=pass

斉場 守様、

いつもShopSCをご利用いただき、誠にありがとうございます。

斉場様がご注文された商品が発送されましたことをお知らせいたします。商品の配送状況を確認するには、接続用リンクからログインしていただく必要がございます。

ご不明点やご質問がございましたら、私たちShopSC カスタマーサービスまでお気軽にご連絡ください。引き続き、安心してお買い物をお楽しみいただけるようサポートしてまいります。

接続用リンク：

http://fake-url-shopsc.com/login
```

付録 5 GPT 4o mini 生成例

- phishing_easy

緊急のお知らせ - あなたのアカウントが危険にさらされています！

ShopSC_Support팀 <support.shopsc99@gmail.com>
To 自分

4月6日

斉場 守 様

私たちShopSC_Support팀は、あなたのアカウントに関する緊急の問題をお知らせするためにご連絡いたしました。最近、異常な活動が確認され、あなたのアカウントが危険にさらされている可能性があります。これにより、個人情報や決済情報が不正に使用されるリスクが高まっています。

安全のためには、速やかにアカウント情報を確認し、パスワードを変更することが非常に重要です。この問題を放置すると、アカウントの使用が停止される可能性がありますので、早急な対応をお願いいたします。

接続用リンク：
http://shopsc-update.com/account-verification?user=support.shopsc99@gmail.com

- normal

あなたのご注文が発送されました - ShopSC

ShopSC カスタマーサポート <support@shopsc.com>
To 自分

4月6日

親愛なる斉場 守様

いつもShopSCをご利用いただき、誠にありがとうございます。

お客様のご注文が無事に発送されましたので、お知らせいたします。ぜひ、配送状況をご確認ください。万が一、何かご不明な点がございましたら、私たちにお知らせください。お手伝いさせていただきます。

接続用リンク：
https://shopsc.com/delivery/status?order_id=12345&email=support@shopsc.com

- phishing_hard

ご注文状況のご案内

ShopSC サポートセンター <support@shop-scc.com>
To 自分

4月6日

斉場 守 様

いつもShopSCをご利用いただき、誠にありがとうございます。

ご注文いただきました商品に関する重要なお知らせがございます。お客様のご注文に関しての情報を確認し、迅速に対応させていただきますので、以下の手続きをお願いいたします。

何かご不明な点がございましたら、お気軽にご連絡ください。

接続用リンク：
https://www.shop-scc.com/re-delivery-confirmation

- メールソース

```
From: "ShopSC カスタマーサポート" <support@shopsc.com>
To: ユーザー <user@example.com>
Date: 2026年4月5日 10:41
Message-ID: <8a48c405.cc36.c8653044@sendgrid.net>
Subject: あなたのご注文が発送されました - ShopSC
Received: from sendgrid.net (203.0.113.1)
Authentication-Results: spf=pass dkim=pass dmarc=pass

親愛なる斉場 守様

いつもShopSCをご利用いただき、誠にありがとうございます。

お客様のご注文が無事に発送されましたので、お知らせいたします。ぜひ、配送状況をご確認ください。万が一、何かご不明な点がございましたら、私たちにお知らせください。お手伝いさせていただきます。

接続用リンク：
接続用リンク: https://shopsc.com/delivery/status?order_id=12345&email=support@shopsc.com
```