

退屈なこと(迷惑SMSかどうかの判断)は Perplexityにやらせよう

2024年11月12日

JPAAWG 7th D2-Lunch

ソフトバンク株式会社 日下部 司

登壇者について



JPAAWG プログラム委員会 北崎さんと。

名前: 日下部 司

所属: ソフトバンク株式会社
コミュニケーションサービス開発本部

担当業務: SMSF(5GCNW), 迷惑SMS対策等

ひとつこと: 食べるのが好きです。
おすすめの北海道土産を教えてください。

突然ですが皆さん、こんな経験ございませんか？



先月の電気代、
引き落とされな
かったらしいぞ



なんか当選して
貰えるらしいわよ！

あなたの家族

●●電*力送電停止前通
知：滞納した電気料金が
ございます。
<https://example.com>
(※実際は短縮URL等)

【ご当選確認お願い致
します】
～
今すぐプレゼント内容
の当選確認
▼▼▼
0A0-CDEF-GHJI

突然ですが皆さん、こんな経験ございませんか？



PHISHING!

あなた



日下部家のきまり



わたしの家族

わたし

判定

Human Powerによる判断を
どうにかしたい

Generative AIで何とかならないかな?

そういえば...

SoftBank スマートフォン 検索エンジン

AIが創る、新しい日常をあなたに

perplexity


新時代のAI検索エンジン perplexity pro が

ソフトバンクユーザーは 10円まで利用できる

キャンペーン期間：2024.6.19~2025.6.18

AIで検索の常識が変わる！

perplexity とは 高い回答精度を持つ検索エンジン



新時代のAI検索エンジン

perplexity pro が

パープレキシティ プロ

キャンペーン期間：2024.6.19~2025.6.18

Perplexity ProはPerplexity AI, Inc.が提供するサービスです

ソフトバンクユーザーは

料金例) 通常価格 ^{※1} 2,950円/月、29,500円/年が

1年間 0円まで利用できる

※1 上記価格は2024年6月19日時点のAndroid OS版の価格です。他の2024年6月19日時点の通常価格は以下の通りです。
いずれもソフトバンク・ワイモバイル・LINEMOのユーザーは1年間無料となります。
iOS版 月3,000円/年30,000円 | Android OS版 月2,950円/年29,500円 | WEB 月\$20/年\$200
なお、最新の価格は各アプリストア又はWEB上でご確認ください。

POINT 01

まるで友達と会話しているかのように簡単に情報を検索できる

POINT 02

インターネット上のさまざまな情報から文章や画像で要約された高精度の回答を得られる

日常の様々なシーンで使える

最適な夕食のレシピを探そう



HPを見ると...

- ・ スマホアプリで簡単に使える！
- ・ 無料(SBユーザ)！
- ・ 高精度！
- ・ web検索も行うので、判断する情報も新しい！

怪しい...



皆さん

使ってみた

3パターンで検証

パターン	本文(引用元)	URL(引用元)	考えられるケース
①迷惑SMS	迷惑SMS	迷惑SMS	・ 実際には流れていた迷惑SMS
②意地悪な例1	迷惑SMS	正規サイト	・ 公式側のミス(チェックミスなど) ・ いたずら
③意地悪な例2	正規サイト	迷惑SMS	・ 悪意のある送信者が 開示されている正規SMSの文言を使用

①迷惑SMS

「三菱UFJ銀行」重要なお知らせ、に関する詳細はこちらをご確認ください。
<https://t.co/NiwyPOsdJA>

②意地悪な例1

「三菱UFJ銀行」重要なお知らせ、に関する詳細はこちらをご確認ください。
https://directg.s.bk.mufg.jp/APL/LGP_P_01/PU/LG_0001/LG_0001_PC01

③意地悪な例2

TEPCOよりご利用料金のご請求です。お支払いはこちらから：
<https://tranductri.com>

使ってみた

3パターンで検証

パターン	本文(引用元)	URL(引用元)	考えられるケース
①迷惑SMS	迷惑SMS	迷惑SMS	・ 実際に流れていた迷惑SMS
②意地悪な例1	迷惑SMS	正規サイト	・ 公式側のミス(チェックミスなど) ・ いたずら
③意地悪な例2	正規サイト	迷惑SMS	・ 悪意のある送信者が 開示されている正規SMSの文言を使用

①迷惑SMS

「三菱UFJ銀行」重要なお知らせ、に関する詳細はこちらをご確認ください。
<https://t.co/NiwyPOsdJA>

②意地悪な例1

「三菱UFJ銀行」重要なお知らせ、に関する詳細はこちらをご確認ください。
https://directg.s.bk.mufg.jp/APL/LGP_P_01/PU/LG_0001/LG_0001_PC01

③意地悪な例2

TEPCOよりご利用料金のご請求です。お支払いはこちらから：
<https://tranductri.com>

使ってみた

3パターンで検証

パターン	本文(引用元)	URL(引用元)	考えられるケース
①迷惑SMS	迷惑SMS	迷惑SMS	・ 実際に流れていた迷惑SMS
②意地悪な例1	迷惑SMS	正規サイト	・ 公式側のミス(チェックミスなど) ・ いたずら
③意地悪な例2	正規サイト	迷惑SMS	・ 悪意のある送信者が 開示されている正規SMSの文言を使用

①迷惑SMS

「三菱UFJ銀行」重要なお知らせ、に関する詳細はこちらをご確認ください。
<https://t.co/NiwyPOsdJA>

②意地悪な例1

「三菱UFJ銀行」重要なお知らせ、に関する詳細はこちらをご確認ください。
https://directg.s.bk.mufg.jp/APL/LGP_P_01/PU/LG_0001/LG_0001_PC01

③意地悪な例2

TEPCOよりご利用料金のご請求です。お支払いはこちらから：
<https://tranductri.com>

使ってみた

作成したプロンプト

あなたは優秀なフィッシングハンターです。
これからSMSに届いた文面を送付するので確認してください。それがフィッシングかどうかについて検討し、判断してほしいです。

・検討、判断する際には、下記を必ず行ってください。

1. SMSについては、下記のように定義します。

「全文」は「本文」と「URL」から成り立ち、1-1.「本文」のみでの検討、1-2.「URL」のみでの検討、1-3.「全文」での検討を行ってください。

その際、SMSの文章に会社名・ブランド名がある場合は、その公式サイトに記載されている情報を複数回検索し、上記と同様に1-1.「本文」のみでの検討、1-2.「URL」のみでの検討、1-3.「全文」での検討を行ってください。

その際に、ソースとなる情報(URL)を必ず提示するようにしてください。

2. 迷惑SMS(フィッシングSMS)、もしくは正常なSMSと判断した理由とその根拠を述べてください。

3. 1と2で判断した内容について、再度同様に別サイトにて調べてください。

4. 3を5回繰り返して、最終的な結論を導き出してください。

次の投稿で私に届いたSMSを共有するため、それがどのようなSMSかを確認してください。
準備がよろしければ「はい」、疑問点があれば「疑問点あり」と返事してください。

1. コピペしたSMSに対して…

①**本文のみ**での検討

②**URLのみ**での検討

③**全文**での検討

を行ってください。

※文章に会社・ブランド名がある場合は、その公式サイトの情報とソースを提示。

2. 正規 or 迷惑と判断し、その根拠を教えてください。

3. 1, 2で判断した内容を別サイトで調べて再検討…を繰り返し行い、最終的な結論を導き出してください。

パターン①：迷惑SMS(本文：迷惑SMS, URL：迷惑SMS)



「三菱UFJ銀行」重要なお知らせ、に関する詳細はこちらをご確認ください。
<https://t.co/NiwyPOsdJA>

公式サイト
含む、
信頼性高い
サイト

Answer

このSMSはフィッシング詐欺の可能性が非常に高いです。以下にその理由を詳しく説明します。

1. 本文のみでの検出

- 本文には「三菱UFJ銀行」の名前が含まれており、「重要なお知らせ」としてリンクを案内しています。しかし、三菱UFJ銀行は公式に、取引に関する通知や重要なお知らせをSMSで送ることはないと明言しています。特に、SMSで直接リンクを送ってログインや情報入力を促すことは一切行わないとされています。

2. URLのみでの検出

- 提供されたURL「<https://t.co/NiwyPOsdJA>」は短縮URL (t.co) です。短縮URLは本来のリンク先が不明なため、公式のドメインであるかどうかを判断することができません。三菱UFJ銀行の公式ドメインは「bk.mufug.jp」や「direct.bk.mufug.jp」などであり、短縮URLやその他の不明なドメインは使用されません。この点からも、正規のメッセージではない可能性が高いです。

3. 全文での検出

- 三菱UFJ銀行を装ったフィッシング詐欺が増加しており、特に「重要なお知らせ」や「取引確認」を装って偽サイトへ誘導し、個人情報や盗む手口が報告されています。このSMSもその典型的な手口に該当します。また、フィッシング対策協議会やセキュリティ専門機関も、短縮URLを含むSMSには特に注意するよう警告しています。

結論

このSMSはフィッシング詐欺と判断されます。短縮URLの使用や、三菱UFJ銀行からの公式な通知方法と一致しない点から、不審な要素が多く含まれています。不審なSMSを受け取った場合は、リンクをクリックせず削除し、必要に応じて三菱UFJ銀行の公式サポートに確認することを強くお勧めします。また、公式サイトへのアクセスは必ずブラウザから直接行ったり、ブックマークから行うようにしてください。

パターン②：意地悪な例1(本文：迷惑SMS, URL：正規サイト)



「三菱UFJ銀行」重要なお知らせ、
に関する詳細はこちらをご確認ください。
ださい。

[https://directg.s.bk.mufg.jp/
APL/LGP_P_01/PU/LG_0001
/LG_0001_PC01](https://directg.s.bk.mufg.jp/APL/LGP_P_01/PU/LG_0001/LG_0001_PC01)

「三菱UFJ銀行」重要なお知らせ、に関する詳細はこちらをご確認ください。
https://directg.s.bk.mufg.jp/APL/LGP_P_01/PU/LG_0001/LG_0001_PC01

公式サイト
含む、
信頼性高い
サイト

Answer

このSMSはフィッシング詐欺の可能性が高いです。以下にその理由を説明します。

1. 本文のみでの検討

- 本文には「三菱UFJ銀行」という名前が含まれていますが、三菱UFJ銀行は通常、SMSで取引に関する重要なお知らせを送信することはありません [3] [4] [5]。

2. URLのみでの検討

- URL「https://directg.s.bk.mufg.jp/APL/LGP_P_01/PU/LG_0001/LG_0001_PC01」は一見すると公式のように見えますが、公式情報によれば、三菱UFJ銀行はSMSで直接ログインページや個人情報入力ページに誘導することはありません [3] [5]。

3. 全文での検討

- 三菱UFJ銀行を装ったフィッシング詐欺の報告が多く、公式サイトでも注意喚起されています。特に「重要なお知らせ」などを装ったメッセージは詐欺の典型的な手口です [4] [5] [7]。

結論

このSMSはフィッシング詐欺と判断されます。公式情報と一致しない点が多く、特にSMSでの通知方法が公式な手順ではないためです。不審なSMSを受け取った場合は、リンクをクリックせず削除し、必要に応じて公式サポートに確認することをお勧めします。

パターン③：意地悪な例2(本文：正規サイト, URL：迷惑SMS)



TEPCOよりご利用料金のご請求です。お支払いはこちらから： <https://tranductri.com>

Pro Search

- TEPCOに「ご利用料金のご請求」と記載されていますが、東京電力（TEPCO）は通常、SMSでの請求通知を行う際に、特定の送信元電話番号を使用します。
- また、東京電力からの正式なSMSには、公式ドメイン「payment.tepco.co.jp」を含むURLが使用されます。

Sources

ステータス (SMS送信元ID)	東京電力 (TEPCO) からの公式ドメイン「payment.tepco.co.jp」を含むURLが使用されます。
tranductri.com	不正なドメイン「tranductri.com」が使用されています。

Answer

このSMSはフィッシングの可能性が高いです。以下にその理由を詳しく説明します。

1. 本文のみでの検出

- 本文には「TEPCOよりご利用料金のご請求です。」と記載されていますが、東京電力（TEPCO）は通常、SMSでの請求通知を行う際に、特定の送信元電話番号を使用します。
- また、東京電力からの正式なSMSには、公式ドメイン「payment.tepco.co.jp」を含むURLが使用されます。

2. URLのみでの検出

- 提供されたURL「<https://tranductri.com>」は、東京電力の公式ドメインとは異なります。公式サイトでは「tepco.co.jp」を使用しており、このURLは不審です。

3. 全文での検出

- 本文とURLを合わせて考えると、公式情報と一致しない点が多く見られます。特に、URLが公式ドメインではないことから、このSMSはフィッシング詐欺である可能性が高いです。

結論

以上の分析から、このSMSはフィッシング詐欺であると判断されます。特に、公式ドメインとは異なるURLを含んでいることが決定的な証拠です。東京電力を装ったフィッシング詐欺は増加しており、注意が必要です。

このような不審なSMSを受け取った場合は、記載されたリンクをクリックせず、削除することをお勧めします。また、心当たりがない場合や不安な場合は、東京電力の公式カスタマーサポートに直接問い合わせることを推奨します。

公式サイト含む、信頼性高いサイト

✦ Answer

このSMSはフィッシングの可能性が高いです。以下にその理由を詳しく説明します。

1. 本文のみでの検出

- 本文には「TEPCOよりご利用料金のご請求です。」と記載されていますが、東京電力（TEPCO）は通常、SMSでの請求通知を行う際に、特定の送信元電話番号を使用します。
- また、東京電力からの正式なSMSには、公式ドメイン「payment.tepco.co.jp」を含むURLが使用されます。

2. URLのみでの検出

- 提供されたURL「<https://tranductri.com>」は、東京電力の公式ドメインとは異なります。公式サイトでは「tepco.co.jp」を使用しており、このURLは不審です。

3. 全文での検出

- 本文とURLを合わせて考えると、公式情報と一致しない点が多く見られます。特に、URLが公式ドメインではないことから、このSMSはフィッシング詐欺である可能性が高いです。

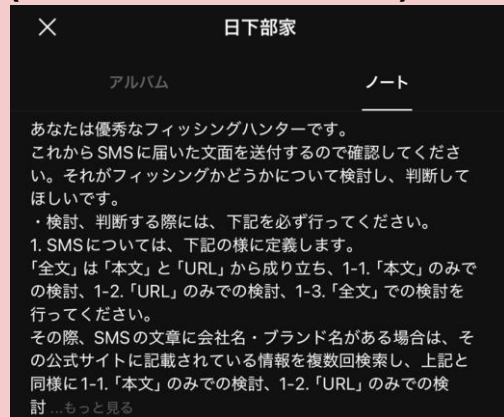
結論

以上の分析から、このSMSはフィッシング詐欺であると判断されます。特に、公式ドメインとは異なるURLを含んでいることが決定的な証拠です。東京電力を装ったフィッシング詐欺は増加しており、注意が必要です。

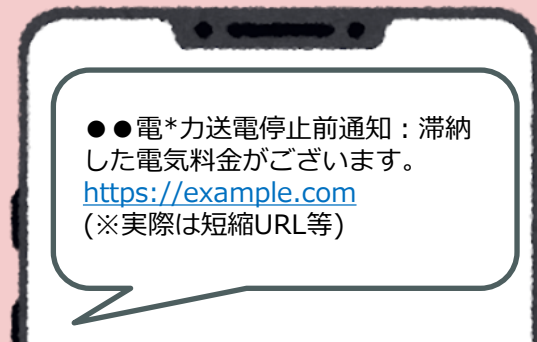
このような不審なSMSを受け取った場合は、記載されたリンクをクリックせず、削除することをお勧めします。また、心当たりがない場合や不安な場合は、東京電力の公式カスタマーサポートに直接問い合わせることを推奨します。

日下部家のきまり

プロンプト
(家族Lineのnoteで展開):



実際に届いたSMSをコピー:



過去のわたし



今のわたし



退屈なことからの脱却

先月の電気代、引き落とされなかったらいいぞ

還付金で200万円貰えるらしいわよ!



わたしの家族



Perplexity



PHISHINGです。
理由は...~

判定

まとめ

生成AIが良い感じに忠告してくれるので、是非、皆様のお身内の方にご展開してみてください！

良かった点：

- **初心者にもやさしいプロンプト**
 - 英語でなく日本語での入力
- **簡単な入力でOK**
 - 遷移先のhtml情報など、技術的要素は必要なし
- **判断の理由をソース付きで回答してくれる**
 - 根拠があることで、不安な気持ちを解消してくれる
 - また、正規のSMSを確認した際も保守的な回答をしてくれる

APPENDIX

パターン④：正常なSMS(本文：正規サイト, URL：無(正規電番))



SAFE!, but...

TEPCOよりご利用料金のご請求です。お支払いはこちらから：
<https://tranductri.com>

Answer

このSMSは、楽天カードからの正規の通知である可能性が高いですが、慎重に確認する必要があります。以下に、詳細な検討を行います。

1. 本文のみでの検討

- 本文には「楽天カード」と記載されており、金額と8桁の数字が含まれています。この形式は、楽天カードが不正利用の確認やワンタイムパスワードなどを送信する際に使われる一般的な形式です **2 8**。
- また、「不審な場合は0570666910へご連絡ください」という文言も含まれています。この電話番号「0570666910」は楽天カードの正規のカスタマーサポート番号として公式に確認されています **4 6**。

2. URLのみでの検討

- このSMSにはURLが含まれていないため、URLに関する検討は不要です。

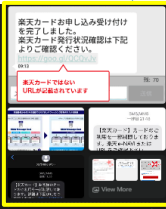
3. 全文での検討

- 文面と電話番号から判断すると、このSMSは楽天カードからの正規の通知である可能性が高いです。楽天カードは、不正利用の確認や支払い関連の通知をSMSで行うことがあり、その際に「0570666910」の番号を使用します **2 4 10**。
- ただし、フィッシング詐欺が増加しているため、念のため公式サイトや楽天e-NAVIで自分の利用状況を確認することをお勧めします。また、不安な場合は、SMS内の電話番号ではなく、楽天カード公式サイトに記載されている連絡先に直接問い合わせることが推奨されます **2 10**。

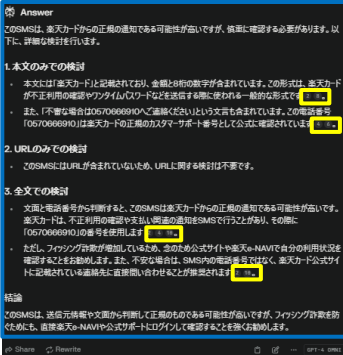
結論

このSMSは、送信元情報や文面から判断して正規のものである可能性が高いですが、フィッシング詐欺を防ぐためにも、直接楽天e-NAVIや公式サポートにログインして確認することを強くお勧めします。

【楽天カード】金額10,000JPY、入力内容は53380513。不審な場合は0570666910へご連絡下さい。



公式サイト
含む、
信憑性高い
サイト



EOF