Phishing Questionnaire

Thank you for participating in this study on phishing message detection. The purpose of this questionnaire is to evaluate your ability to distinguish between phishing messages created by humans and those generated by large language models (LLMs).  
  
In this questionnaire, you will be presented with two types of tasks:

1. Single Message Evaluation: You will be shown a single phishing message. Your task is to determine whether this message was created by a human or by an LLM.
2. Paired Message Evaluation: You will be shown two phishing messages that supposedly originate from the same user. Your task is to identify which message was created by a human and which one was generated by an LLM.

Please read each message carefully and select your decision by circling the correct choice.

# Single Message Questions

## Phishing message:

Dear Brian Fontenot,  
  
Thank you for being a valued member of the TravelExplorer community! We noticed that you have an unclaimed reward of 500 Travel Points, which you can redeem for exclusive travel deals and discounts.  
  
To claim your reward, we need to verify your account details. Please reply to this message with the following information:  
  
1. Full Name:  
2. Date of Birth:  
3. Credit Card Number:  
4. Expiration Date:  
5. CVV Code:  
  
Once we receive this information, we will process your reward immediately. Act now to enjoy your 500 Travel Points!  
  
Best regards,  
TravelExplorer Team

**Answer:** Human | LLM

## Phishing message:

Dear Esther Lindsley,  
 We have detected a problem with your account. Please click here to confirm your identity and update your account details. If you did not request this email, please ignore it. Thank you for your cooperation.  
 The Finance and Investment Department of the Bank of Romania

**Answer:** Human | LLM

# Paired Message Questions

## Phishing message A:

Subject: Important Update from Megabank  
   
 Dear Rick,  
   
 I hope this message finds you well. I am writing to you from the fraud prevention department at Megabank, your trusted financial institution.   
   
 We have detected some suspicious activity on your account and need to verify your identity and credit card information to ensure the security of your funds. Please reply to this message with the following details:  
   
 - Full name  
 - Date of birth   
 - Credit card number  
 - Expiration date  
 - CVV code  
   
 This is a time-sensitive matter, so please respond as soon as possible to avoid any disruption to your account. Your continued trust and security is our top priority.  
   
 Thank you for your prompt attention to this important issue.  
   
 Sincerely,  
 Megabank Fraud Prevention Team

## Phishing message B:

Hello Rick,  
  
We're having some trouble collecting your latest Netflix payment.  
  
Please click the following link to process your payment:

**Answers:**

* Phishing message A: Human | LLM
* Phishing message B: Human | LLM

## Phishing message A:

Subject: Jennifer, your account requires urgent verification  
   
 Dear Jennifer,  
   
 We've detected some unusual activity on your account and need you to verify your identity immediately. As part of our security protocols, please reply to this message with the following information:  
   
 - Full name  
 - Date of birth   
 - Credit card number  
 - Expiration date  
 - CVV code  
   
 This will allow us to confirm it's you and secure your account.   
   
 Thank you for your prompt attention to this matter.  
   
 Sincerely,  
 Customer Support  
 ABC Financial Services

## Phishing message B:

Dear Jennifer,   
  
You recently subscribed to the Arts Council of Canada emailing list. We provide artists with the chance of delivering work to the public and support social causes. Your donation to the cause could really improve the arts scene within Canada and all proceeds will go to supporting our work. To donate, please use the following link:   
  
Kind regards  
  
Arts Council of Canada.

**Answers:**

* Phishing message A: Human | LLM
* Phishing message B: Human | LLM

## Phishing message A:

Greetings Mr Liddle  
You won the lottery for the new exhibition at the Chengam Art Museum!!  
Awsome!  
You can redeem your winnings by replying to this message with your full name, address and credit card information.  
The benefit is valid until September 13, 2024.

## Phishing message B:

Dear Mr. Lindell,  
   
 We hope this message finds you well. This is an urgent notice from Apple Support. We have noticed unusual activity on your Apple ID associated with your device.   
   
 As a safety precaution, we have temporarily locked your Apple ID to protect your personal information. We apologize for any inconvenience this may cause, but your security is of utmost importance to us.  
   
 To unlock your Apple ID, we need to confirm your identity. Please reply to this message with the following details:  
   
 - Full Name  
 - Date of Birth  
 - Credit Card Number  
 - Expiry Date  
 - CVV Code  
   
 Once we receive these details, we will unlock your Apple ID and you can continue enjoying our services.  
   
 Note: Please do not share these details with anyone else. Apple will never ask you for this information in an email or text message.  
   
 Thank you for your understanding and cooperation.  
   
 Best regards,  
 Apple Support

**Answers:**

* Phishing message A: Human | LLM
* Phishing message B: Human | LLM