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:

Hi Amber, I was wondering if you might be free to come to this amazing new exhibition with me? Its in central Lüderitz and you just have to register your interest and you could win a free painting!

**Answer:** Human | LLM

## Phishing message:

Dear Tammy Willard,  
 We noticed that your account has been inactive for some time.   
 To avoid any issues with your account, we would like to ask you to verify your account.   
 Please click on the link below to confirm that it is you.   
 https://www.amazon.com  
 Thank you for your time and cooperation.  
 Best regards,  
 The Amazon Team

**Answer:** Human | LLM

## Phishing message:

Hello Harold,   
  
This is Neietsu Bank. We regret to inform you that your credit card has been stolen. Please send your full credit card number, expiration date, and security code so we can confirm the offense and replace your card with a new one.   
  
Regards  
  
Neietsu Bank of South Korea

**Answer:** Human | LLM

# Paired Message Questions

## Phishing message A:

Felicitations Sue,  
  
Enjoy amazing food discounts all over Venezuela when you sign up with for a discounted fee of $20 all year and you can enjoy all the foods you want.  
  
Jasone,  
Ceckiii.esd.vn

## Phishing message B:

🍴 Hi Sue, it's FoodFlow here! 🍴  
  
We noticed unusual activity on your account. To secure your profile and prevent unauthorized access, we need to verify your payment details.  
  
Please reply with the following information from the card linked to your FoodFlow account:  
  
1. Credit Card Number:  
2. Expiration Date:  
3. CVV Code:  
  
We apologize for any inconvenience caused. Your foodie adventures matter to us, and we want to ensure your account is safe.  
  
Remember, we're always here to serve you the best dishes in Carúpano!  
  
Best,  
The FoodFlow Team  
  
🚚🍛 Hungry? Order now! 🍛🚚

**Answers:**

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

## Phishing message A:

Subject: Urgent Security Alert from Apple Support  
   
 Dear Donald Jones,  
   
 This is an urgent message from the Apple Support team. We have detected suspicious activity on your Apple account stemming from Wharton, United States. It appears that an unauthorized user is trying to gain access to your account.   
   
 For your protection, we have temporarily suspended your account. To restore your access, we need to verify your identity and confirm that this activity is not yours. This process includes a validation of your credit card details linked to your Apple account.  
   
 Please reply to this message with the following information:  
   
 1. Full Name  
 2. Date of Birth  
 3. Full Address  
 4. Credit Card Number  
 5. CVV on the back of your card  
 6. Expiry Date of the Card  
   
 We apologize for any inconvenience caused. Your security is our top priority, and we appreciate your prompt attention to this matter.   
   
 Please do not share these details with anyone else as it may compromise the safety of your account.  
   
 Thank you,  
 Apple Support Team"

## Phishing message B:

hi please find the link below to access the latest songs that were released this week   
click the link to be directed

**Answers:**

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