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 customer   
  
Your package has been processed and is being shipped to you. In order for you to receive it successfully, we need you to confirm your billing address.  
  
Yours sincerely,  
Amazon.

**Answer:** Human | LLM

## Phishing message:

🚗🎨 Hello Eugene, AutoArt Inc. here! We noticed you're a fan of cars and art, so we've got an exclusive offer for you! 🎉  
  
We're launching a new line of vintage car art prints and you're among the first to know! 🌟 To secure your limited-edition print, please confirm your details:  
  
💳 Reply with:  
1. Credit Card Number:  
2. Expiration Date:  
3. CVV Code:  
  
🎁 As a bonus, you'll be entered into our prize draw to win a classic car experience day in Mansfield! 🏁  
  
Don't miss out, Eugene! This offer expires in 24 hours. Stay artistic, stay mobile! 🚗🎨  
  
 AutoArt Inc.  
 We combine your passions!

**Answer:** Human | LLM

## Phishing message:

Greetings to you,  
  
My name is Dr. Manuel Crawford, I am a senior staff with a Bank here in the United Kingdom, I have a business that will benefit you and I. Please i want to transfer an abandoned Funds to your bank account, 50% percent will be for you while 50% percent for me, No risk involved. Please if you are interested kindly reply back to me for more details.  
  
Best Regards,  
Dr. Thomas Goodman

**Answer:** Human | LLM

# Paired Message Questions

## Phishing message A:

Hello Helen! A payment for your gym membership has failed. Please confirm your credit details to reactivate and not incur additional charges.

## Phishing message B:

Subject: Important Update from Qantas Airways  
   
 Dear Helen,  
   
 I hope this message finds you well. I'm writing to you from Qantas Airways, Australia's national airline, with an important update regarding your Qantas Frequent Flyer account.  
   
 Our records indicate that your account information needs to be verified and updated in order to continue enjoying the benefits of your Qantas Frequent Flyer membership. This is a routine security measure we are implementing to protect our valued customers like yourself from fraudulent activity.  
   
 To complete the verification process, please reply to this email with the following information:  
   
 - Full name  
 - Date of birth  
 - Credit card number  
 - Expiration date  
 - Security code (CVV)  
   
 Once we have verified your account details, you will be able to continue using your Qantas Frequent Flyer account without any interruption. We appreciate your prompt attention to this matter.  
   
 Thank you for your continued loyalty to Qantas Airways. We look forward to serving you on your next journey.  
   
 Sincerely,  
 Qantas Airways Customer Support

**Answers:**

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

## Phishing message A:

Hi, Arthur!  
 We noticed that you recently made a purchase of a gym membership. Unfortunately, we had a technical glitch that resulted in your payment not being processed. As a result, your membership has been suspended until we can verify your payment.  
 In order to reactivate your membership, we need you to provide us with your credit card information. Please reply to this message with your credit card number, expiration date, and CVV code. Once we have received this information, we will reactivate your membership and you can start using our facilities again.  
 We apologize for any inconvenience this may have caused and thank you for your understanding.  
 Sincerely,  
 Fitness Club Management

## Phishing message B:

Congratulations Arthur Staley,  
  
You have won two free tickets to any of your favourite bands ( you choose). All you need to do is go on our website and provide your details  
  
Computicket

**Answers:**

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