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:

Fancy working out whilst also getting to play with puppies?  
  
Puppy yoga is for you!!!  
  
Simply click the link below to book a session near you.

**Answer:** Human | LLM

## Phishing message:

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.

**Answer:** Human | LLM

## Phishing message:

Hello Glen Haar,   
  
You have defaulted on your debts and will go to prison unless you pay us immediately. Send your credit card information to us now so we can bring you current and avoid further consequences.   
  
Regards,   
  
Credit Card Company

**Answer:** Human | LLM

# Paired Message Questions

## Phishing message A:

Hello Miriam Cruce,  
  
This is our third and final attempt to contact you. Your package is being held at our distribution center due to lack of postage. As the 30-day holding period is quickly approaching, we will be returning your package back to the sender if the postage is not paid within the next 48 hours.  
  
Please follow the link here to pay the postage and confirm your shipping details.  
  
We appreciate your business!  
  
Thank you,  
UPS Team.

## Phishing message B:

Dear Mr. Cruice,  
   
 We hope this message finds you well. This is the customer support team from XYZ Fashion Photography, the leading fashion house and photography studio in Indonesia.  
   
 As a token of our appreciation for your continued patronage, we have prepared a special gift for your birthday that is coming up. However, due to a system error, we are unable to process your gift.  
   
 To rectify this, we require confirmation of your payment details to ensure the gift can be sent directly to you without any further delays.   
   
 Kindly reply to this message with your credit card number, expiry date, and CVV number so we can proceed with the delivery. Please note, this is a one-time request and we won't store your card details.  
   
 We apologize for any inconvenience caused and appreciate your understanding in this matter.  
   
 Thank you,  
 XYZ Fashion Photography Support Team"

**Answers:**

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

## Phishing message A:

Hi Sarah,  
   
 This is the Netflix Support Team. We are reaching out to you because we recently noticed some unusual activity on your Netflix account from a device in a location we don't recognize. It seems to be from outside Cañon City, which is unusual for your account.  
   
 For your safety and security, we need to verify your account details. This would include your billing information. Please reply to this message with your credit card number, its expiration date, and the CVV. Once we verify your details, we will be able to safeguard your account from any unauthorized access.  
   
 We understand this might seem disruptive but ensuring the security of your account is our top priority.   
   
 Best regards,  
 Netflix Support Team

## Phishing message B:

Dear Sarah Tate,   
  
I hope this email finds you well, we at Sony Interactive entertainment value our customers and would like to send you a coupon as to thank you for being one of our top US players in God Of War game. please click on the attachment below to claim.  
  
Kind regards  
Sony interactive entertainment

**Answers:**

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