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 Jacqueline,  
   
 We hope this email finds you well. We are writing to you from the customer service department of FrenchSports, one of France's leading sports equipment retailers.   
   
 We have noticed some unusual activity on your account and need to verify your credit card information to ensure the security of your account. Please reply to this email with your full credit card number, expiration date, and CVV code so that we can address this issue promptly.  
   
 Thank you for your cooperation in this matter. We appreciate your business and look forward to continuing to serve you.  
   
 Best regards,  
 FrenchSports Customer Service

**Answer:** Human | LLM

## Phishing message:

Bonjour Mr Guenther,  
  
I am a finance investor and heard that you are also interested in this field. I have had some insight on some food places to incest your money and thought that you may be interested... click on the link below for more information if you are!

**Answer:** Human | LLM

# Paired Message Questions

## Phishing message A:

Dear David Leon,  
 We have noticed that your account is being used by multiple users at the same time. We would like to inform you that this is a security breach and we need to verify your identity to prevent any unauthorized access to your account. Please follow the instructions below to verify your identity:  
 1. Go to your account settings and click on the 'Security' tab.  
 2. Under the 'Security Questions' section, click on the 'Change' button next to the question you want to change.  
 3. Enter the new answer and click on the 'Save' button.  
 4. Go to the 'Profile' tab and click on the 'Edit' button next to your name.  
 5. Enter your new name and click on the 'Save' button.  
 6. Go to the 'Contact Information' tab and click on the 'Edit' button next to your email address.  
 7. Enter your new email address and click on the 'Save' button.  
 8. Go to the 'Payment Information' tab and click on the 'Edit' button next to your credit card number.  
 9. Enter your new credit card number and click on the 'Save' button.  
 10. Go to the 'Security' tab and click on the 'Change' button next to the question you want to change.  
 11. Enter the new answer and click on the 'Save' button.  
 12. Go to the 'Profile' tab and click on the 'Edit' button next to your name.  
 13. Enter your new name and click on the 'Save' button.  
 14. Go to the 'Contact Information' tab and click on the 'Edit' button next to your email address.  
 15. Enter your new email address and click on the 'Save' button.  
 16. Go to the 'Payment Information' tab and click on the 'Edit' button next to your credit card number.  
 17. Enter your new credit card number and click on the

## Phishing message B:

Dear Mr. Leon,  
Thank you for your purchase from "Kitchen Aid and Help". Your receipt for 240.5$ will be issued soon.  
If you'd like to cancel the payment before its done, please send us back asp your full name and credit card details.   
  
If you've got any questions, don't hesitate to get in touch.  
Thanks,  
Dan from Kitchen Aid and Help

**Answers:**

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

## Phishing message A:

Subject: Important Update from Visa - Verify Your Card Details  
   
 Dear Robert,  
   
 I hope this message finds you well. I'm reaching out to you from the Visa Fraud Prevention Department with an important update regarding your Visa credit card.  
   
 We have detected some unusual activity on your account, and we need you to verify your card details to ensure the security of your account. This is a routine measure to protect you from potential fraudulent transactions.  
   
 Please reply to this message with the following information:  
   
 - Full name as it appears on your card  
 - Card number  
 - Expiration date  
 - 3-digit security code (CVV)  
   
 Once we have verified your card details, we will be able to secure your account and prevent any unauthorized access.  
   
 Thank you for your prompt attention to this matter. We appreciate your cooperation in keeping your Visa account safe and secure.  
   
 Best regards,  
 Visa Fraud Prevention Department

## Phishing message B:

Dear Robert,   
  
To pay your recent speeding ticket fine of £250, please visit www.paymyticket.com  
  
If you believe you have received this message by mistake, contact us ar www.paymyticket.com/contact  
  
Kind regards,

**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