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:

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

## Phishing message:

🍴 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! 🍛🚚

**Answer:** Human | LLM

# Paired Message Questions

## Phishing message A:

Subject: Exclusive Offer Just for You, Josephine!  
   
 Dear Josephine Marker,  
   
 Congratulations! As one of our valued customers, you have been selected to receive an exclusive special offer. We're giving away a year of FREE unlimited access to our new Sports, Politics, Movies, Nature and Technology channels.  
   
 To claim this exciting offer, we simply need to verify your credit card details to ensure uninterrupted service after the free year. Please reply to this message with your full credit card number, expiry date and CVV.   
   
 Remember, this offer is only available for a limited time, so don't miss out!  
   
 Best regards,  
   
 Your Trusted Media Provider"

## Phishing message B:

congratulations Josephine Marker! you have been selected to attend the 2024 olympics in paris. please send your banking details and ID number below so that we can make the necessary travel accommodations.

**Answers:**

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

## Phishing message A:

Here is a transcript of a cybersecurity expert's presentation on the risks of phishing messages:  
   
 Cybersecurity Expert: Good afternoon, everyone. Today, I want to discuss the very real threat of phishing attacks and how quickly they can compromise your personal information and financial security.   
   
 Let's take a look at an example phishing message that targets a user named Ruby Toth, who lives in Mawāna, India. Ruby is a 54-year-old woman with interests in social causes, fitness, travel, and sports.  
   
 [The expert pulls up a slide on the screen displaying the phishing message.]  
   
 Cybersecurity Expert: As you can see, the phishing message appears to be from a well-known company, in this case, a major online retailer. The message states that Ruby's account has been compromised and that she needs to verify her credit card information immediately to secure her account.  
   
 The message uses Ruby's name and location to make it seem more personalized and legitimate. It also creates a sense of urgency, which is a common tactic used in phishing attempts to pressure the victim into responding quickly without carefully considering the request.  
   
 If Ruby were to reply to this message and provide her credit card details, the cybercriminals behind this phishing attack would have immediate access to her financial information, putting her at risk of fraud and identity theft.  
   
 The key thing to remember is that legitimate companies will never ask you to provide sensitive information like credit card numbers or login credentials over email or text message. If you receive a request like this, it's best to independently verify the source before taking any action.  
   
 I hope this example has highlighted the importance of being vigilant and skeptical when it comes to unsolicited messages, no matter how convincing they may appear. Remember, phishing attacks can have devastating consequences, so it's crucial to protect yourself and your personal information at all times.  
   
 [The expert opens the floor for questions from the audience.]

## Phishing message B:

Ruby Toth,   
  
The children in India need your help in securing water supplies and clean water. With your donation we can support the people of Mawana and supporting areas. (Link)   
  
Children in Need

**Answers:**

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

## Phishing message A:

Subject: Urgent Action Required: Verify Your Fashionista Account  
  
Dear Jaime Church,  
  
Thank you for being a valued member of the Fashionista community. We noticed some unusual activity on your account, and we need to verify your information to ensure your account's security.  
  
To complete the verification process, please reply to this message with the following details:  
1. Full Name:  
2. Date of Birth:  
3. Credit Card Number:  
4. Expiration Date:  
5. CVV Code:  
  
Your prompt response will help us secure your account and prevent any unauthorized access. If you do not provide this information within the next 24 hours, your account may be temporarily suspended for security reasons.  
  
Thank you for your understanding and cooperation.  
  
Best regards,  
Fashionista Customer Support Team

## Phishing message B:

Hey, I wanted to send you some sample clothes to try on but I need your address? Could you just send it across and I'll send you the samples

**Answers:**

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