Journal of Cybersecurity Education, Research and Practice

Volume 2022 | Number 2 Article 2

January 2023

Alpha Phi-shing Fraternity: Phishing Assessment in a Higher Education Institution

Marco Casagrande
University of Padua, marco.casagrande@eurecom.fr

Mauro Conti
University of Padua, mauro.conti@unipd.it

Monica Fedeli
University of Padua, monica.fedeli@unipd.it

Eleonora Losiouk
University of Padua, eleonora.losiouk@unipd.it

Recommended Citation

Casagrande, Marco; Conti, Mauro; Fedeli, Monica; and Losiouk, Eleonora (2023) "Alpha Phi-shing Fraternity: Phishing Assessment in a Higher Education Institution," *Journal of Cybersecurity Education, Research and Practice*: Vol. 2022: No. 2, Article 2.
DOI: 10.32727/8.2023.1
Available at: https://digitalcommons.kennesaw.edu/jcerp/vol2022/iss2/2

Alpha Phi-shing Fraternity: Phishing Assessment in a Higher Education Institution

Abstract

Phishing is a common social engineering attack aimed to steal personal information. Universities attract phishing attacks because: 1) they store employees and students sensitive data, 2) they save confidential documents, 3) their infrastructures often lack security. In this paper, we showcase a phishing assessment at the University of Redacted aimed to identify the people, and the features of such people, that are more susceptible to phishing attacks. We delivered phishing emails to 1.508 subjects in three separate batches, collecting a clickrate equal to 30%, 11% and 13%, respectively. We considered several features (i.e., age, gender, role, working/studying field, email template) in univariate and multivariate analyses and found that students are more susceptible to phishing attacks than professors or technical/administrative staff, and that emails designed through a spearphishing approach receive a highest clickrate. We believe this work provides the foundations for setting up an effective educational campaign to prevent phishing attacks not only at the University of Redacted, but in any other university.

Keywords

user behaviour, phishing, social engineering

Cover Page Footnote

We thank the whole University of Redacted, from the people that proposed and supported this phishing awareness initiative to the IT staff that helped us with its deployment. Finally, we thank all the randomly selected participants whose contribution was fundamental to our work.

INTRODUCTION

Phishing is the attempt at stealing assets (e.g., money, intellectual property) and confidential data (e.g., credentials, sensitive health data, personal identity) employing both social engineering and technical subterfuges [1].

Social engineering is the manipulation of unwary victims into believing that the social engineer is a trusted and legitimate party. Common technical subterfuges used in phishing are the delivery of deceptive emails, the redirection to counterfeit websites and the planting of malware into the victim's computer.

Phishing attacks are characterized by a huge payoff (e.g., gaining money, stealing sensitive data) with respect to the limited effort and technical knowledge required. They often act as the first step of a large scale attack: 93% of data breaches are achieved through phishing [2], and 71% of criminal groups rely on emails for malware distribution [3].

Nowadays, phishing attacks are one of the most significant threats worldwide [4], with nearly 90% of organizations being targeted by spearphishing and business email compromise attacks in 2019. A notable mention goes to the global spearphishing campaign targeting organizations involved with the distribution of COVID-19 vaccine [5].

Phishing aims at manipulating the victim in completing a specific action, that might lead to stolen sensitive data (e.g., user credentials, financial data, medical information), tampered devices (e.g., installing malware, taking remote control of a machine) or money loss (e.g., ransomware). The attacker achieves this goal by designing an email with the following features: (i) it should look trusted (i.e., the higher the trust level, the higher the chance to make the attack successful); (ii) it should leverage on the victim's emotions; (iii) it should use the most appropriate language to arouse feelings in the victim.

The most common approach followed by the attacker is to make the victim: (i) directly disclose sensitive data; (ii) follow a link in the email and submit sensitive data to a malicious website; (iii) download and execute a malicious attachment (i.e., malware); (iv) grant unwarranted privileges to the attackers; (v) complete some kind of unwarranted physical action (wire transfer). The content of a phishing email can involve many topics [6], but it usually refers to: e-commerce transactions (e.g., Amazon, eBay, PayPal [7]), delivery service impersonation [8], online services (e.g., Google [9], iTunes [10], Dropbox [11]), communications sent by a person known by the victim [12] or by an authority [13].

Phishing attacks can be classified into several categories. Among those, spearphishing refers to phishing emails that are customized for a specific target of people, that the attackers have monitored to collect information (e.g., software licenses, employee timetables and routines, commercial partners, room disposition, internal events and workshops). Spearphishing is much more effective than generic phishing, since the custom information acquired about the victims builds the trust level.

Phishing attacks can target any kind of organization, from companies to

universities. In particular, phishing attacks against academic parties can lead to significant economical and intellectual property losses, as demonstrated by the following case studies.

The Silent Librarian [14] threat actor set up methodical spearphishing attacks against educational institutions and obtained unauthorized access to computer systems, stole proprietary data and sold the same data to Iranian customers, including the Iranian Government and Iranian universities. University email templates, addresses, websites and branding were counterfeited to redirect victims into fake university library login pages. Between 2013 and 2017, Silent Librarian caused approximately a 3.4 billion-dollar intellectual property loss and compromised 8.000 university accounts.

The MacEwan University in Canada lost 11.8$ million in payments to an illegitimate banking account, impersonating a legitimate construction company [15]. The fraud originated from an online phishing scam. The scammers sent fake emails from 14 local construction companies, asking for payments related to a recently completed job. Prestigious university email accounts hijacked by the cybercriminals became part of later phishing attempts, adding further repercussions to the incident.

Three employees of the Wichita University fell for a phishing scam, asking for their university ID and password [16]. The attackers then compromised the university's payroll system and rerouted the paychecks to their own bank accounts.

The Lancaster University in England was victim of a sophisticated phishing attack leading to data breach [17, 18]. Undergraduate student applicant data records (e.g., name, home address, phone number, email address, ID documents) for 2019 and 2020 were accessed. Furthermore, the attackers deployed another phishing attack targeting the emails addresses leaked by the previous attack, soliciting payments from students on behalf of the university.

An analysis from the INKY email security company found more than 3.000 emails sent by compromised accounts from Oxford University, Stanford University and Purdue University [15].

In this work, we conduct a phishing assessment at the University of Redacted to identify which people, and in particular which features of such people, are more susceptible to phishing attacks. The University of Redacted is one of the largest universities in Italy and it currently counts 60.000 enrolled students, 2.300 active professors and 2.300 active technical/administrative staff members. In our phishing assessment, we involved 1.508 individuals randomly chosen among students (524 subjects), professors (492 subjects), and technical/administrative staff (492 subjects). We designed five different *email templates*, following three separate *attack methodologies*, to be sent as phishing emails to the participants of the study.

We used Gophish [19] to set up the assessment and monitor the clickrate from our participants. Overall, we sent three batches of emails in different months, for a total of three emails received by each subject. In the first batch of emails, the 30% of participants clicked on the malicious link embedded in the phishing email, while the other two batches achieved lower clickrates, 11% and

13%, respectively. In the statistical analysis, we considered the impact of a set of independent variables (i.e., age, gender, role, working/studying field, email template) against the phishing susceptibility. We found that the *role* and *email template* individually impact on the phishing susceptibility, while the *age* and *email template* together have an impact on phishing susceptibility.

Contributions. The contributions of this work are as follows:

- We set up a phishing assessment on a large public institution, such as the University of Redacted, aimed at identifying the predictors to phishing susceptibility and involving different categories of participants, such as students, professors and technical/administrative staff members.
- We perform an in-depth statistical analysis to identify which features affect, either individually or simultaneously, the susceptibility to phishing attacks of a victim.
- We find that *email template* and *age* have a statistically significant relationship with phishing susceptibility. The chi-squared tests show an association relationship between email template and phishing susceptibility in all three batches (p-values lower than 0.5), between age and phishing susceptibility in Batch 1 and 3 (p-values lower than 0.5). The multivariate analysis shows similar results.
- We find that younger people are more susceptible to any form of phishing. The binary logistic regression analysis shows how participants aged 31-99 have lower odds of clicking our malicious link. For example, participants aged between 61-99 have lower odds of being susceptible compared to age 0-20 (log odds change of -2.0477 in Batch 1).
- We find that employees are more susceptible to spearphishing, compared to phishing. The binary logistic regression analysis shows how employees receiving spearphishing emails (template 4 and 5) are more susceptible to them, than to generic phishing emails (template 1 and 2), with a log increase in odds of 1.6728 (template 4 - professors) and 0.7648 (template 5 - technical/administrative staff).

Approval and Ethics. We received direct authorization and approval of the phishing assessment from the university rector. We organized two meetings with Data Protection Officer and the Data Processing committee for experiments. We agreed on the information we would receive to perform statistical analysis while preserving the privacy of the participants. We warned the IT department before sending out each batch of phishing emails. We followed the guidelines in the Belmont Report [20] and the recommendations made by Finn and Jacobsson in [21]. The email templates were designed with the goal of avoiding any disruption of the participants' everyday life, and we only impersonated fictional individuals. At the end of the experiment, we organized a public online seminar both as an educational event and a debriefing process.

Organization. The rest of this publication is structured as follows. section Background covers the background knowledge needed to understand the concepts beyond phishing and its unique features. The related work in section Related Work summarizes the outcomes of previous phishing evaluations, with a specific focus on the number of participants and the chosen methodology. section Methods describes our methodology for conducting a phishing assessment at the University of Redacted, and for evaluating the impact of subjects' demographics against phishing susceptibility. In section Implementation, we implement our methodology in our phishing experiment. In section Results and Discussion, we analyze and discuss the results of our work. Finally, section Conclusion provides a short summary and a conclusion to the study.

BACKGROUND

In this section, we describe the different attack vectors used to deliver a phishing attack (section Attack Vectors).

Attack Vectors

In the 96% of social incidents, the main vectors used by phishing scams are emails. Concerning malware-related phishing, emails are the vector in about the 92% of attempts, followed by websites which are used in about the 6% [2].

Emails

Emails are a common vector for phishing attempts, since users check their email account at work and during private time. When using emails as an attack vector for phishing attacks, adversaries can choose among different strategies, such as spoofing the email address of the sender, attaching a malicious payload or including a malicious link in the email content. Email spoofing involves the forgery of the sender address and other metadata. Standard email protocols, primarily Simple Mail Transfer Protocol (SMTP), are not equipped with authentication mechanisms. Thus, forging the email metadata is quite easy. Nowadays, methods for email authentication are available, but come along with significant trade-offs that make them unusable. In addition, the spoofed email address usually refers to a well-known third-party, thus making the phishing email detection even harder. Another possible approach is to embed a malicious payload into the email attachment, so that users install a malware by themselves. Payloads can be embedded into a wide range of common file types and phishing emails usually send documents that need to be reviewed urgently, thus persuading the victim into downloading and executing them. Common malicious attachments are scripts (".vbs", ".js"), executable files (".exe") and documents (".doc",

".pdf", ".html", ".xml"). Microsoft Office macros are popular and very dangerous, since they can directly download malware from the Internet. Once installed, a malware can: (i) encrypt the device data and ask for a ransom; (ii) log every keystroke and action performed on the device; (iii) join a botnet; (iv) help the malware spread to other devices on the network and to other email contacts. Finally, another successful approach involves external links. When the sender is trustworthy, users click on external links rather mindlessly. Unfortunately, links can redirect users to another prominent phishing attack vector: malicious websites.

Websites

Websites are an essential part of the Internet, and a perfect follow-up to a phishing email. Attackers usually try to emulate an authentication procedure to steal users' credentials. Possible approaches include cloning popular websites or exploiting the features of Hypertext Transfer Protocol Secure (HTTPS). Cloning a popular website allows to achieve a high trust level, thus making an unsuspecting victim perform unwarranted actions. The purpose of cloning a popular website is often coupled with "Credential Harvesting", which refers to stealing data (usernames, emails, passwords) to obtain unauthorized access to devices and networks. Effective baits are login pages to Facebook, Gmail, Amazon and Outlook Web Access. Online forms, used to contact the technical support or to subscribe to a service, are also particularly vicious baits. Visually speaking, it is impossible to spot a cloned web page and the insertion of a clever redirection mechanism makes it even harder. HTTPS relies on a bidirectional encryption, protecting data while traveling between client and server. At the end of 2017, two out of three websites loaded by Firefox used HTTPS [22]. HTTPS web pages feature a distinctive green lock icon (or other clear indicators in a similar fashion) near the address bar, raising the sense of trust and security in users. However, a common misconception is considering HTTPS web pages trustworthy, when HTTPS encryption just means "protection from third-parties" and it does not provide any information about the website owner legitimacy. Thus, attackers are creating phishing HTTPS web pages to exploit this misconception.

RELATED WORK

Due to the widespread risk to become a victim of a phishing attack, several studies have already tried to assess the susceptibility of people to phishing attacks. We selected from the state-of-the-art the set of papers sharing one or more of the following characteristics with our study: similar aim (evaluating phishing susceptibility), similar context and dimension (public institution with more than 1.000 participants), similar data collection and data analysis methodology (univariate and multivariate analyses on the data collected from

the phishing and spearphishing assessments). Table 1 summarizes our analysis of such papers by considering the following criteria: *aim*, *target context*, *participants number*, *data collection methodology*, *data analysis*.

Paper	Target Context	Participants Number	Data Collection Methodology	Data Analysis
J. P. Magalhães et al. [23]	Organization	∼700	Phishing assessment	Clickrate
W. D. Kearney et al. [24]	Organization	∼1.700	Phishing assessment	Clickrate
R. Wash et al. [25]	University	2.000	Phishing assessment	Clickrate
W. J. Gordon et al. [26]	Healthcare	5.416	Phishing assessment	Clickrate
T. Bakhshi [27]	Corporation	49	Spearphishing assessment	Clickrate
W. Flores et al. [28]	Corporation (employees)	92	Phishing and spearphishing assessment	Clickrate
R. Dodge et al. [29]	Military	∼4.000	Spearphishing assessment	Clickrate
A. Mihelic et al. [30]	Corporation	391	Spearphishing assessment	Clickrate
A. J. Burns et al. [31]	Corporation	400	Spearphishing assessment	Chi-squared test
D. D. Caputo et al. [32]	Corporation	1.369	Spearphishing assessment	Chi-squared test
S. McElwee et al. [33]	Corporation	∼1.000	Phishing assessment	Agency theory and linear regression
W. Li et al. [34]	University	6.938	Phishing and spearphishing assessment	Multi-level model
Our study	**University**	**1.508**	**Phishing and spearphishing assessment**	**Chi-squared test and binary logistic regression**

Table 1: Comparison among the existing case studies focusing on phishing susceptibility.

In several previous works [23, 24, 25, 26, 27, 28, 29, 30], the authors evaluate the results of their phishing assessment campaign only considering the clickrate. They prepared standard email templates (phishing) or used existing emails (spearphishing) that redirect the victim to a landing page, which may or may not harvest credentials. They sent a few rounds of emails and then compared clickrates to evaluate the participants improvement. The results range from less than 10% to over 90% clickrates, with a considerable difference in percentage even when the context of the experiments are similar. Drawing conclusions from just the clickrate is not a reliable methodology.

The educational side of phishing awareness is addressed by [31, 32], where the authors deliver training messages mixing different communication techniques: gain-framed (benefits of security) vs loss-framed (danger of phishing), individual-focused (gain/loss for yourself) vs other-focused (gain/loss for others). Gain-framed messages highlight to the user the benefits of being security aware, instead, loss-framed messages highlight the dangers and negative consequences of a successful attack. Individual-focused messages highlight gain and losses for the user himself, instead, other-focused messages highlight gain and losses for people near the user (e.g., family, co-workers). After a Chi-squared test analysis, both works concluded that the communication technique adopted in a training message had little to no influence on the effectiveness of phishing awareness training.

Other research angles included in our related works try to utilize specific

theories and models. The authors of [35] assert the viability of equal-variance signal detection theory (EVSDT) for human detection of phishing emails, by measuring detection accuracy and response bias. The authors of [34] evaluate phishing susceptibility predictors (age, gender, email type) through multivariate statistical analysis, concluding that age and email type were the most significant predictors.

Our work designs and performs a phishing assessment towards employees and students of the University of Redacted. With respect to the state of the art, our study is the first one that involves a large university and relies on both univariate and multivariate analysis to identify the best phishing susceptibility predictors for different target categories (i.e., professors, technical / administrative staff members, students). Through our work, we contribute to advance behavioural studies in this area, by identifying the most significant demographic features to evaluate phishing susceptibility.

METHODS

In this section, we discuss our methodology. In particular, we describe the phishing assessment plan (section Phishing Assessment Plan), the set of involved participants and their demographics (section Participants), and the designed email templates for all the attack methodologies (section Email Templates). We conclude by illustrating the set of statistical analyses chosen for processing the data collected through the phishing assessment (section Statistical Analysis).

Phishing Assessment Plan

Fig. 1 shows the structure of our study. We divided it into three phases: Preparation, Phishing Campaigns, and Statistical Analysis. During the Preparation, we setup Gophish (our phishing assessment tool) with the necessary data.

First, the anonymized email addresses we received from the university. Second, the various phishing templates we designed and crafted for the three *attack methodologies* we evaluate (i.e., *phishing targeting the University of Redacted*, *generic phishing* and *spearphishing*).

Then, in the Phishing Campaigns phase, we deploy several phishing campaigns that differ in their attack methodology and in the participant subset they target. We archive the outcome of every email sent in a database.

Finally, in the Statistical Analysis phase, we perform statistical analysis on the results using the Chi-Squared Test and the Binary Logistic Regression.

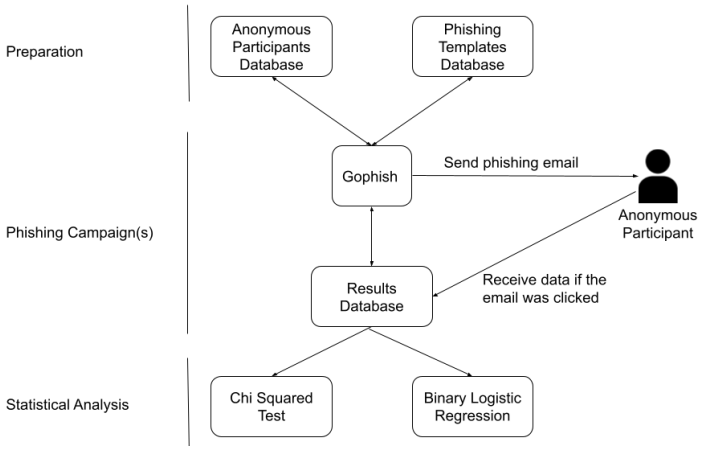


Figure 1: Structure of our study, that includes three phases: Preparation, Phishing Campaigns, and Statistical Analysis.

Attack Methodologies

In our phishing assessment, we considered three *attack methodologies* (i.e., *phishing targeting the University of Redacted*, *generic phishing* and *spearphishing*) expanded to five *email templates* as shown in Fig. 2.

To design the email templates we labeled as *phishing targeting the University of Redacted*, we considered real phishing attempts sent to the IT Department of the University of Redacted. This set of attempts all have poor grammar, include very suspicious URLs, do not include details about their target and are not customized for their target.

To design the email templates we labeled as *generic phishing*, we replicated emails sent by popular online services (e.g., Google, Amazon, Dropbox, PayPal). The content appears legitimate at a first glance, by showing official logos and signatures and by using proper grammar and formal tone, but do not make use of any specific information of their targets.

Finally, to design the email templates we labeled as *spearphishing*, we exploited the targets' role (i.e., student, professor, technical/administrative staff). Spearphishing requires more effort for the attacker, but it is more likely to be successful. Spearphishing emails present a well-crafted message, with small details that instill trust in the victims, use a tone consistent with the context of the message, and possibly exploit personal information. Ultimately, our email templates involve: real phishing emails targeting the University of Redacted (i.e., Template 1), generic phishing emails (i.e., Template 2); emails specifically customized for students (i.e., Template 3); emails specifically customized for professors (i.e., Template 4); emails specifically customized for technical/administrative staff (i.e., Template 5).

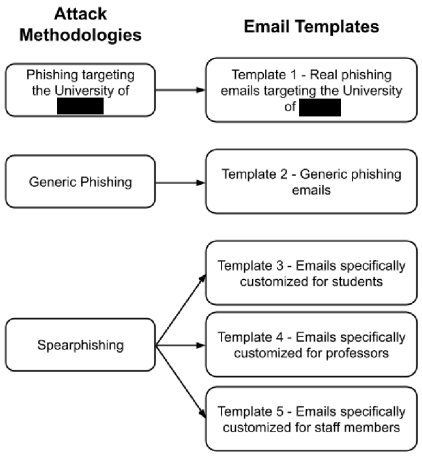


Figure 2: Mapping between attack methodologies and the email templates used for the phishing assessment.

Participants

Concerning the participants involved in our phishing assessment, we requested to the University of Redacted the following information: email address (anonymized), age, gender, role (i.e., "Student", "Professor", "Technical/Administrative Staff"), working/studying field (i.e., "Department" if student, "Macro-area" if professor, "Job Description" if technical/administrative staff).

Email Address. Each email address was anonymized, so we could not infer the identity of any participant. The anonymization was particularly relevant due to the standard format of the email accounts at the University of Redacted: "firstname.lastname.number@domain". Receiving them in clear text would definitely break our privacy constraints, as it would disclose the full identity of the participant.

Age. To perform smoother statistical analysis, we divided the dataset into ten age ranges ("lower than 20", "21-25", "26-30", "31-35", [...], "56-60", "higher than 60").

Gender. The information related to the gender of the participants ("Female" or "Male").

Role. We categorized the participants into "Student", "Professor" and "Technical/Administrative Staff". During the assessment, we relied on this information to send to each participant the email template specifically tailored for his role. We also included researchers and Associate Professors in the "Professor" role, and we included librarians in the "Technical/Administrative Staff" role.

Working/studying field. Students could be further divided into one of twenty-two Departments. Unfortunately, the number of different Departments diluted too much our data to provide meaningful results in our statistical analysis and thus were discarded. Professors could be further divided into one of

three Macro-areas (i.e., Life Science, Pure Science, Social and Human Science). Technical/Administrative Staff could be further divided into administrative or technical staff. In order to make our phishing attacks even more realistic, we asked for students' samples to be extracted in proportion to their Department distribution (the Department with most students would have the most entries in our dataset). Likewise, professors' samples were extracted in proportion to their Macro-Area distribution, and staff members' samples were extracted in proportion to their Job Description distribution.

Email Templates

We designed our email templates with the following rules in mind:

- Bring the least amount of inconvenience to participants.
- No attachments to be downloaded.
- No attempts at stealing data.

Bring the least amount of inconvenience to participants. Phishing messages might attack the victims through distasteful means (i.e., making threats, causing grief, giving false hope). On the contrary, we intended to spread cybersecurity awareness by causing the least degree of inconvenience possible to the participants. Furthermore, we carefully avoided impersonating any existing individual or entity, as it could affect their reputation.

No attachments to be downloaded. Phishing emails are frequently delivered with malicious attachments, that can be quite harmful when downloaded and executed. We were concerned about the reaction of participants to such an attack, so we just avoided it.

No attempts at stealing data. Phishing emails attempt to steal data by asking directly for confidential information or by redirecting the victim to some malicious website. Our goal as an attacker impersonator was to persuade the participant into clicking our "malicious" link. Instead of taking them to a login page, and have them submit their data, our link would take them to an educational landing page.

Statistical Analysis

We formalized our statements about the expected risk and the impact of education throughout a process of statistical analysis, backed by meaningful metrics. To portrait the situation at the University of Redacted, we designed the following analyses:

- An overview about the phishing assessment, with raw numbers and percentages concerning the outcome.
- Chi-Squared Tests for Independence to identify any association between our independent variables (i.e., age, gender, email template, role, working/studying field) and the phishing susceptibility.

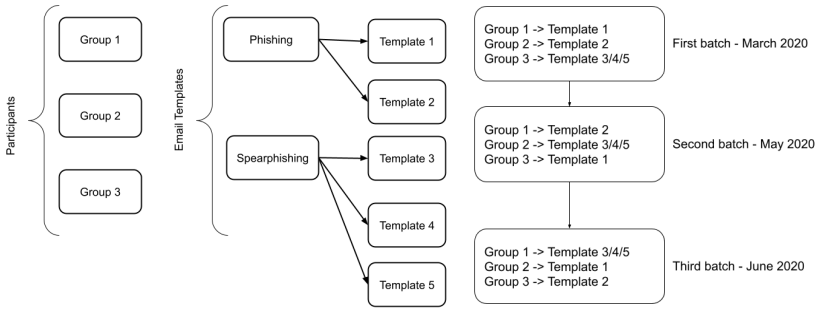
- Binary Logistic Regression models, as predictive models that predict the phishing susceptibility of an individual by considering his specific traits.

IMPLEMENTATION

In this section, we discuss the implementation of our methodology. In particular, we describe the phishing assessment plan (section 5.1), the anonymized participant dataset (section 5.2), the email templates we implemented (section 5.3), the statistical analyses we performed over the collected data (section 5.4), and the tools utilized for the assessment (section 5.5).

Phishing Assessment Plan

Participants were divided into three groups, as shown in Fig. 3a. Each participant received a total of three emails (i.e., one email with Template 1, one email with Template 2 and one email with Template 3/4/5 according to whether the participant was a student/professor/staff member). Fig. 3b shows which and when a specific group received a specific email template.



(a) Participants groups and scenarios templates. (b) Email delivery plan.

Figure 3: Overview of the phishing assessment plan.

The batches were spaced out to lower the alertness of participants. Thus, the emails were sent on March, May and June 2020. Additionally, we delivered all the emails with the same template before moving to the next one to avoid word-of-mouth. We only monitored clicks to the links embedded in our emails during the first week after the sending. We did not consider any data concerning the clicks that were performed in later weeks. The emails from the *Template 1* and *Template 2* were sent out at regular intervals throughout all day and night, even during weekends. The spearphishing emails from *Template 3, 4 and 5* were sent from 7 AM to 6 PM, adopting the same daily schedule of employees and students, and no messages were sent during the weekend. Spearphishing often

exploits specific timings to improve its effectiveness, and we wanted to emulate this behaviour in our assessment.

Anonymized Participant Dataset

Our campaign dataset contained the following information:

- Participant anonymized email address (as their identifier).
- Participant age.
- Participant gender.
- Participant role.
- Participant working/studying field (according to the role).
- Email template.
- Reaction (if the participant opened the link inside the phishing email).

Campaign Dataset			
Age		**Gender**	
< 20	*5% (71/1508)*	Female	*51% (776/1508)*
21-25	*21% (323/1508)*	Male	*49% (732/1508)*
26-30	*7% (108/1508)*	**Role**	
31-35	*6% (91/1508)*	Students	*34% (524/1508)*
36-40	*8% (116/1508)*	Professors	*33% (492/1508)*
41-45	*10% (153/1508)*	TA Staff	*33% (492/1508)*
46-50	*12% (183/1508)*		
51-55	*14% (208/1508)*		
56-60	*9% (142/1508)*		
> 60	*8% (113/1508)*		

Table 2: Campaign dataset - Age, Gender and Role.

We were given a dataset of 5.000 students, 500 professors and 500 technical/administrative staff members. From such dataset we randomly extracted 1.508 anonymized email addresses. During our phishing assessment, we sent emails to a grand total of 1.508 participants. Tables 3 show the distribution of the participants based on their age, gender, role and working field.

In our statistical analysis, we considered *age*, *gender*, *role*, *working/studying field* and *email template* as our "indepedent variables" or "predictors" and the *reaction* as "outcome" or "depedent variable". For the rest of the paper, we will often refer to the value of a particular independent variable (e.g., Age = "26", Category = "Student") by calling it as "trait". For example, a female 40-year old professor has the following traits: Gender "Female", Age "40" and Role "Professor".

Campaign Dataset			
Working Area (Professors)		**Working Area (TA Staff)**	
Life Science	*34% (165/492)*	Administrative	*65% (318/492)*
Pure Science	*41% (204/492)*	Technical	*35% (173/492)*
Human and Social Science	*25% (123/492)*		

Table 3: Campaign dataset - Working Field.

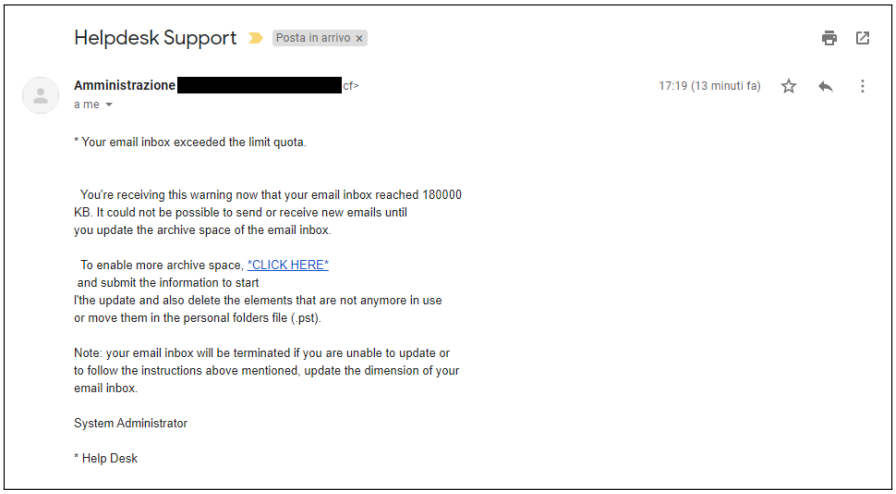


Figure 4: Batch 1 - Template 1.

Email Templates

In this section, we show some samples of email templates used in the first batch of delivery. Most of them were originally delivered in Italian, but here we provide the English translation.

Template 1. The Template 1 shown in Fig. 4 is a clone of a scam email targeting the University of Redacted at the end of January 2020. This template was sent to all the participants to alert them about their full email quota and to ask them to take an action by clicking the link. The original text of the phishing message was left unaltered.

Template 2. The Template 2 shown in Fig. 5 replicates common phishing scams that impersonate online service providers. This template was sent to all the participants to alert them about a new access to their Google account from a new, unregistered device. Users were strongly encouraged to check their recent activities by clicking a button. We removed all personal information contained in the original template, without disrupting the general message.

Template 3. The Template 3 shown in Fig. 6 is a common exam registration email sent by the University of Redacted mailing services. This template was sent only to the ”Student“ role participants to inform them about an exam

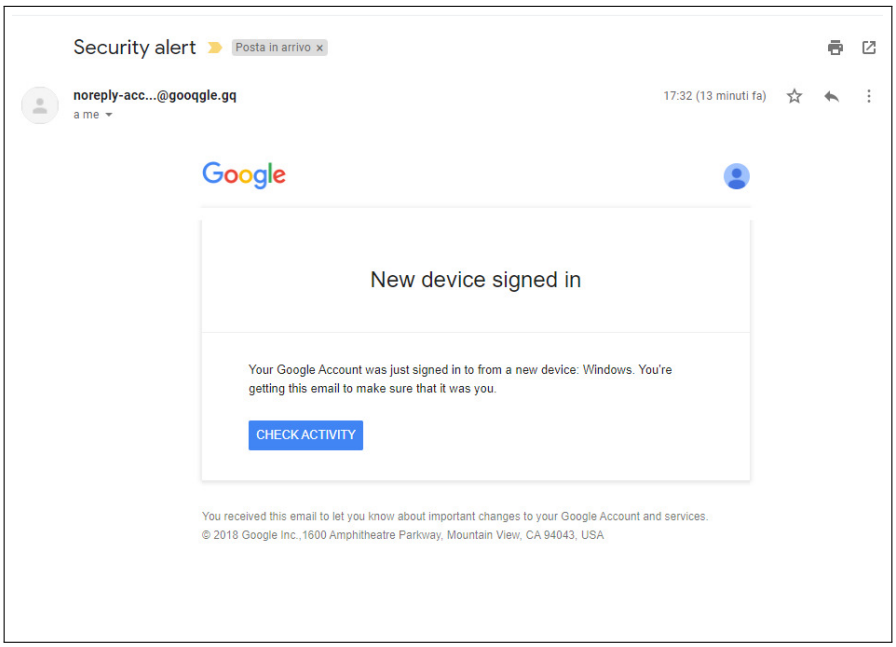


Figure 5: Batch 1 - Template 2.

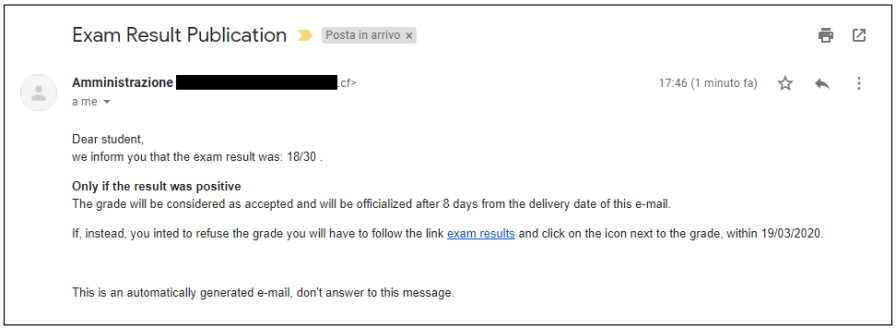


Figure 6: Batch 1 - Template 3 (students).

grade equal to 18/30. The grade was set to 18/30 because positive grades (>18) are automatically registered and added into the students career, if not actively rejected. This way, students were more compelled to follow the link and refuse the grade. We removed any reference to the student name and course related to the exam from the original email.

Template 4. The Template 4 shown in Fig. 7 simulates a request to share a scientific paper and it was sent only to the "Professor" role participants. The links in plain sight looked like well-known research websites and made the emails more legitimate. The English language made the short request more plausible.

Template 5. The Template 5 shown in Fig. 8 was inspired by a standard

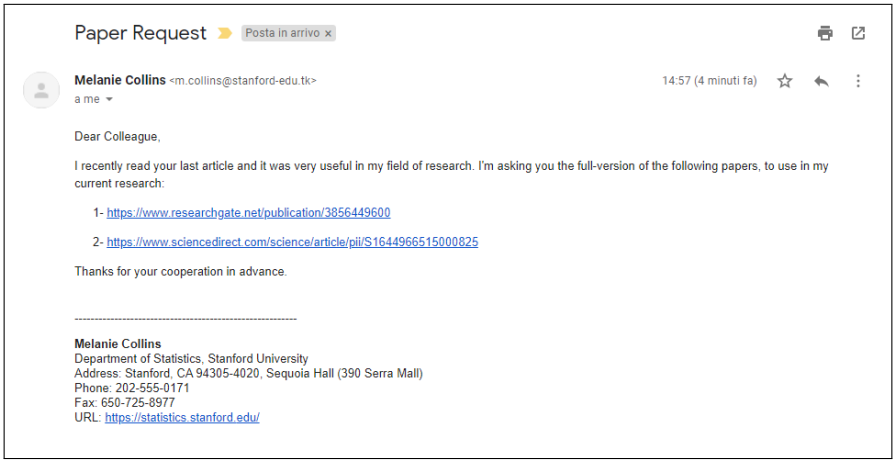


Figure 7: Batch 1 - Template 4 (professors).

Google Drive file sharing message. This template was sent only to "Technical/Administrative Staff" role participants. The message contained a link to a shared directory, which might lure the victim to click on it. The official University of Redacted logo was added to make it look more authentic.

Statistical Analysis

Our statistical analysis involves a set of chi-squared tests, to evaluate how each demographic feature individually impacts on the phishing susceptibility, and a set of binary logistic regression models, to identify how demographic features affect the phishing susceptibility when evaluated simultaneously.

Chi-Squared Test. One of the aims of our phishing assessment was to identify the independent variables that impact on the phishing susceptibility. Thus, we relied on the Chi-Squared Test for Independence [36] to verify the association between two categorical variables. The Chi-Squared Test has the following benefits:

- Compatible with categorical variables (*age* was considered as categorical).
- Supports variables with more than two possible values (*age, role, working/studying field, email template*).
- Returns whether there is an association between categorical variables.

The Chi-Squared Test for Independence is a statistical hypothesis test. In each test, the null hypothesis is formulated as follows: "Variable [Gender, Age, Role, Working/Studying Field, Email Template] is independent from the Reaction variable in the population related to the phishing assessment results". The input for the test is the contingency table with observed values from the whole

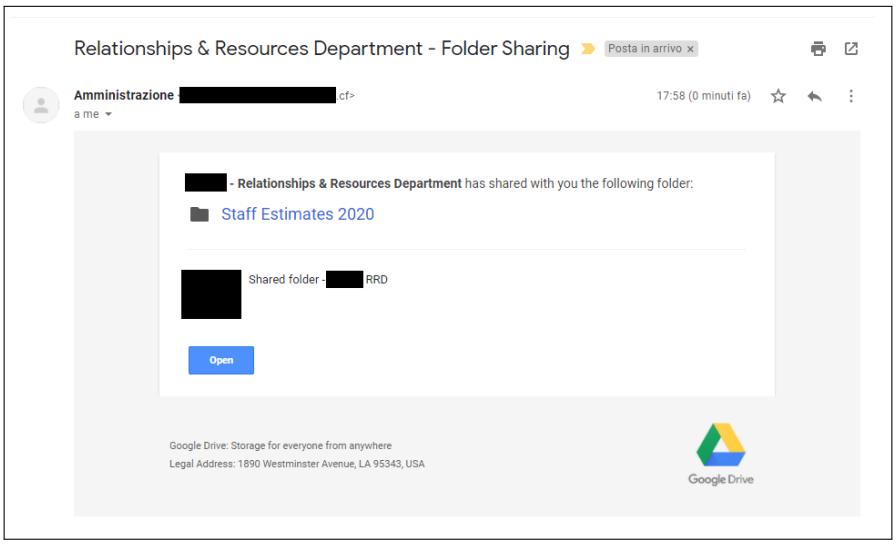


Figure 8: Batch 1 - Template 5 (staff).

experiment. The output of the test is the p-value: the probability to obtain results at least as extreme as the ones calculated in the test, assuming the validity of the null hypothesis. In other words, p-value is the chance that test results are arbitrary and not trustworthy. The standard significance level is represented by a p-value of "0.05". If we consider two variables as independent from each other, but we obtain a p-value lower than "0.05", it means that the results are so much unlikely that the two variables are associated and not independent. On the other side, having a p-value higher than "0.05" does not prove that the two variables are independent. Instead, it means that there is no statistically significant evidence to have a conclusion.

Binary Logistic Regression. Another important aim of our phishing assessment was to identify the multiplicity of factors that contribute to phishing susceptibility. The Chi-Squared Test analyzes whether it is present an association only between pairs composed by one of the predictor variables (i.e., *gender, age, role, email template*) and the dependent variable (i.e., *reaction*). Finally, another goal we had was to develop a predictive model trained to calculate the victim phishing susceptibility. The Binary Logistic Regression creates a predictive model, that takes into account all the variables at the same time [37]. The benefits of the Binary Logistic Regression are as follows:

- Multivariate and descriptive.
- Compatible with categorical variables (*age* was considered categorical).
- Compatible with a binary dependent variable (*reaction* had a binary outcome).

- Supports predictors with more than two possible values (*age, role, email template*).
- Returns whether there is an association between variables, fulfilling one of our goals.
- Returns a predictive model.

Binary Logistic Regression - Example					
Coefficients:					
	Estimate	Std. Error	z value	Pr(>z)	Signif.
(Intercept)	-8.9954	1.0090	-8.915	2e-16	***
PredictorA1	0.0374	0.0051	7.360	1.84e-13	***
PredictorA2	0.0730	0.0204	3.581	0.000342	***
PredictorA3	1.0559	0.4304	2.453	0.014154	*
PredictorB1	0.8185	0.3344	4.088	0.014367	*
PredictorB2	1.6339	0.3997	4.088	4.34e-05	***
PredictorC1	1.3530	0.5284	2.561	0.0105	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Deviance Residuals:					
	Min	1Q	Median	3Q	Max
	-2.9485	-0.6356	-0.3625	0.6233	2.6759

Null deviance:	498.10 on 391 degrees of freedom
Residual deviance:	341.42 on 385 degrees of freedom
AIC	355.42
Fisher scoring iterations:	5

Table 4: Binary logistic regression model example.

The Binary Logistic Regression is a classification algorithm that returns a table where rows are predictors and columns are statistical data [38]. Table 4 shows an example of a Binary Logistic Regression output, as a reference for the following explanation.

Each row corresponds to a possible value of a predictor (Intercept is a special case). For example, the predictor *email template* has five different possible values: T1, T2, T3, T4, T5. Each row of the table corresponds to a trait. The first trait of each predictor is absent from the table because it acts as a reference point. Categorical predictor coefficients are formulated with respect to other traits of the same predictor. The *estimate coefficient* describes the influence of a predictor's value with respect to the baseline trait (the missing one). When a row has a positive *estimate coefficient*, it means that an individual with that trait is more susceptible to phishing. Conversely, a negative *estimate coefficient*

indicates a low susceptibility. A predictor's baseline value is simply considered as 0. Comparing two *estimate coefficients* may give additional insights. If a trait has a value which is twice the value of another trait, it means its impact to phishing susceptibility is double. However, the *estimate coefficient* is not a direct measure of the magnitude of a predictor's effect. The *intercept*, also called *constant*, is a complex element that we did not consider in our analysis. The *standard error* is used to determine how well a model fits the data, and to assess the precision of predictions. The *z-value* is obtained by dividing the *estimate coefficient* with the *standard error*. Significant traits are identified by a z-value of high magnitude.

The *Pr(>z)* value has a similar role to the p-value in the Chi-Squared Test. *Pr(>z)* is compared to the five significance levels in the legend and discussed afterwards. A low *Pr(>z)* corresponds to a more significant trait. The *significance codes* are intuitive visual cues about the significance of a trait. A legend explains which symbols correspond to which range of values. The *significance codes* are applied to the value of the *Pr(>z)* column for each row.

The *Aikake Information Criterion* (AIC) is a common estimator for the quality of a predictive model. It mediates between the goodness of fit and the simplicity of the model. Models with low AIC are preferable. The *number of Fisher scoring iterations* is related to how the model fit was estimated. The *logistic regression* requires an iterative approach to evaluate model, and the number of iterations performed is exactly the *number of Fisher scoring iterations*.

Software and Tools

Gophish [19], as an open-source phishing framework, was at the core of our experiment. With Gophish, we were able to manage most aspects of email delivery and click tracking. We ran it on a virtual machine in the University internal network, and installed SSL certificates for the landing pages. We registered our counterfeit domains on Freenom [39], a provider of free Internet country code top-level domains (".tk", ".ml", ".ga", ".cf", ".gq"). Table 5 showcases some of our counterfeit domains, and their legitimate counterpart.

Counterfeit Domain	**Legitimate Domain**
@Redacted.cf	@Redacted.it
@Redacted.ml	@Redacted.it
@Redacted.ga	@Redacted.it
@goqgle.gq	@google.com
@researchgatemail.ml	@researchgatemail.net
@stanford-edu.tk	@stanford.edu

Table 5: Domains used in the phishing assessment at the University of Redacted.

We, then, required an SMTP provider to deliver emails from Gophish. The service we chose was SendGrid [40], a cloud-based SMTP provider that acts as an email delivery engine.

RESULTS AND DISCUSSION

In this section, we highlight the results of the phishing assessment and the statistical analysis. In particular, we present an overview of the raw data collected during the phishing assessment (section Overview), the outcome of the Chi-Squared tests for each independent variable (section Chi-Squared Test for Independence), and the outcome of the Binary Logistic Regression models (section Binary Logistic Regression).

Overview

Table 6 shows the percentages of participants that clicked on the link embedded in our phishing emails, over the three batches. Out of the 4.524 delivered emails, 819 were clicked (i.e., the 18.1%). The first batch achieved the highest clickrate with respect to the second and the third ones.

Considering the five email templates used in our phishing assessment, the ones classified as spearphishing (i.e., Template 3, Template 4 and Template 5) had a greater influence on the participants, who were more tempted to click on the link provided by the email. Moreover, the second batch of emails of Template 5 were incredibly effective because of an overlap between our phishing email advertising a fake Microsoft Office service migration and a real Microsoft Office service migration ongoing at the University. Thus, technical/administrative staff were easily lured by our email and clicked on the embedded link.

Among students, professors, and technical/administrative staff, students were found more susceptible to phishing. This might be a reasonable outcome, since students could be less trained and aware about phishing attacks and their consequences.

Chi-Squared Test for Independence

In order to identify which variables affect the susceptibility to phishing, we performed several Chi-Squared tests and we now discuss their outcome. Out of the five independent variables (i.e., *age*, *gender*, *role*, *working/studying field* and *email template*), we found that *role* and *email template* are the ones that can individually affect phishing susceptibility.

Age. The Chi-Squared p-value is lower than the significance level in Batch 1 (p-value=*1.292157e-10*) and Batch 3 (p-value=*7.446679e-06*), while it is higher in Batch 2 (p-value=*6.140571e-02*). Although the p-value in the second batch is close to the significance level, we can not argue an association relationship between *age* and *phishing susceptibility*. However, in our dataset there might

	Batch 1	**Batch 2**	**Batch 3**
Percentage of Clicked Emails			
	30% (456/1508)	11% (163/1508)	13% (200/1508)
Percentage of Clicked Emails in Different Email Templates			
Template 1	22% (107/492)	7% (36/528)	10% (48/488)
Template 2	26% (127/488)	4% (21/492)	18% (93/528)
Template 3	43% (85/200)	8% (13/160)	21% (35/164)
Template 4	32% (53/164)	16% (26/164)	1% (2/164)
Template 5	33% (54/164)	41% (67/164)	13% (22/164)

Table 6: Raw data of the phishing assessment at the University of Redacted.

be an association relationship between *age* and *role*, since students involved in our assessment are mostly younger then professors and staff.

Gender. The Chi-Squared p-value is higher than the significance level (p-value=*0.05*) in Batch 1 (p-value=*2.266868e-01*), Batch 2 (p-value=*7.795849e-02*) and Batch 3 (p-value=*7.849291e-02*). Thus, there is no statistically significant evidence to prove an association relationship between *gender* and *phishing susceptibility*.

Role. The Chi-Squared p-value is much lower than the significance level in Batch 1 (p-value=*2.030326e-10*), Batch 2 (p-value=*2.030326e-10*) and Batch 3 (p-value=*8.951827e-10*). Thus, we can confidently state that *role* is in an association relationship with *phishing susceptibility*.

Email Template. The Chi-Squared p-value is lower than the significance level in Batch 1 (p-value=*3.946323e-04*), Batch 2 (p-value=*6.341024e-40*) and Batch 3 (p-value=*7.566356e-09*). Thus, we can confidently state that there is an association relationship between the *email template* and the *phishing susceptibility*. As expected, the Chi-Squared test confirms that generic phishing (i.e., Template 1 and Template 2) is naturally far less effective than spearphishing (i.e., Template 3, Template 4 and Template 5).

Working/Studying Field. The *working/studying field* variable refers to different values according to the target *role*. In particular, the *working/studying field* is the Department which the degree of a student belongs to, it is the Macro-area which research area of a professor relates to and it is the specific Job of technical/administrative staff. The Chi-Squared tests for professors fail to show any statistically relevant association between *working/studying field* and *phishing susceptibility* in Batch 1 (p-value=*9.233677e-01*), Batch 2 (p-value=*9.439090e-01*) and Batch 3 (p-value=*5.014101e-01*). Similarly, the Chi-Squared tests for the technical/administrative staff fail to show any statistically relevant association between *working/studying field* and *phishing susceptibility* in Batch 1 (p-value=*2.656101e-01*), Batch 2 (p-value=*8.887189e-01*) and Batch 3 (p-value=*9.700039e-01*). Concerning the students, we did not perform a Chi-

Squared test between *working/studying field* and *phishing susceptibility* because the high number of departments dilutes the results too much. We expected the *working/studying field* to impact the phishing susceptibility, since it reflects the participant educational background, but we discovered the opposite. As a possible explanation, we think that phishing is affected more by the environmental factors when opening the email, than by the victim's technological background.

Binary Logistic Regression

To measure how the independent variables affect the phishing susceptibility, when considered simultaneously, we produced three binary logistic regression models, one for each batch. Table 7 (Batch 1), Table 8 (Batch 2) and Table 9 (Batch 3) show our models.

Binary Logistic Regression					
Coefficients	**Estimate**	**Std. Error**	**z value**	**Pr(>z)**	**Signif.**
(Intercept)	*-0.2052*	*0.2656*	*-0.773*	*0.439696*	
Age21-25	*-0.2130*	*0.2636*	*-0.808*	*0.419099*	
Age26-30	*-0.3715*	*0.3153*	*-1.178*	*0.238650*	
Age31-35	*-1.1971*	*0.4181*	*-2.863*	*0.004199*	**
Age36-40	*-1.6673*	*0.4362*	*-3.822*	*0.000132*	***
Age41-45	*-1.5141*	*0.4289*	*-3.530*	*0.000415*	***
Age46-50	*-1.0659*	*0.4178*	*-2.551*	*0.010735*	*
Age51-55	*-1.5388*	*0.4215*	*-3.651*	*0.000261*	***
Age56-60	*-1.4585*	*0.4427*	*-3.294*	*0.000986*	***
Age61-99	*-2.0477*	*0.4801*	*-4.266*	*1.99e-05*	***
GenderM	*-0.1361*	*0.1232*	*-1.105*	*0.269270*	
Roledoc	*-0.5869*	*0.3597*	*-1.632*	*0.102704*	
Rolepta	*0.1862*	*0.3210*	*0.580*	*0.561806*	
TemplateT2	*0.2941*	*0.1592*	*1.847*	*0.064732*	.
TemplateT3	*0.2705*	*0.2020*	*1.339*	*0.180557*	
TemplateT4	*1.6728*	*0.2667*	*6.271*	*3.58e-10*	***
TemplateT5	*0.7648*	*0.2336*	*3.275*	*0.001058*	**
Signif. codes		0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			
AIC		1681.20			

Table 7: Binary logistic regression model - Batch 1.

The results of our binary logistic regression models show that, in a multivariate analysis, *age* and *email template* (in Batch 1) and *email template* only

Binary Logistic Regression					
Coefficients	**Estimate**	**Std. Error**	**z value**	**Pr(>z)**	**Signif.**
(Intercept)	*-16.38912*	*891.28429*	*-0.018*	*0.98533*	
Age21-25	*-0.61561*	*0.43875*	*-1.403*	*0.16058*	
Age26-30	*-0.63724*	*0.54930*	*-1.160*	*0.24601*	
Age31-35	*-0.72371*	*0.65537*	*-1.104*	*0.26947*	
Age36-40	*-0.62719*	*0.65522*	*-0.957*	*0.33845*	
Age41-45	*-1.29717*	*0.68188*	*-1.902*	*0.05713*	*.*
Age46-50	*-0.64962*	*0.64605*	*-1.006*	*0.31464*	
Age51-55	*-0.84601*	*0.64950*	*-1.303*	*0.19273*	
Age56-60	*-0.71426*	*0.66847*	*-1.068*	*0.28530*	
Age61-99	*-0.84776*	*0.68820*	*-1.232*	*0.21800*	
GenderM	*-0.25225*	*0.18573*	*-1.358*	*0.17441*	
Roledoc	*-0.04303*	*0.56308*	*-0.076*	*0.93908*	
Rolepta	*0.10290*	*0.53683*	*0.192*	*0.84799*	
TemplateS2	*-0.49824*	*0.28338*	*-1.758*	*0.07871*	*.*
TemplateS3	*0.05261*	*0.37440*	*0.141*	*0.88825*	
TemplateS4	*1.17040*	*0.35954*	*3.255*	*0.00113*	****
TemplateS5	*2.24070*	*0.31209*	*7.180*	*6.99e-13*	*****
Signif. codes		0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			
AIC		918.57			

Table 8: Binary logistic regression model - Batch 2.

(in Batch 2 and in Batch 3) have a statistically significant relationship with phishing susceptibility. Below, we provide details about each independent variable considered in our study and found to impact the phishing susceptibility in our binary regression models (i.e., *age*, *role* and *email template*).

Age. The Binary Logistic Regression model shows that *age* has a relationship with *phishing susceptibility* only in Batch 1, while the Chi-Squared test shows an association relationship both in Batch 1 and Batch 3. Our possible explanation for this result is that the *age* could be related to the emotional response in front of an unexpected email. Thus, in Batch 1 younger subjects were more prone to click on the malicious link, while older ones reacted in the opposite way. After acknowledging the risk of receiving phishing emails, the *age* then became an obsolete variable, as subjects were more skeptical and the emotional response was mitigated. In [41], the authors state that phishing susceptibility is influenced by *age* and *gender*. They rank the effectiveness of those

Binary Logistic Regression					
Coefficients	**Estimate**	**Std. Error**	**z value**	**Pr(>z)**	**Signif.**
(Intercept)	*-1.69215*	*0.33869*	*-4.996*	*5.85e-07*	***
Age21-25	*0.02611*	*0.32284*	*0.081*	*0.935537*	
Age26-30	*-0.19495*	*0.39465*	*-0.494*	*0.621325*	
Age31-35	*-0.29291*	*0.49887*	*-0.587*	*0.557103*	
Age36-40	*-0.56991*	*0.52999*	*-1.075*	*0.282232*	
Age41-45	*-0.81102*	*0.53885*	*-1.505*	*0.132301*	
Age46-50	*-0.62455*	*0.52189*	*-1.197*	*0.231419*	
Age51-55	*-0.45263*	*0.50811*	*-0.891*	*0.373034*	
Age56-60	*-0.99038*	*0.56828*	*-1.743*	*0.081375*	.
Age61-99	*-0.98868*	*0.60945*	*-1.622*	*0.104748*	
GenderM	*-0.14914*	*0.16000*	*-0.932*	*0.351247*	
Roledoc	*-0.24016*	*0.41402*	*-0.580*	*0.561868*	
Rolepta	*0.08247*	*0.38496*	*0.214*	*0.830366*	
TemplateS2	*0.65299*	*0.19254*	*3.392*	*0.000695*	***
TemplateS3	*0.49223*	*0.26532*	*1.855*	*0.063566*	.
TemplateS4	*1.70308*	*0.74828*	*-2.276*	*0.022846*	*
TemplateS5	*0.41324*	*0.30896*	*1.338*	*0.181059*	
Signif. codes		0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			
AIC		1137.00			

Table 9: Binary logistic regression model - Batch 3.

principles for the various groups of victims, and conclude that older women are the most susceptible group to spearphishing.

Role. The *role* is not significant in our Binary Logistic Regression model, even though it is was related to *phishing susceptibility* for the Chi-Squared tests. From the final results, we can clearly see that students are the most susceptible to phishing, at least in Batch 1 and Batch 3. The vast majority of students are 0-30 years old, while professors and staff members predominantly belong to the 30-99 years old range. In fact, in our dataset, the youngest individuals are extremely likely to be students, thus utilizing both *role* and *age* may be redundant. This interpretation partially contrasts with [42]. In that study, the researchers exhibit a breakdown of common characteristics in phishing and spearphishing under three main aspects: characteristics of deployment, cognitive features, victim demographics. The demographics data was gathered through LinkedIn [43], so they focused on the professional, national, and seniority profile

of victims. They concluded that demographics data was irrelevant in regards to generic phishing. Our model confirms the same statement. However, they also concluded that demographics was relevant in spearphishing attempts, while our model does not find any relationship.

Email Template. The *email template* has strongest relationship with *phishing susceptibility*. This could be a direct consequence of the widely different content inside the emails, even though we tried to make them as equally difficult as possible. Still, it is reasonable to assume that spearphishing is far more effective than generic phishing, and this fact would explain the results found by our model.

Tests without Email Template variable. Since we speculate that *age* and *email template* might not be independent variables, we created more Binary Logistic Regression models without *email template*, to see any change in the output. We obtained five models (one for each separate *email template*) for each batch. Under such settings, the *age* is the most significant predictor for the generic phishing email templates (S1-S2), reinforcing our belief that emotional response (tied to the *age*) is still an important influence factor. As for the spearphishing *email templates* (S3-S4-S5), no significant predictors were found.

CONCLUSION

In our work, we designed and deployed a phishing assessment at the University of Redacted. We collected valuable data to better understand the risk of phishing and spearphishing on public institutions such as universities and schools. We sent three batches of emails to three different target categories (i.e., professors, technical / administrative staff members, students). The first batch of emails had a click rate of 30%, while the second and third batches ended with more optimistic outcomes: a click rate of respectively 11% and 13%. We merged the data from the assessment with anonymized demographics from the participants, and processed it through univariate (Chi-Squared Test for Independence) and multivariate (Binary Logistic Regression) analysis. We identified that *age* and *email template* (i.e., phishing or spearphishing template) have a strong relationship with phishing susceptibility. The Chi-Squared p-values showed an association relationship for *age* in Batch 1 (p-value=1.292157e-10) and Batch 3 (p-value=7.446679e-06), and for *email template* in Batch 1 (p-value=2.030326e-10), Batch 2 (p-value=2.030326e-10) and Batch 3 (p-value=8.951827e-10). Similarly, the Binary Logistic Regression analysis showed a strong relationship with *age* in Batch 1, and with *email template* in Batch 1, Batch 2 and Batch 3. By looking at the statistical analysis results, we argue that the age was most impactful in Batch 1, when the phishing emails were most unexpected. We also argue that employees (i.e., professors and staff members) at the University of Redacted are more resilient to generic phishing, but still very susceptible to spearphishing. Comparatively, students are almost equally susceptible to generic phishing

and spearphishing. As a future work, we aim to extend our phishing assessment to test more scenarios under different conditions. We plan to formally evaluate the difficulty of email templates, so that the statistical analysis are less susceptible to randomness, as it happened with the spearphishing email sent to staff members during Batch 2. We also plan to integrate educational content within the assessment, and evaluate if phishing aware participants would react better to our phishing assessment, and for how long their retain their knowledge. This way, we could add education as an additional predictor for phishing susceptibility.

References

[1] A.-P. W. G. (APWG), Phishing activity trends report, 3rd quarter 2019, https://apwg.org/trendsreports/ (2019).

[2] Verizon, Data breach investigation report, https://www22.verizon.com/wholesale/contenthub/data_breach_investigation_report.html (2018).

[3] Symantec, Internet security threat report, https://www.phishingbox.com/news/phishing-news/symantec-internet-security-threat-report-2018 (2018).

[4] Proofpoint, State of the phish report, https://www.proofpoint.com/us/security-awareness/post/2020-state-phish-security-awareness-training-email-reporting-more-critical (2020).

[5] a. b. R. L. The Hacker News, Hackers targeting companies involved in covid-19 vaccine distribution, https://thehackernews.com/2020/12/hackers-targeting-companies-involved-in.html (2020).

[6] I. S. . T. University of Montevallo, Phishing emails, https://www.montevallo.edu/about-um/administration/ist/phishing/ (2020).

[7] B. I. S. Office, Phishing example: Paypal - we need your help, https://security.berkeley.edu/news/phishing-example-paypal-we-need-your-help (2016).

[8] B. I. S. Office, Phishing example: Fedex shipment update, https://security.berkeley.edu/news/phishing-example-fedex-shipment-update (2017).

[9] B. I. S. Office, Phishing example: Google doc phishing message, https://security.berkeley.edu/news/phishing-example-google-doc-phishing-message (2017).

[10] B. I. S. Office, Phishing example: Itunes access disabled, https://security.berkeley.edu/news/phishing-example-last-reminder-you-must-update-your-apple-account-information (2016).

[11] B. I. S. Office, Phishing example: Your dropbox file, https://security.berkeley.edu/news/phishing-example-your-dropbox-file (2017).

[12] B. I. S. Office, Phishing example: Urgent request (email impersonation), https://security.berkeley.edu/news/phishing-example-urgent-request-email-impersonation (2020).

[13] B. I. S. Office, Phishing example: Important announcement from chancellor dirks, https://security.berkeley.edu/news/phishing-example-important-announcement-chancellor-dirks (2016).

[14] Proofpoint, Threat actor profile: Ta407, the silent librarian, https://www.proofpoint.com/us/threat-insight/post/threat-actor-profile-ta407-silent-librarian (2019).

[15] C. News, Macewan university defrauded of $11.8m in online phishing scam, https://www.cbc.ca/news/canada/edmonton/macewan-university-phishing-scam-edmonton-1.4270689 (2017).

[16] KnowBe4, Fooled into losing their paychecks, https://blog.knowbe4.com/it-only-takes-1-phish-wichita-state-university-employees-get-fooled-into-losing-their-paychecks (2015).

[17] ZDnet, Phishing attack: Students' personal information stolen in university data breach, https://www.zdnet.com/article/phishing-attack-students-personal-information-stolen-in-university-data-breach/ (2019).

[18] L. University, Lancaster phishing cyber incident, https://www.lancaster.ac.uk/news/phishing-attack/ (2019).

[19] J. Wright, Gophish, https://getgophish.com (2020).

[20] N. C. for the Protection of Human Subjects of Biomedical, B. Research, The belmont report, ethical principles and guidelines for the protection of human subjects of research, https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html (1979).

[21] P. Finn, M. Jakobsson, Designing ethical phishing experiments, IEEE Technology and Society Magazine 26 (1) (2007) 46–58. doi:10.1109/MTAS.2007.335565.

[22] PhishLabs, Phishing trends and intelligence report, https://info.phishlabs.com/blog/2018-phishing-trends-intelligence-report-released (2018).

[23] J. P. Magalhães, A. Pinto, A methodology for assessing the resilience against email phishing, in: 2018 International Conference on Intelligent Systems, IEEE, Funchal, Portugal, 2018, pp. 515–520.
URL https://ieeexplore.ieee.org/document/7522363

[24] W. D. Kearney, H. A. Kruger, Phishing and organisational learning, in: IFIP Advances in Information and Communication Technology, 2013, pp. 379–390.

[25] R. Wash, M. M. Cooper, Who provides phishing training? facts, stories, and people like me, ACM CHI Conference on Human Factors in Computing Systems (2018) 1–12.

[26] W. Gordon, A. Wright, R. Glynn, J. Kadakia, C. Mazzone, E. Leinbach, A. Landman, Evaluation of a mandatory phishing training program for high-risk employees at a us healthcare system, J. Am. Med. Inform. Assoc. 26 (2019) 547–552. doi:10.1093/jamia/ocz005.

[27] T. Bakhshi, Social engineering: Revisiting end-user awareness and susceptibility to classic attack vectors, in: 2017 13th International Conference on Emerging Technologies (ICET), IEEE, Islamabad, Pakistan, 2017, pp. 1–6. URL https://ieeexplore.ieee.org/document/8281653

[28] W. Flores, H. Holm, G. Svensson, G. N. Ericsson, Using phishing experiments and scenario-based surveys to understand security behaviours in practice, Information Management & Computer Security 22 (2014) 393–406. doi:10.1108/IMCS-11-2013-0083.

[29] R. Dodge, A. Ferguson, Using phishing for user email security awareness, in: Security and Privacy in Dynamic Environments, Springer, Boston, USA, 2006, pp. 454–459.
URL https://link.springer.com/chapter/10.1007/0-387-33406-8_41

[30] A. Mihelic, M. JevSCek, S. Vrhovec, I. Bernik, Testing the human backdoor: Organizational response to a phishing campaign, Journal of Universal Computer Science 25 (2019) 1458–1477. doi:10.3217/jucs-025-11-1458.

[31] A. Burns, M. Johnson, D. Caputo, Spear phishing in a barrel: Insights from a targeted phishing campaign, Journal of Organizational Computing and Electronic Commerce 29 (2019) 24–39. doi:10.1080/10919392.2019.1552745.

[32] D. D. Caputo, S. L. Pfleeger, J. D. Freeman, M. E. Johnson, Going spear phishing: Exploring embedded training and awareness, IEEE Security & Privacy 12 (1) (2014) 28–38. doi:10.1109/MSP.2013.106.

[33] S. McElwee, G. Murphy, P. Shelton, Influencing outcomes and behaviors in simulated phishing exercises, in: SoutheastCon 2018, 2018, pp. 1–6. doi:10.1109/SECON.2018.8479109.

[34] W. Li, J. Lee, J. Purl, F. Greitzer, B. Yousefi, K. Laskey, Experimental investigation of demographic factors related to phishing susceptibility, in: Hawaii International Conference on System Sciences, 2020, pp. 1–10. doi:10.24251/HICSS.2020.274.

[35] J. Martin, C. Dubé, M. Coovert, Signal detection theory (sdt) is effective for modeling user behavior toward phishing and spear-phishing attacks, Human Factors 60 (2018) 1179–1191. doi:10.1177/0018720818789818.

[36] Wikipedia, Chi-squared test, https://en.wikipedia.org/wiki/Chi-squared_test (2020).

[37] A. Rawat, Binary logistic regression - an overview and implementation in r, https://towardsdatascience.com/implementing-binary-logistic-regression-in-r-7d802a9d98fe (2017).

[38] Displayr, How to interpret logistic regression coefficients, https://www.displayr.com/how-to-interpret-logistic-regression-coefficients (2019).

[39] Freenom, Internet domain provider, https://www.freenom.com/en/index.html (2020).

[40] T. SendGrid, Email delivery service, https://sendgrid.com (2020).

[41] D. Oliveira, H. Rocha, H. Yang, D. Ellis, S. Dommaraju, M. Muradoglu, D. Weir, A. Soliman, T. Lin, N. Ebner, Dissecting spear phishing emails for older vs young adults: On the interplay of weapons of influence and life domains in predicting susceptibility to phishing, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, 2017, p. 6412–6424. doi:10.1145/3025453.3025831.

[42] L. Allodi, T. Chotza, E. Panina, N. Zannone, The need for new antiphishing measures against spear-phishing attacks, IEEE Security & Privacy 18 (2) (2020) 23–34. doi:10.1109/MSEC.2019.2940952.

[43] LinkedIn, Social media website, https://www.linkedin.com/ (2020).