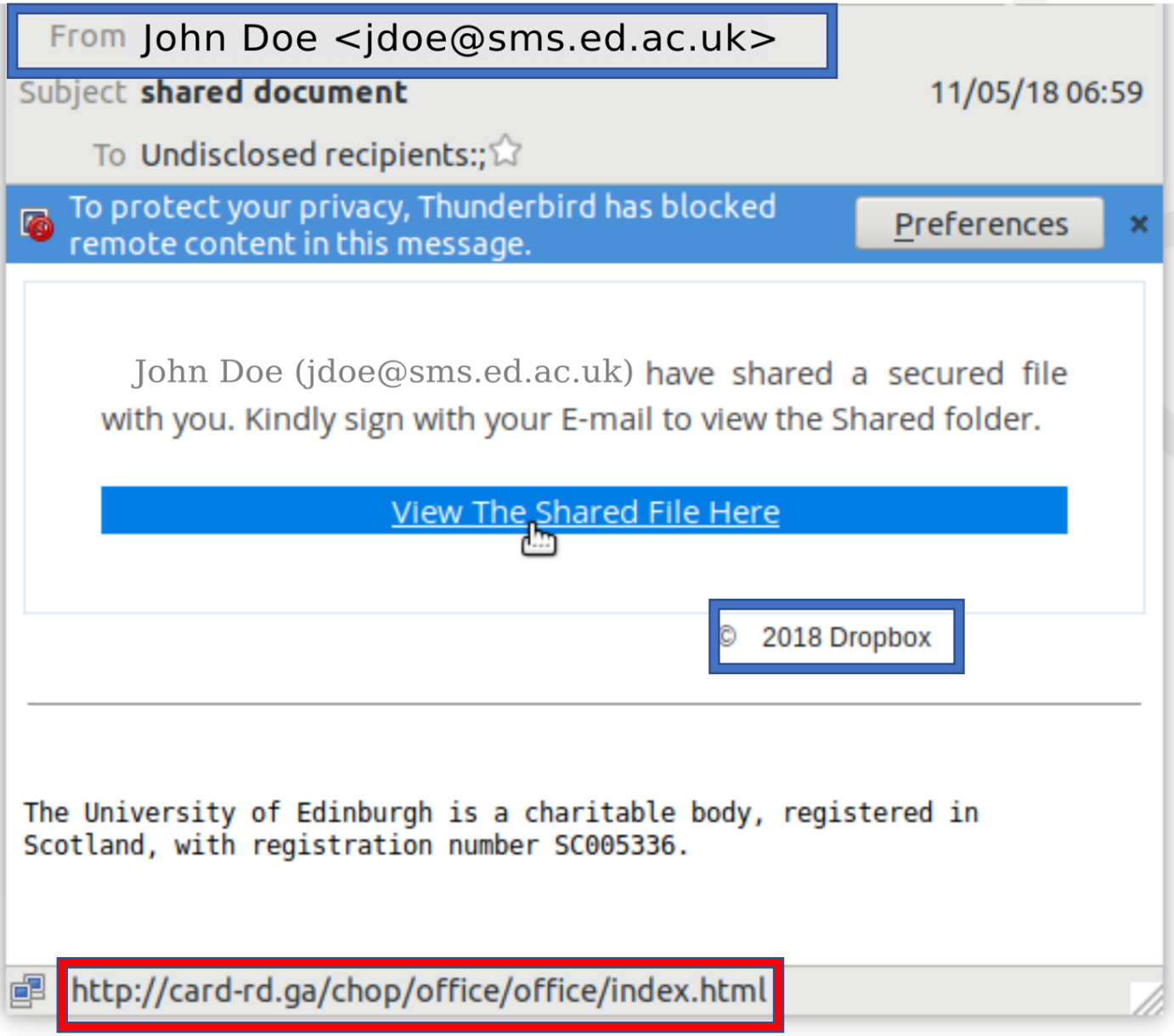
“I DIDN’T CLICK”: WHAT USERS SAY WHEN REPORTING PHISHING

N. PILAVAKIS, **A. JENKINS**, N. KOKCIYAN, K. VANIEA


From John Doe <jdoe@sms.ed.ac.uk>
Subject shared document
11/05/18 06:59
To Undisclosed recipients:;
To protect your privacy, Thunderbird has blocked remote content in this message.
Preferences
John Doe (jdoe@sms.ed.ac.uk) have shared a secured file with you. Kindly sign with your E-mail to view the Shared folder.
View The Shared File Here
© 2018 Dropbox
The University of Edinburgh is a charitable body, registered in Scotland, with registration number SC005336.
http://card-rd.ga/chop/office/office/index.html

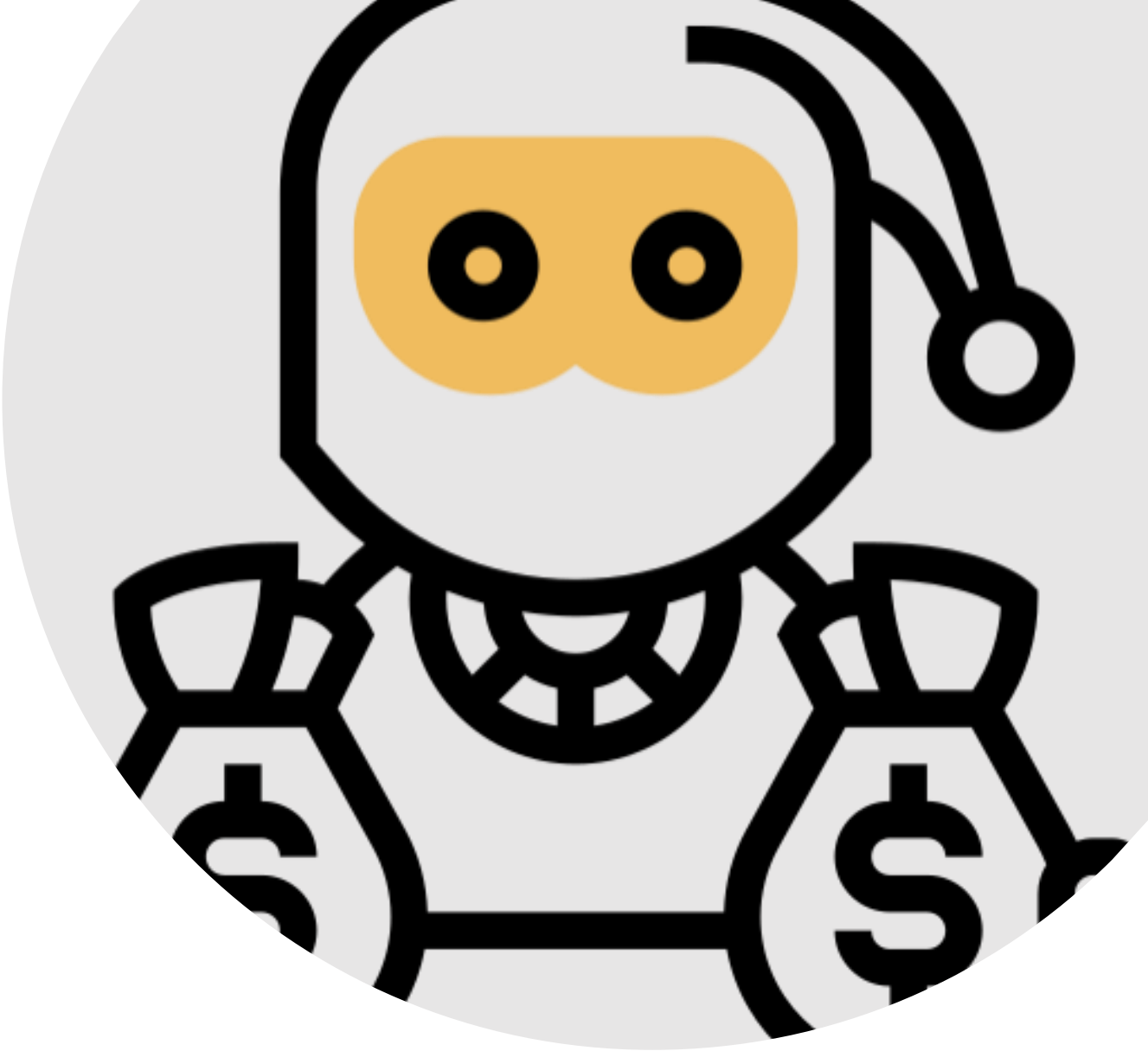
Phishing Impact

Most common threat vector, in the UK alone accounts for **83%** of attacks in 2022[1].

IBM found that phishing is the most expensive vector, at an avg. **of $4.65M**[2].

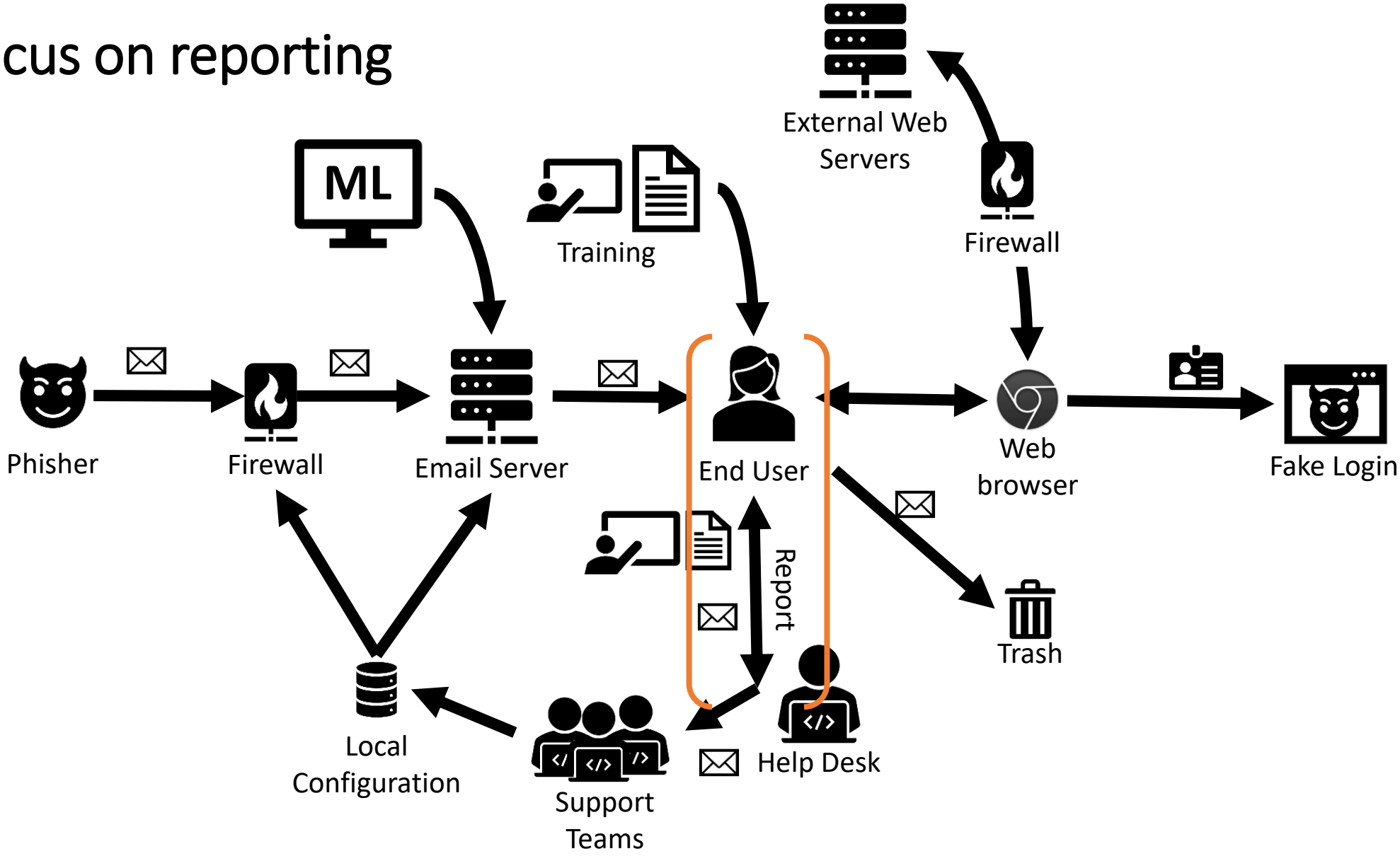
Initial threat causing a cascade of issues such as:

- Malware,
- Data Loss,
- Ransomware, etc.



1. https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2022/cyber-security-breaches-survey-2022
2. https://www.ibm.com/reports/data-breach

We focus on reporting



RESEARCH QUESTIONS

RQ1 - What statements and information do people provide when submitting phishing reports?

RQ2 - What questions do people ask and what kinds of support requests do they make when submitting phishing reports?

TULiPS
Technology Usability Lab in Privacy and Security

University Processes

This work is part of a larger project regarding how organizations handle phishing reports.

Universities have unique features:

- High turnover of staff and students
- Communicate regularly with external orgs.

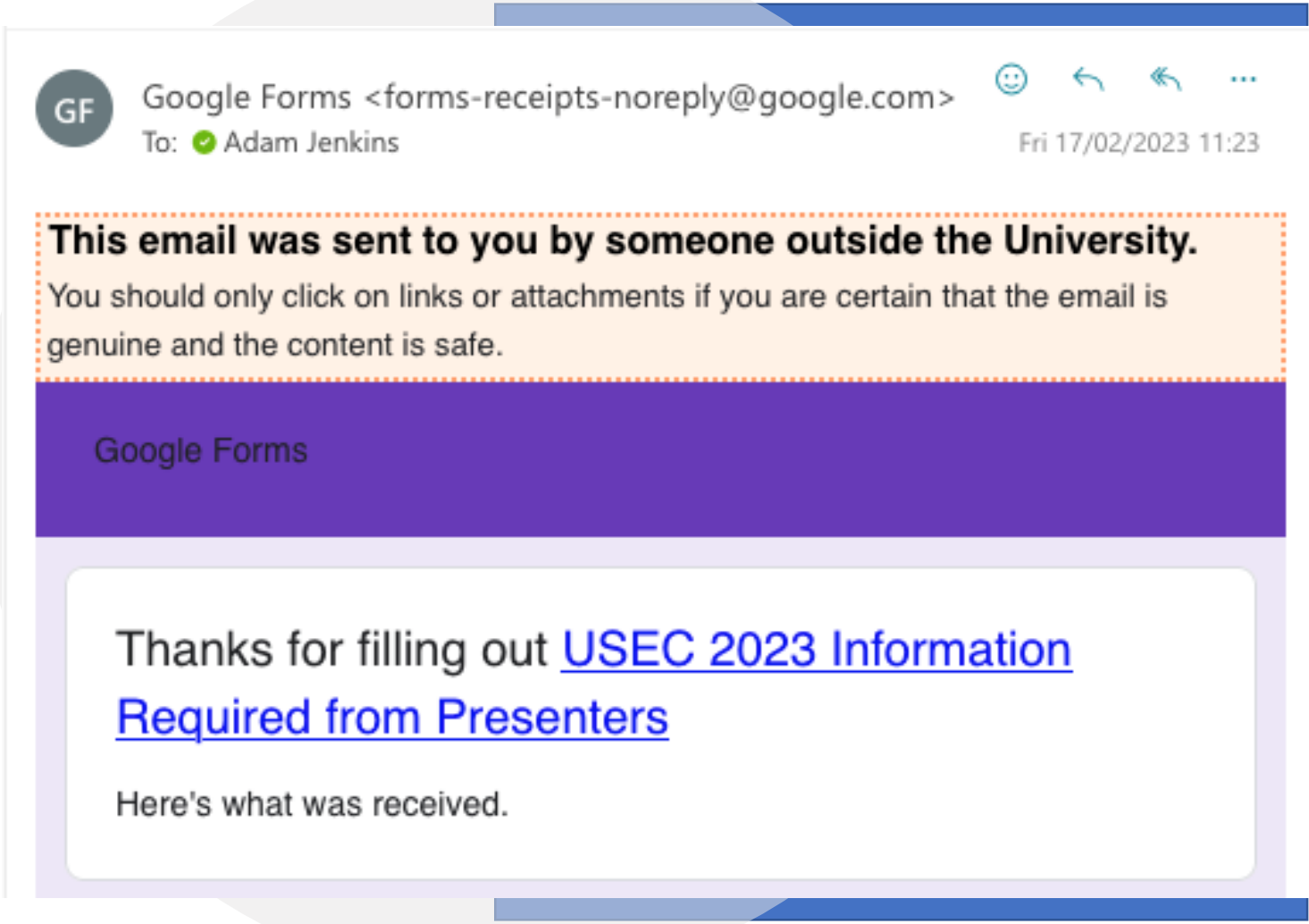
Help Desk system where staff and students may report any IT related issues, using a **'Ticketing'** approach.

METHODOLOGY

TULiPS
Technology Usability Lab in Privacy and Security

Dataset – Help Desk Tickets

- Initial search using keywords, i.e. “phishing”
- Limited search from 27th Oct. 2020, to 2nd Aug. 2021 (9 months)
- Limited due to roll-out of Automatic Banner on external emails
- Final Set was **984** Tickets



Non-Phish & Reporters

NON-PHISH

- Examples where they were confirmed by help desk to be False Positives
- Keyword search, i.e. “legitimate”, resulted in 94
- Total of 22 from manual review

WHO REPORTS

633 unique reporters

- 497 single reports
- 86 reported twice

- Highest reporter provided 71 reports, next being 14
- 82 reports from IT services group

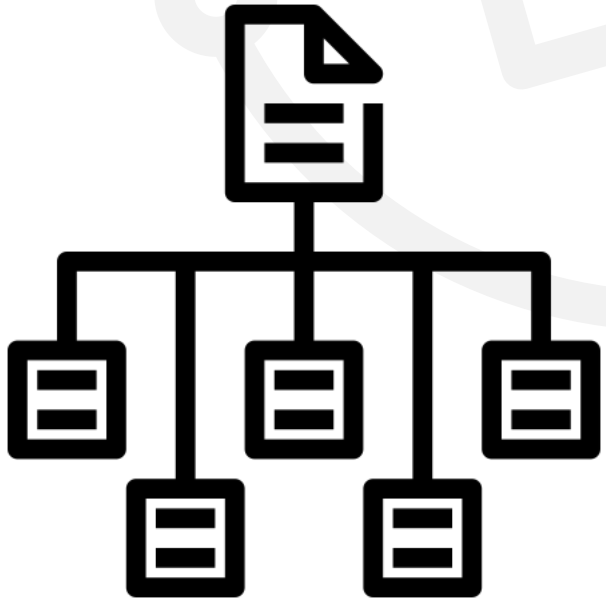
Qualitative Codebook

Initial 100 random tickets were read by lead researcher.

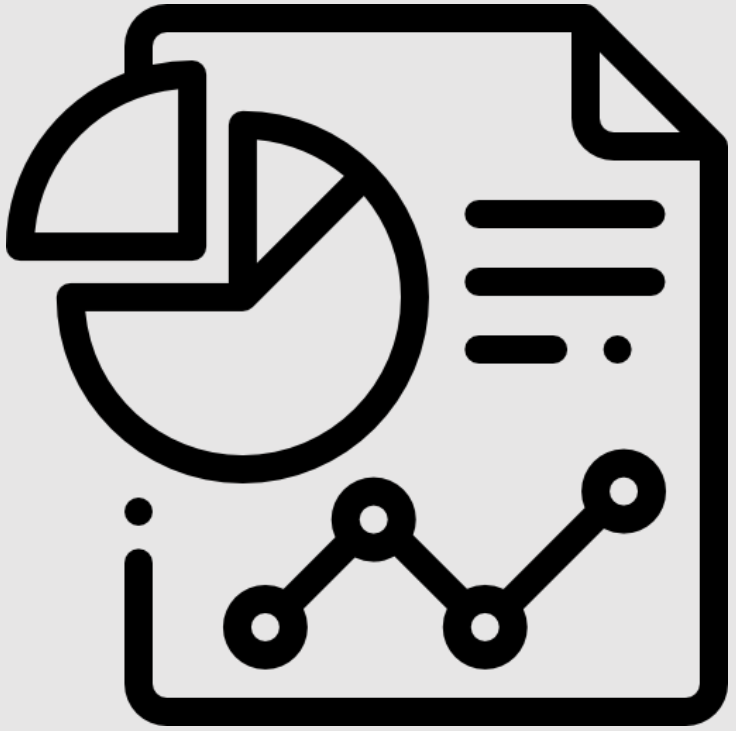
A subset of 300 random tickets for collaborative refinement

Applied to full 300, allowing multiple codes per ticket.

- Krippendorf's Alpha with Jaccard's distance.
- Initial 27 tickets had a score of **0.69**
- Final 31 tickets had a score of **0.77**
- 30 tickets removed as non-phishing related.



RESULTS



TULiPS
Technology Usability Lab in Privacy and Security

Codes, Subcodes & Counts

Just reporting		116
Evidence		82
	From address	53
	Cues	15
	Technical	12
	Unsolicited	11
	Other people	9
	Banner	6
	Tools	5
	Other evidence	2
Impact potential		62
	Repeated emails	21
	Compromised	18
	Convincing	15
	IT systems	8
	Number targeted	7

Actions taken		85
	Clicked	28
	Not clicked	35
	Delete	19
	Change login	8
	Gave data	7
	Not give data	7
	Opened	0
	Not opened	7
	Responded	1
	Not responded	2
	Other action	2
Questions		33
	Next Steps	11
	Other Questions	22

RQ1 - What statements and information do people provide when submitting phishing reports?

Evidence - Subcode	Counts
From address	53
Cues	15
Technical	12
Unsolicited	11
Other people	9
Banner	6
Tools	5
Other evidence	2

“sender looks like a Chemistry Student”

“the time the email was sent compared to the time the call was meant to be received do not align”

”not expecting an email from…”

“thought I would check first in case more users have received this?”

RQ1 - What statements and information do people provide when submitting phishing reports?

Impact Potential - Subcode	Counts
Repeated emails	21
Compromised	18
Convincing	15
IT systems	8
Number targeted	9

"Our [Head of School] has been impersonated again by this address - [attacker email]. In the past you have applied a rule to silently block this email on the mail server.

Please could you do that again?

Some people from our School have engaged and I'd like to cease comms from this address ASAP"

RQ2 - What questions do people ask and what kinds of support requests do they make when submitting phishing reports?

Questions- Subcode	Counts
Next Steps	11
Other Questions	22

“I have received an Outlook calendar invitation from an unknown source, and I think it might be a phishing attempt.

I want to remove it from my calendar. Please can you advise how I can do so safely. I deleted the invitation, but it is still showing as a recurring appointment in my calendar”

DISCUSSION

TULiPS
Technology Usability Lab in Privacy and Security

Discussion

- Similarities with spotting phish
  - Prior work highlight several features that indicate phishing to end-users.
  - *Evidence* and *Impact* potential map well to stages in phishing decisions.

- Self-Efficacy
  - Seeking confirmation regarding state of Phish.
  - Problematic ‘Paranoia’.

Future Work

Reporting Systems and Human AI-collaboration[1]

- Contextual features combined with technical information

Encouraging Phishing Reporting

- Development of systems that do not overwhelm IT staff[2]

Reassuring users

- Provide responses at scale given Help desk workload

1. Jenkins, A., Kokciyan, N., & Vaniea, K. (2022, June). PhishED: Automated contextual feedback for reported Phishing. In *18th Symposium on Usable Privacy and Security*. Usenix.
2. Saka, T., Vaniea, K., & Kökciyan, N. (2022, November). Context-Based Clustering to Mitigate Phishing Attacks. In *Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security* (pp. 115-126).

THE UNIVERSITY of EDINBURGH

REPHRAIN

THANK YOU & QUESTIONS?

ADAM.JENKINS@ED.AC.UK

HTTPS://GROUPS.INF.ED.AC.UK/TULIPS/

HTTPS://WWW.REPHRAIN.AC.UK/PHISHED/

@TULIPSLAB, @ADAMDGJENKINS18

@AJENK@HCI.SOCIAL