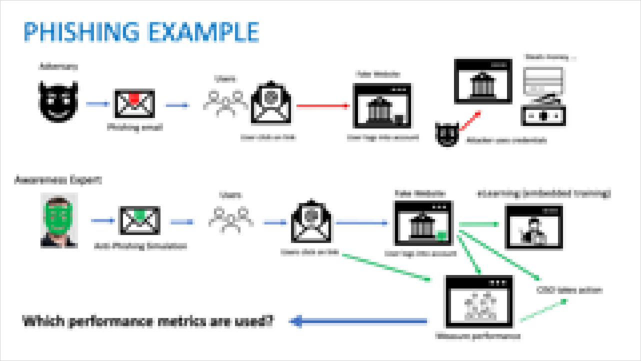
Avoiding The Hook
Phishing Awareness

Zürich, Global Cyber Conference 2023

Thomas Sutter, 15. September



PHISHING EXAMPLE

Adversary

Phishing email

Users

User click on link

Fake Website

User logs into account

Steals money ...

Attacker uses credentials

Awareness Expert

Anti-Phishing Simulation

Users

Users click on link

Fake Website

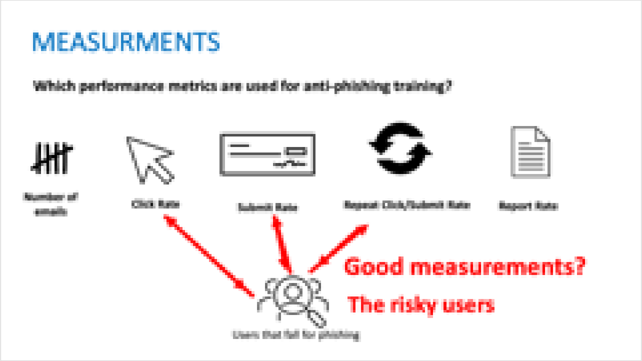
User logs into account

eLearning (embedded training)

CISO takes action

Measure performance

Which performance metrics are used?



MEASURMENTS

Which performance metrics are used for anti-phishing training?

Number of emails

Click Rate

Submit Rate

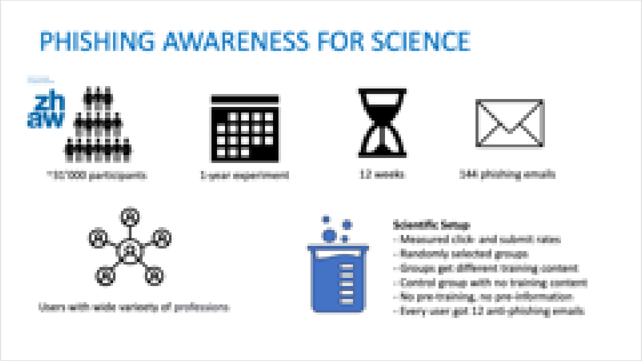
Repeat Click/Submit Rate

Report Rate

Good measurements?

The risky users

Users that fall for phishing



PHISHING AWARENESS FOR SCIENCE

~15'000 participants

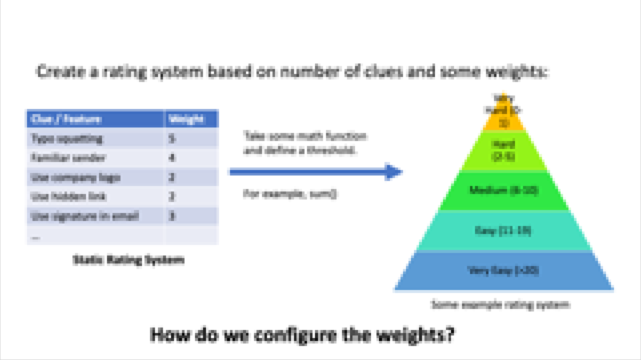
1-year experiment

12 weeks

144 phishing emails

Users with wide variety of professions

Scientific Setup
- Measured click- and submit rates
- Randomly selected groups
- Groups get different training content
- Control group with no training content
- No pre-training, no pre-information
- Every user got 12 anti-phishing emails



Create a rating system based on number of clues and some weights:

Clue / Feature	Weight
Typo squatting	5
Familiar sender	4
Use company logo	2
Use hidden link	2
Use signature in email	3
...	

Static Rating System

Take some math function and define a threshold.

For example, sum()

Very Hard (0-1)

Hard (2-5)

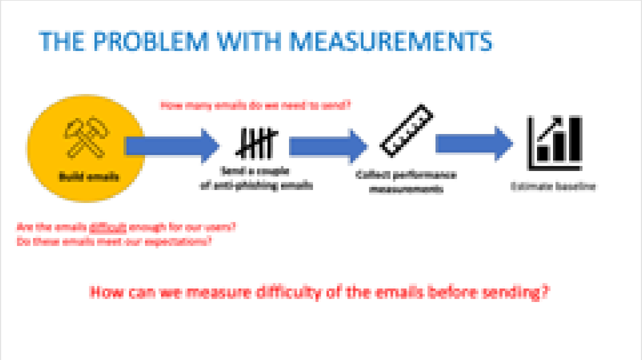
Medium (6-10)

Easy (11-19)

Very Easy (>20)

Some example rating system

How do we configure the weights?



THE PROBLEM WITH MEASUREMENTS

Build emails

How many emails do we need to send?

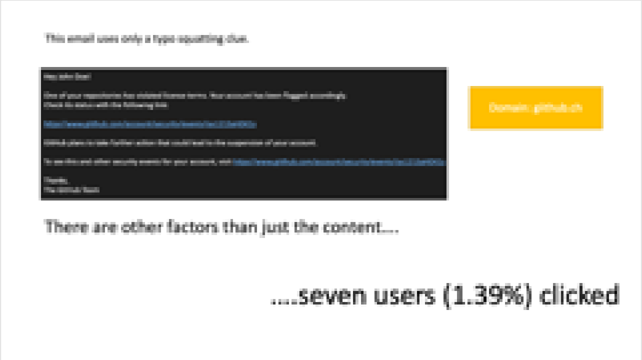
Send a couple of anti-phishing emails

Collect performance measurements

Estimate baseline

Are the emails difficult enough for our users?
Do these emails meet our expectations?

How can we measure difficulty of the emails before sending?



This email uses only a typo squatting clue.

Hey John Doe!

One of your repositories has violated license terms. Your account has been flagged accordingly.
Check its status with the following link:

https://www.github.com/account/security/events/[illegible]

GitHub plans to take further action that could lead to the suspension of your account.

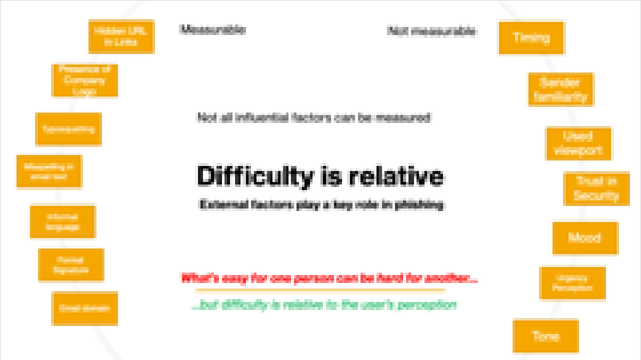
To see this and other security events for your account, visit https://www.github.com/account/security/events/[illegible]

Thanks,
The GitHub Team

Domain: github.ch

There are other factors than just the content....

....seven users (1.39%) clicked



Measurable

Not measurable

Hidden URL in Links

Presence of Company Logo

Typosquatting

Misspelling in email text

Informal language

Formal Signature

Email domain

Timing

Sender familiarity

Used viewport

Trust in Security

Mood

Urgency Perception

Tone

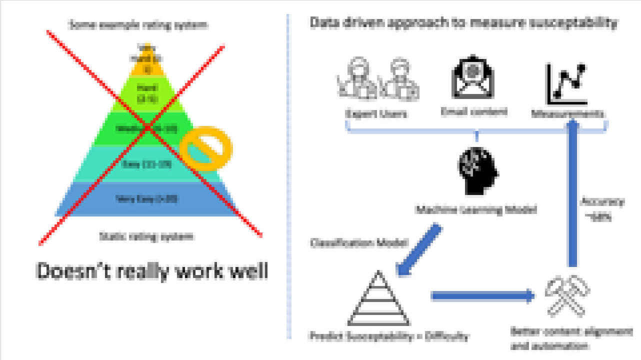
Not all influential factors can be measured

Difficulty is relative

External factors play a key role in phishing

What's easy for one person can be hard for another...

...but difficulty is relative to the user's perception



Some example rating system

Very Hard (0-1)

Hard (2-5)

Medium (6-10)

Easy (11-19)

Very Easy (>20)

Static rating system

Doesn't really work well

Data driven approach to measure susceptability

Expert Users

Email content

Measurements

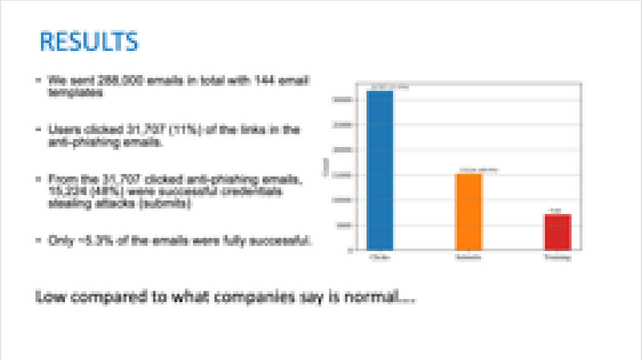
Machine Learning Model

Classification Model

Accuracy ~68%

Predict Susceptability = Difficulty

Better content alignment and automation

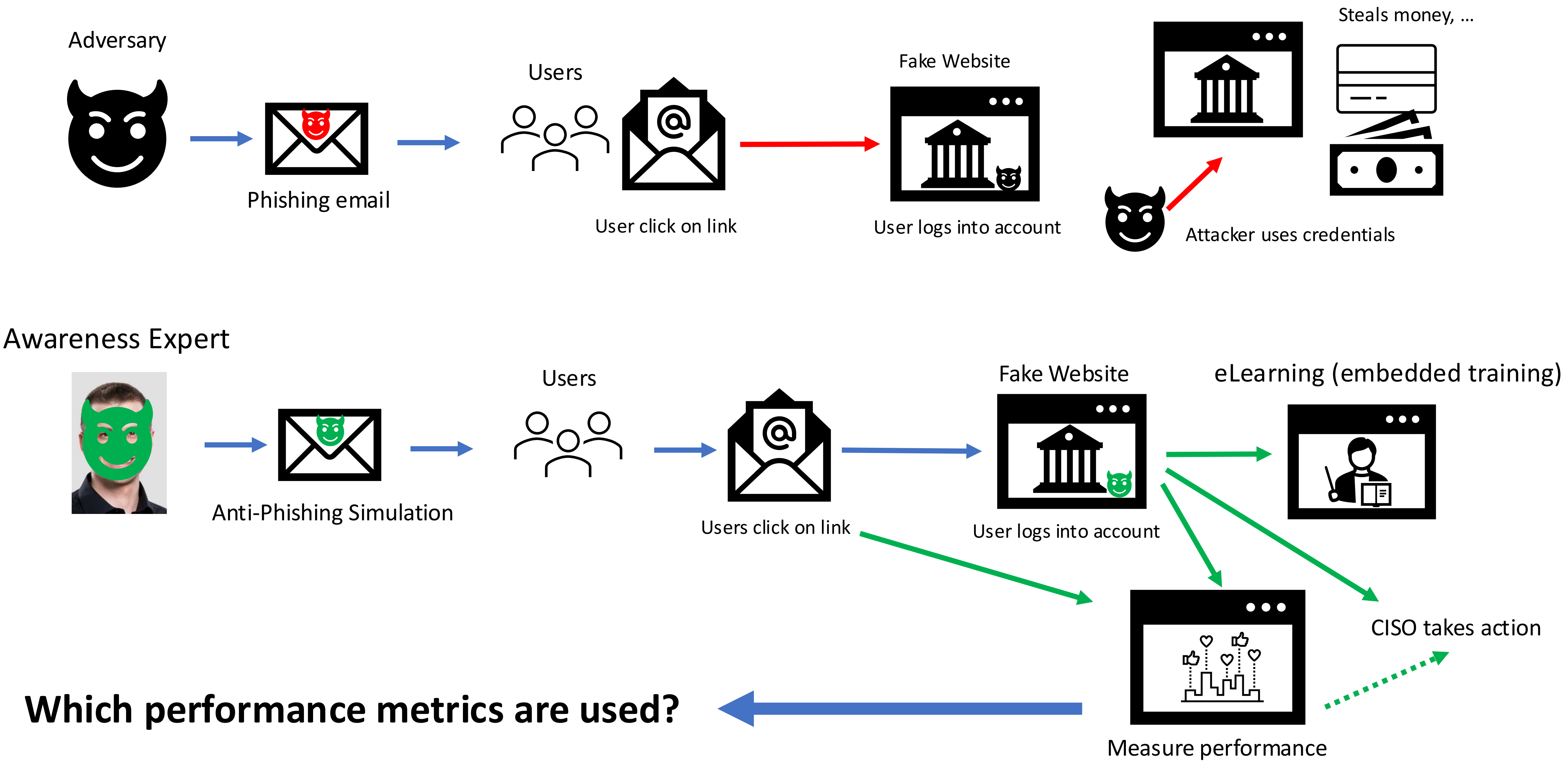


RESULTS

- We sent 288,000 emails in total with 144 email templates
- Users clicked 31,707 (11%) of the links in the anti-phishing emails.
- From the 31,707 clicked anti-phishing emails, 15,224 (48%) were successful credentials stealing attacks (submits)
- Only ~5.3% of the emails were fully successful.

Low compared to what companies say is normal....

PHISHING EXAMPLE



Adversary

Phishing email

Users

User click on link

Fake Website

User logs into account

Attacker uses credentials

Steals money, …

Awareness Expert

Anti-Phishing Simulation

Users

Users click on link

Fake Website

User logs into account

eLearning (embedded training)

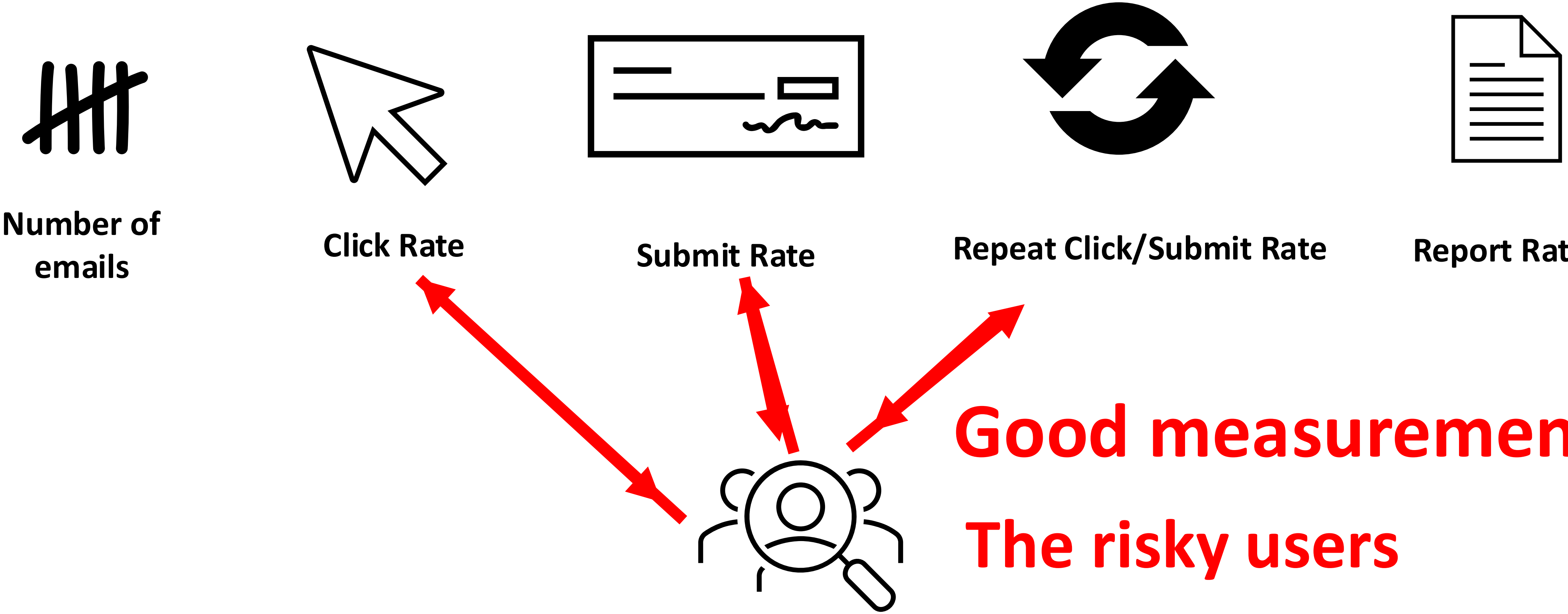
CISO takes action

Measure performance

Which performance metrics are used?

MEASURMENTS

Which performance metrics are used for anti-phishing training?



Number of emails

Click Rate

Submit Rate

Repeat Click/Submit Rate

Report Rate

Users that fall for phishing

Good measurements?

The risky users

PHISHING AWARENESS FOR SCIENCE

Zurich University of Applied Sciences

zh aw

~31’000 participants

1-year experiment

12 weeks

144 phishing emails

Users with wide varieety of professions

Scientific Setup

- Measured click- and submit rates
- Randomly selected groups
- Groups get different training content
- Control group with no training content
- No pre-training, no pre-information
- Every user got 12 anti-phishing emails

RESEARCH OBJECTIVES

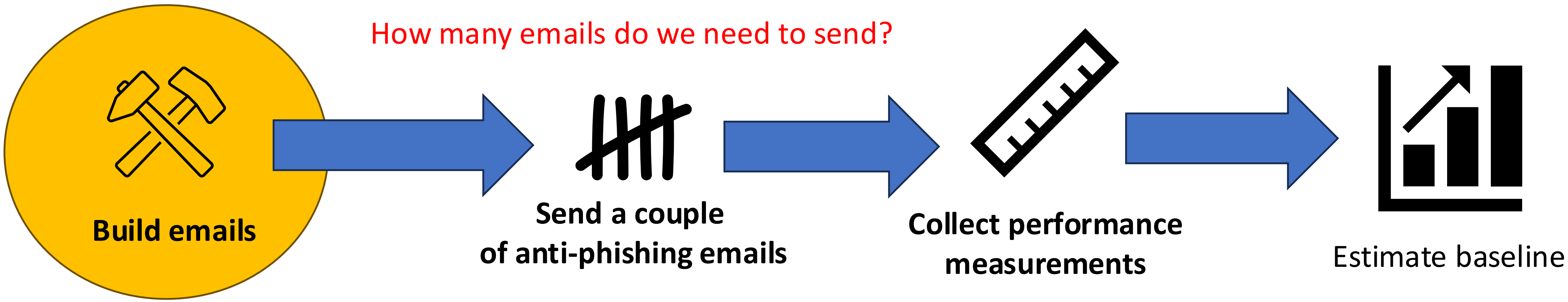
Subject: Training difficulty and phishing susceptibility

Can we predict how many people would click on a specific email?

Subject: Anti-Phishing metrics

How effective is current phishing awareness training?

THE PROBLEM WITH MEASUREMENTS



How many emails do we need to send?

Build emails

Send a couple of anti-phishing emails

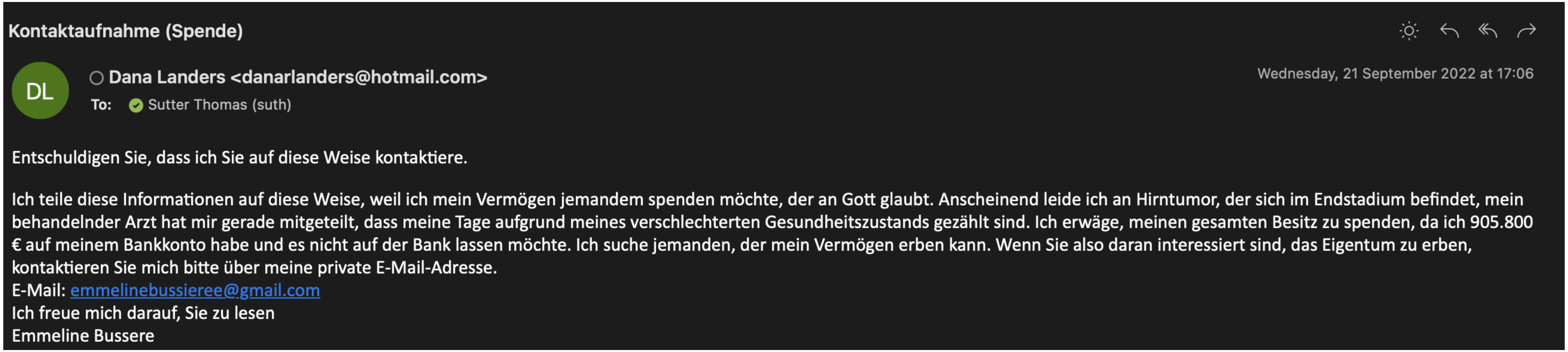
Collect performance measurements

Estimate baseline

Are the emails difficult enough for our users?
Do these emails meet our expectations?

How can we measure difficulty of the emails before sending?

ESTIMATING DIFFICULTY



Kontaktaufnahme (Spende)

○ Dana Landers <danarlanders@hotmail.com>

To: Sutter Thomas (suth)

Wednesday, 21 September 2022 at 17:06

Entschuldigen Sie, dass ich Sie auf diese Weise kontaktiere.

Ich teile diese Informationen auf diese Weise, weil ich mein Vermögen jemandem spenden möchte, der an Gott glaubt. Anscheinend leide ich an Hirntumor, der sich im Endstadium befindet, mein behandelnder Arzt hat mir gerade mitgeteilt, dass meine Tage aufgrund meines verschlechterten Gesundheitszustands gezählt sind. Ich erwäge, meinen gesamten Besitz zu spenden, da ich 905.800 € auf meinem Bankkonto habe und es nicht auf der Bank lassen möchte. Ich suche jemanden, der mein Vermögen erben kann. Wenn Sie also daran interessiert sind, das Eigentum zu erben, kontaktieren Sie mich bitte über meine private E-Mail-Adresse.
E-Mail: emmelinebussieree@gmail.com
Ich freue mich darauf, Sie zu lesen
Emmeline Bussere

Email address is different from sender

Salutation is not personal

Offers money

Urgency tone

...

Clues or hints

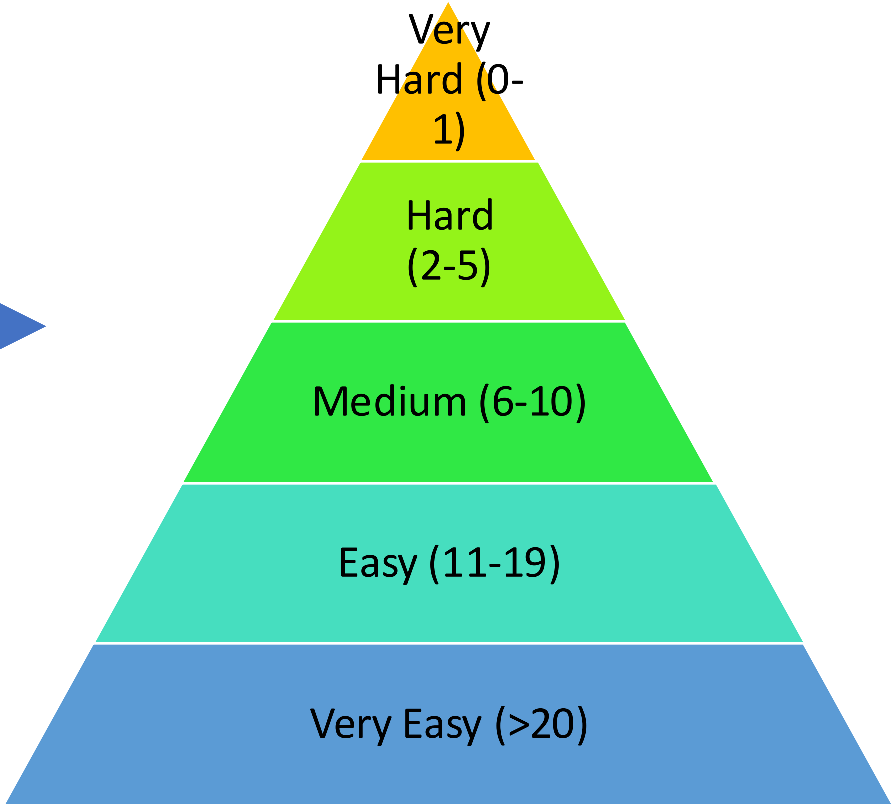
Create a rating system based on number of clues and some weights:

Clue / Feature	Weight
Typo squatting	5
Familiar sender	4
Use company logo	2
Use hidden link	2
Use signature in email	3
…	

Static Rating System

Take some math function and define a threshold.

For example, sum()



Very Hard (0-1)

Hard (2-5)

Medium (6-10)

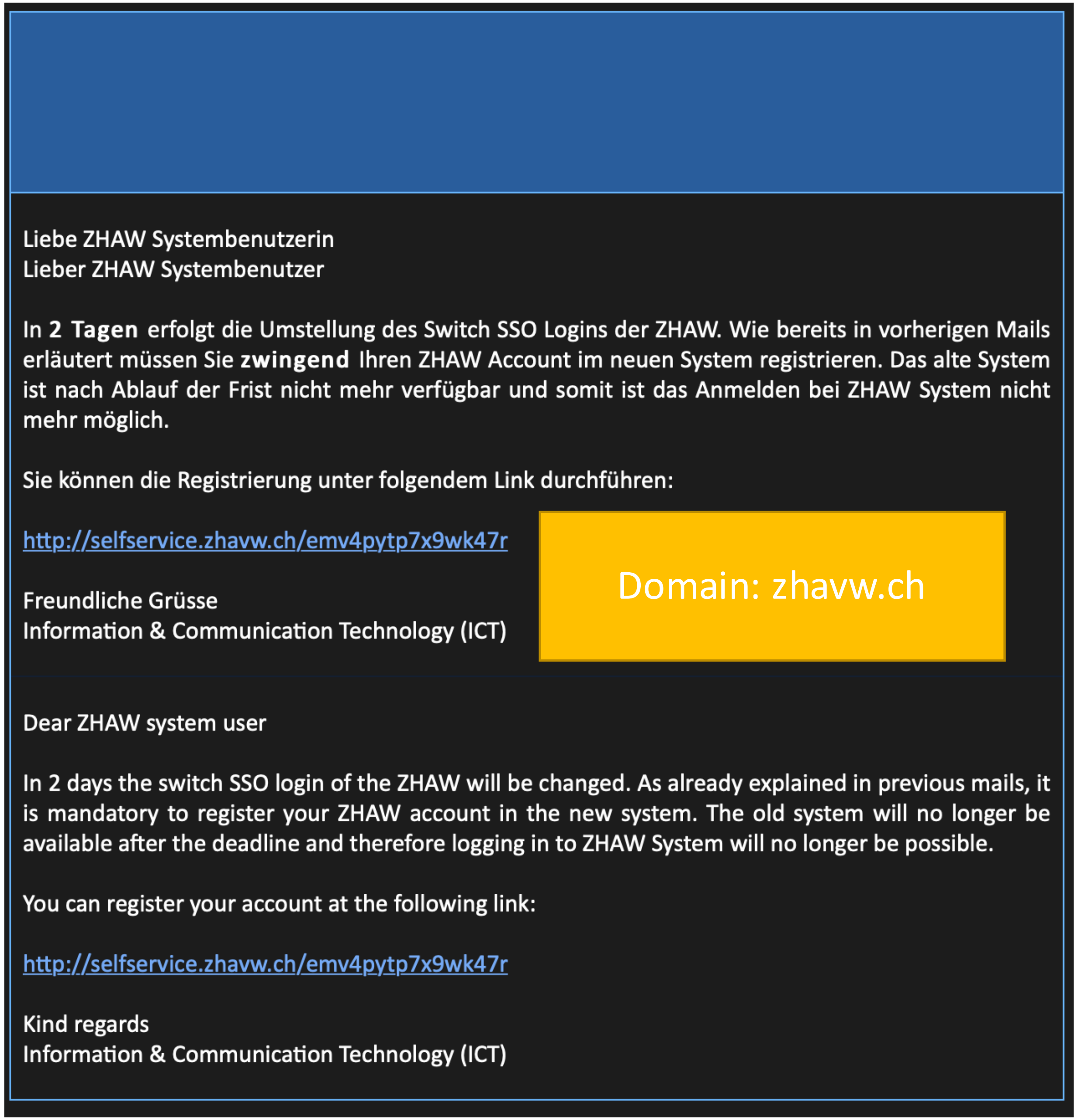
Easy (11-19)

Very Easy (>20)

Some example rating system

How do we configure the weights?

This email uses only a typo squatting clue.



Liebe ZHAW Systembenutzerin
Lieber ZHAW Systembenutzer

In **2 Tagen** erfolgt die Umstellung des Switch SSO Logins der ZHAW. Wie bereits in vorherigen Mails erläutert müssen Sie **zwingend** Ihren ZHAW Account im neuen System registrieren. Das alte System ist nach Ablauf der Frist nicht mehr verfügbar und somit ist das Anmelden bei ZHAW System nicht mehr möglich.

Sie können die Registrierung unter folgendem Link durchführen:

http://selfservice.zhavw.ch/emv4pytp7x9wk47r

Freundliche Grüsse
Information & Communication Technology (ICT)

Domain: zhavw.ch

Dear ZHAW system user

In 2 days the switch SSO login of the ZHAW will be changed. As already explained in previous mails, it is mandatory to register your ZHAW account in the new system. The old system will no longer be available after the deadline and therefore logging in to ZHAW System will no longer be possible.

You can register your account at the following link:

http://selfservice.zhavw.ch/emv4pytp7x9wk47r

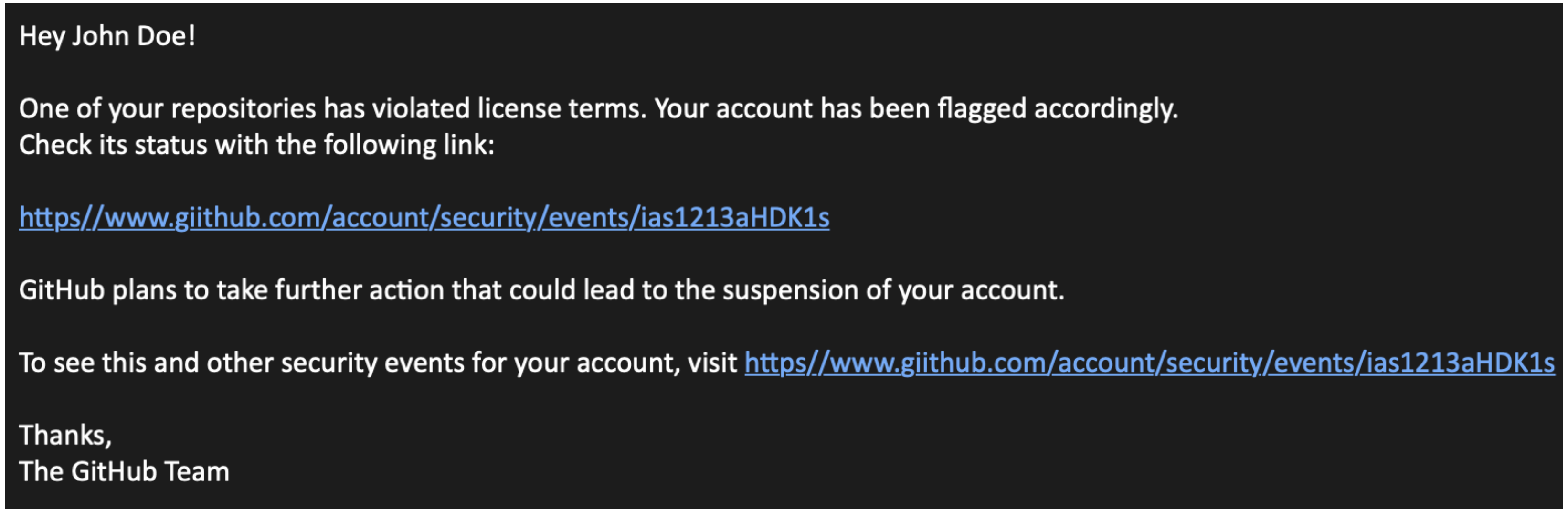
Kind regards
Information & Communication Technology (ICT)

Emails	Clicks
500	225
500	205
500	262
500	283

Click Rate 48.75%

Are all typo squatted emails hard for users to detect?

This email uses only a typo squatting clue.



Hey John Doe!

One of your repositories has violated license terms. Your account has been flagged accordingly.
Check its status with the following link:

https//www.giithub.com/account/security/events/ias1213aHDK1s

GitHub plans to take further action that could lead to the suspension of your account.

To see this and other security events for your account, visit https//www.giithub.com/account/security/events/ias1213aHDK1s

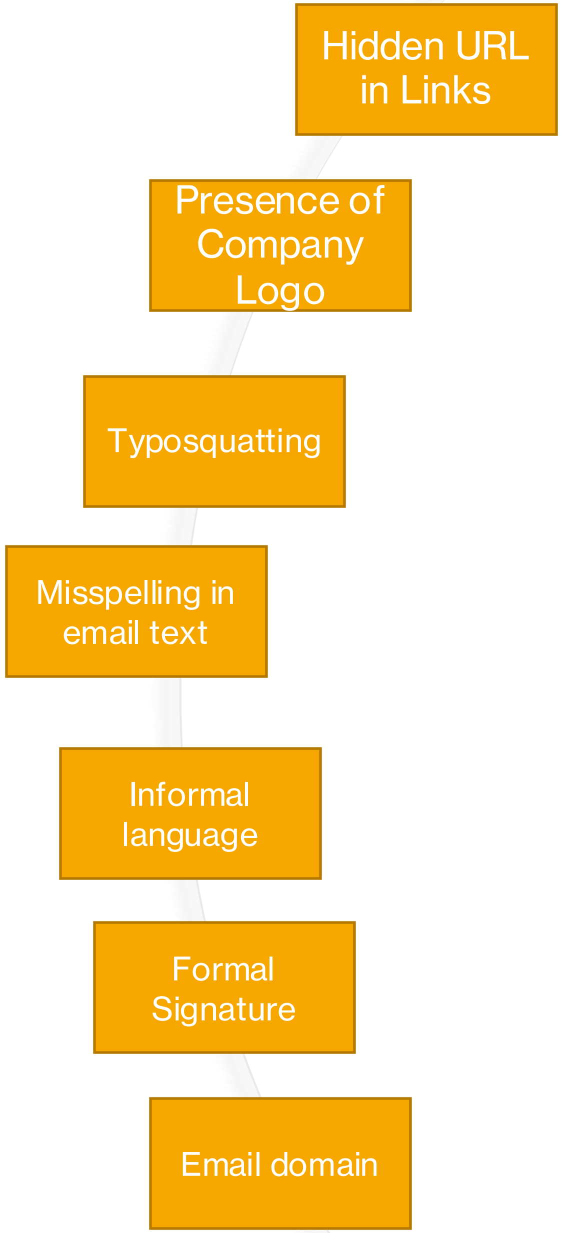
Thanks,
The GitHub Team

Domain: giithub.ch

There are other factors than just the content….

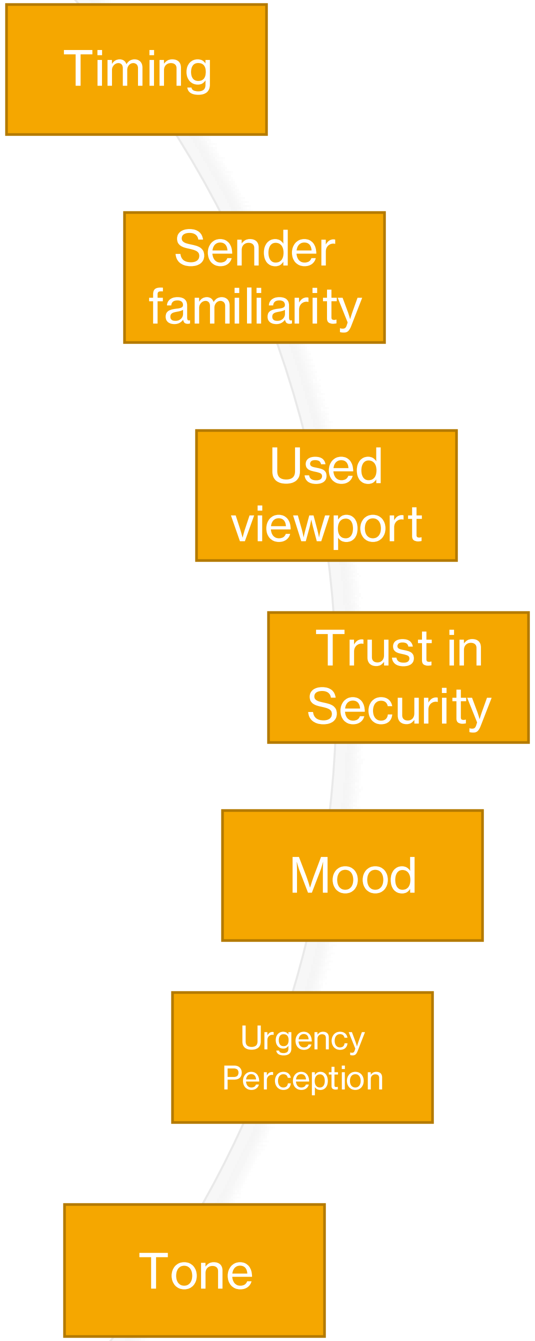
….seven users (1.39%) clicked

Measurable



- Hidden URL in Links
- Presence of Company Logo
- Typosquatting
- Misspelling in email text
- Informal language
- Formal Signature
- Email domain

Not measurable



- Timing
- Sender familiarity
- Used viewport
- Trust in Security
- Mood
- Urgency Perception
- Tone

Not all influential factors can be measured

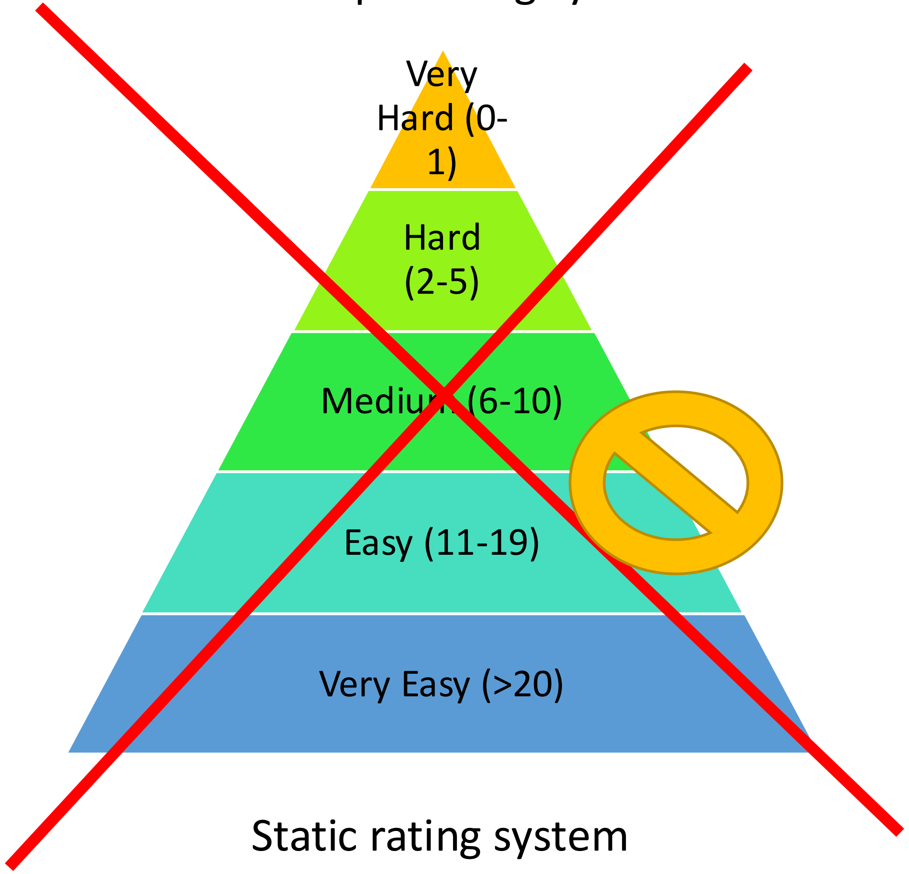
Difficulty is relative

External factors play a key role in phishing

What’s easy for one person can be hard for another…

…but difficulty is relative to the user's perception

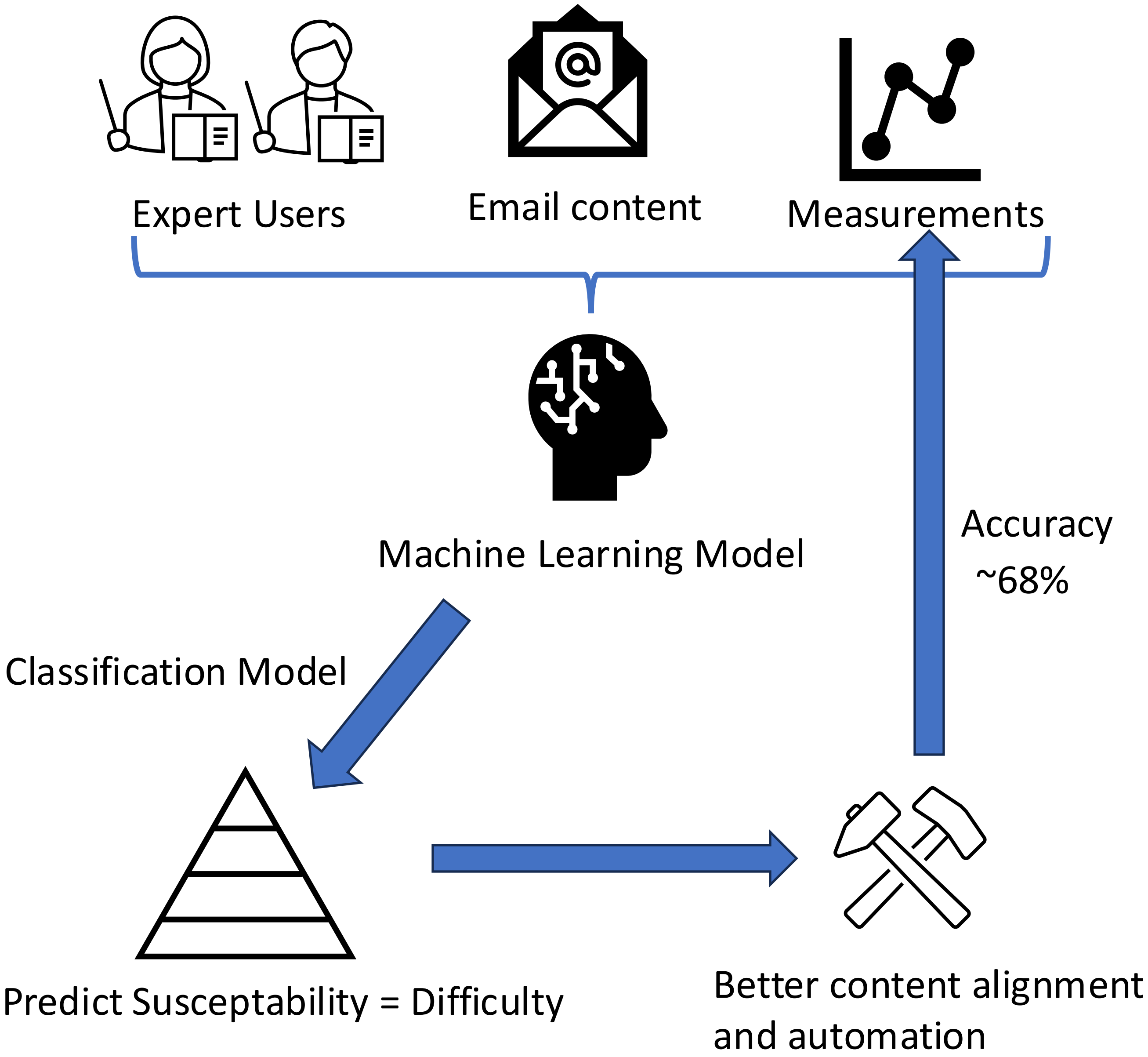
Some example rating system



Static rating system

Doesn’t really work well

Data driven approach to measure susceptability



Expert Users

Email content

Measurements

Machine Learning Model

Classification Model

Accuracy ~68%

Predict Susceptability = Difficulty

Better content alignment and automation

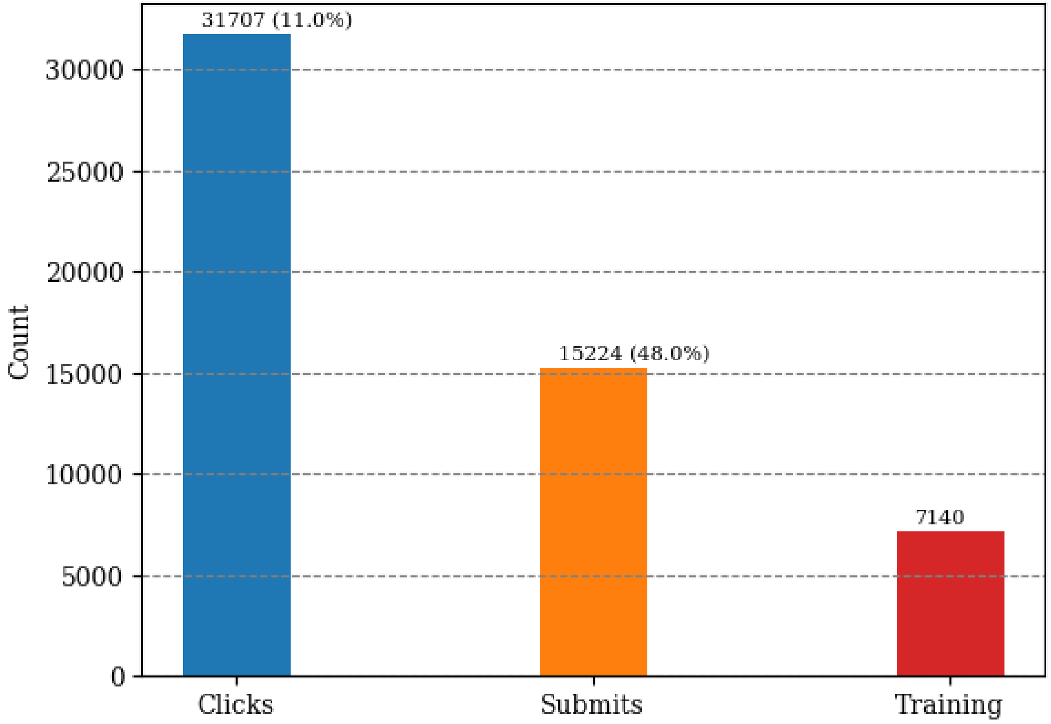
RESULTS

- We sent 288,000 emails in total with 144 email templates

- Users clicked 31,707 (11%) of the links in the anti-phishing emails.

- From the 31,707 clicked anti-phishing emails, 15,224 (48%) were successful credentials stealing attacks (submits)

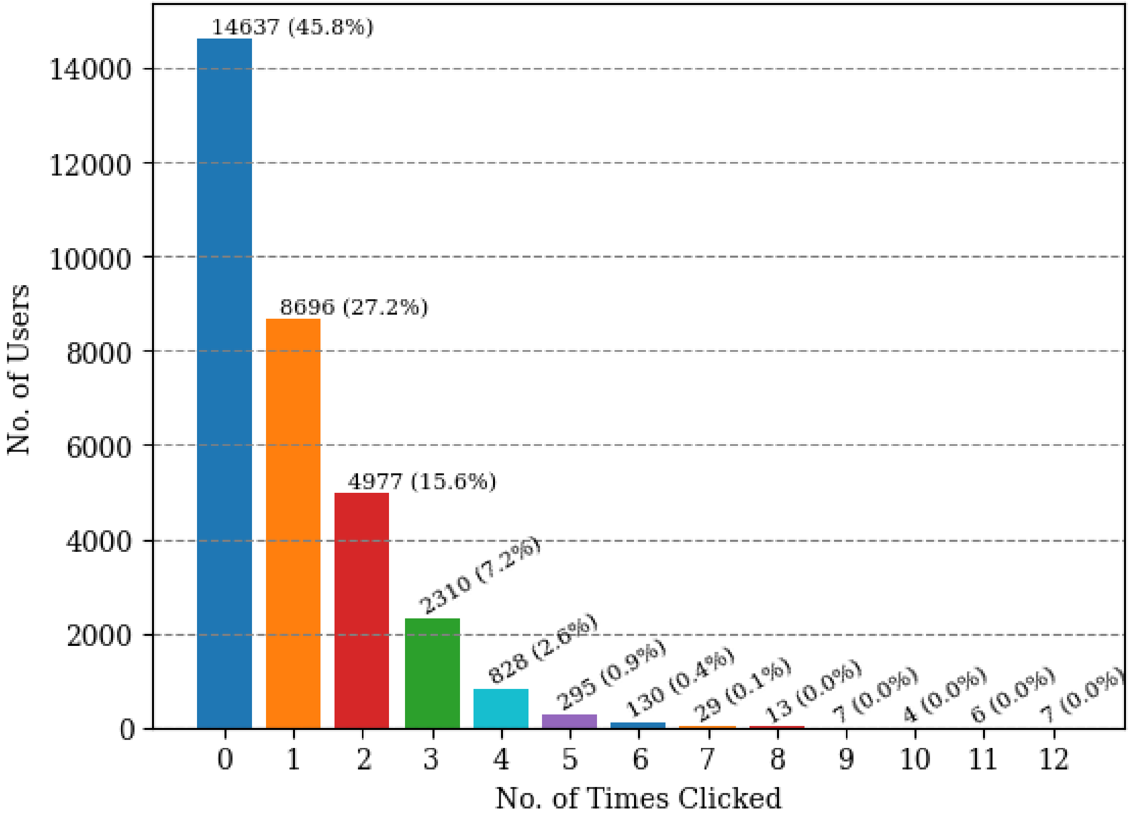
- Only ~5.3% of the emails were fully successful.



Low compared to what companies say is normal….

REPEATED CLICKERS

How many times would the same user click?

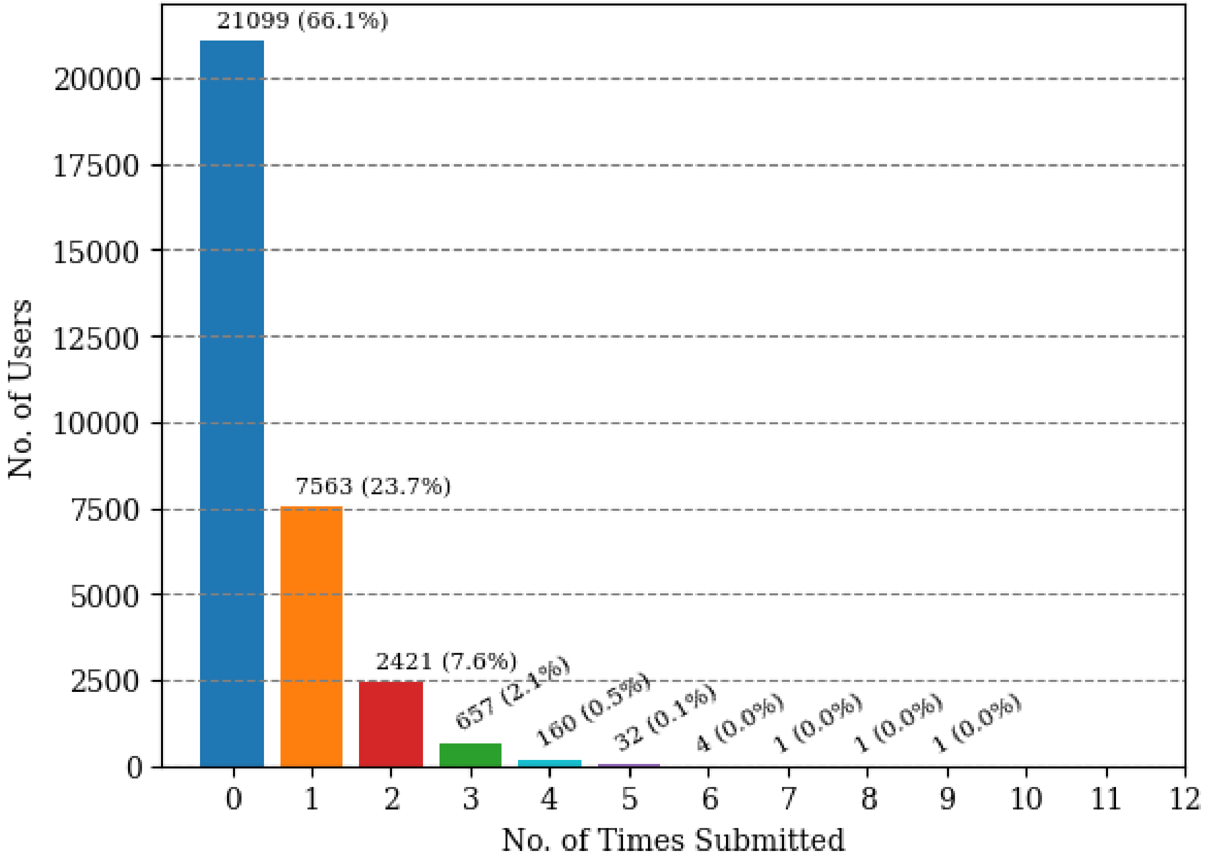


Repeat Click Rate

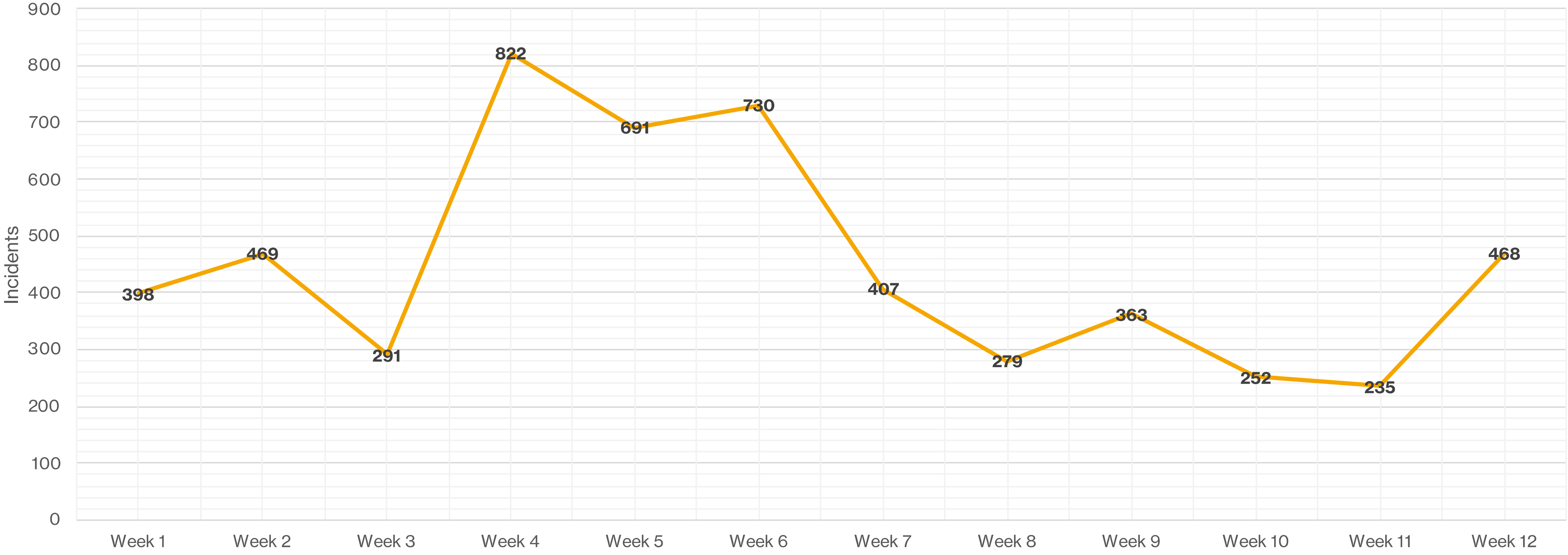
SUBMIT RATE

How many times would the same user submit?

Repeat Submit Rate



ESTIMATE OF SUPPORT TICKETS



LESSONS LEARNED

Subject: Training difficulty and phishing susceptibility

Can we predict how many people would click on a specific mail?

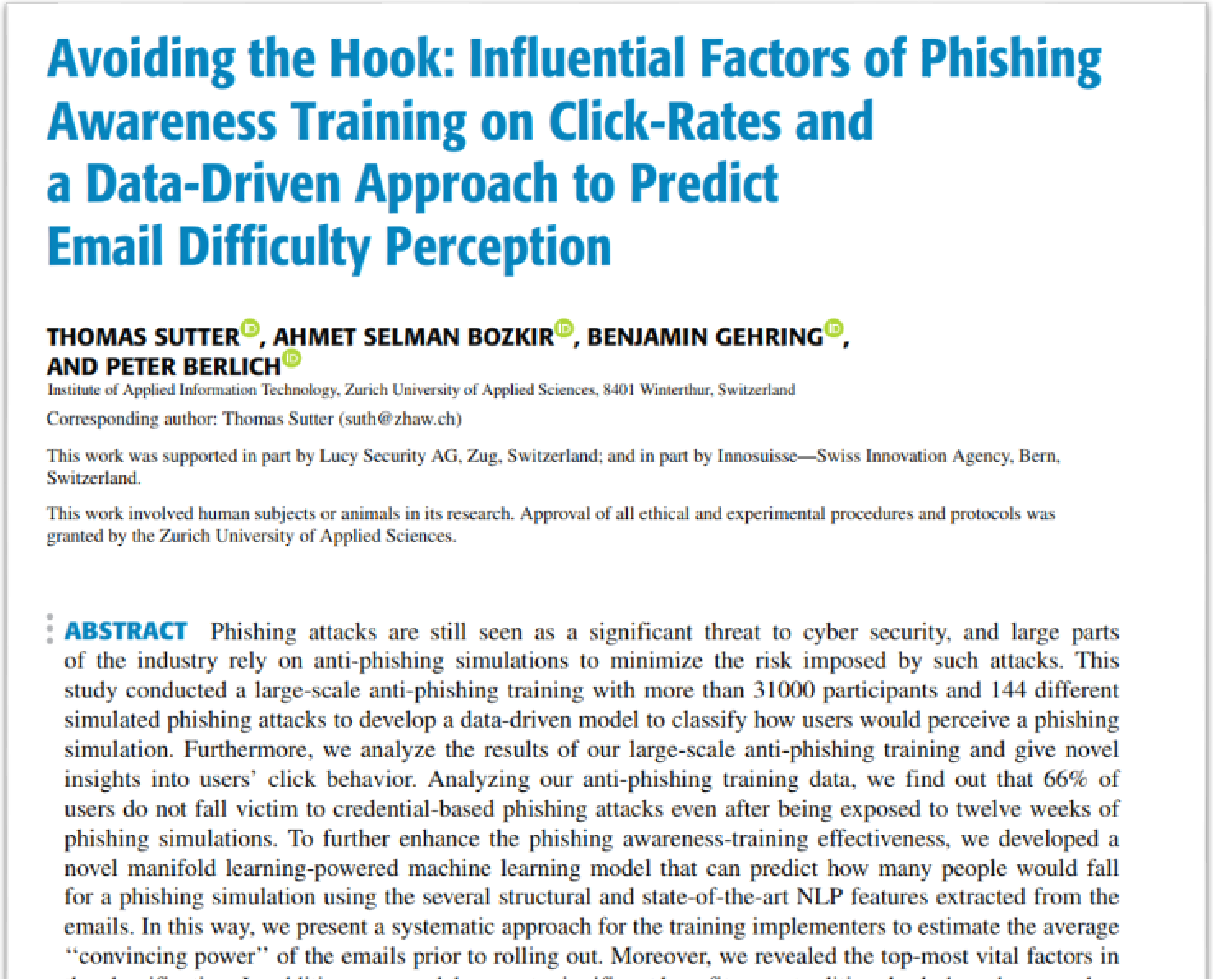
Yes, with a data driven approach and sufficient user data

Subject: Anti-Phishing metrics

How effective is current phishing awareness training?

- Current measurements do not take the difficulty into account
- 66% of users do not fall victim to credential-based phishing attacks even after being exposed to twelve weeks of phishing simulation!
  - Bothering this users with anti-phishing training is a waste of time.

We do innovation research



Avoiding the Hook: Influential Factors of Phishing Awareness Training on Click-Rates and a Data-Driven Approach to Predict Email Difficulty Perception

THOMAS SUTTER, AHMET SELMAN BOZKIR, BENJAMIN GEHRING, AND PETER BERLICH
Institute of Applied Information Technology, Zurich University of Applied Sciences, 8401 Winterthur, Switzerland
Corresponding author: Thomas Sutter (suth@zhaw.ch)

This work was supported in part by Lucy Security AG, Zug, Switzerland; and in part by Innosuisse—Swiss Innovation Agency, Bern, Switzerland.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Zurich University of Applied Sciences.

ABSTRACT Phishing attacks are still seen as a significant threat to cyber security, and large parts of the industry rely on anti-phishing simulations to minimize the risk imposed by such attacks. This study conducted a large-scale anti-phishing training with more than 31000 participants and 144 different simulated phishing attacks to develop a data-driven model to classify how users would perceive a phishing simulation. Furthermore, we analyze the results of our large-scale anti-phishing training and give novel insights into users' click behavior. Analyzing our anti-phishing training data, we find out that 66% of users do not fall victim to credential-based phishing attacks even after being exposed to twelve weeks of phishing simulations. To further enhance the phishing awareness-training effectiveness, we developed a novel manifold learning-powered machine learning model that can predict how many people would fall for a phishing simulation using the several structural and state-of-the-art NLP features extracted from the emails. In this way, we present a systematic approach for the training implementers to estimate the average "convincing power" of the emails prior to rolling out. Moreover, we revealed the top-most vital factors in