

# How to Use Machine Learning for a Phishing Incident Response

---

Erez Harush

Demisto, Palo Alto Networks  
FloCon 2020



# WHOAMI



---

**Erez Harush**

Palo Alto Networks, Demisto

Data scientist

# Agenda

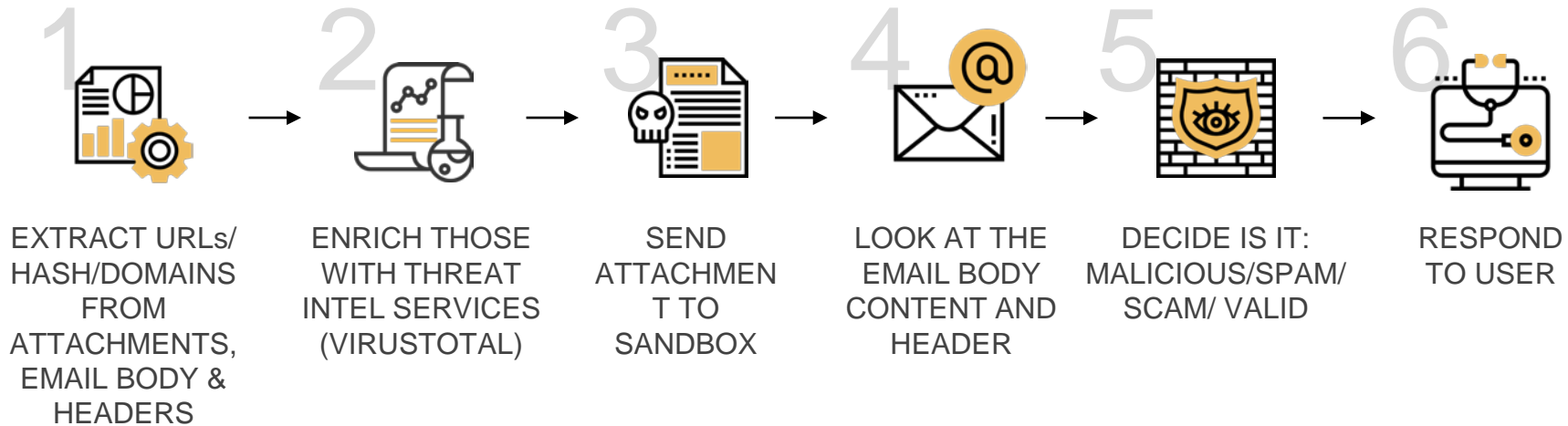
- Introduction to phishing incidents and response processes at SOC teams
- Phishing problem definition
- Datasets used
- Process followed to build the model
- Model deployment
- Q&A

# Phishing Incidents and Response Processes at SOC Teams

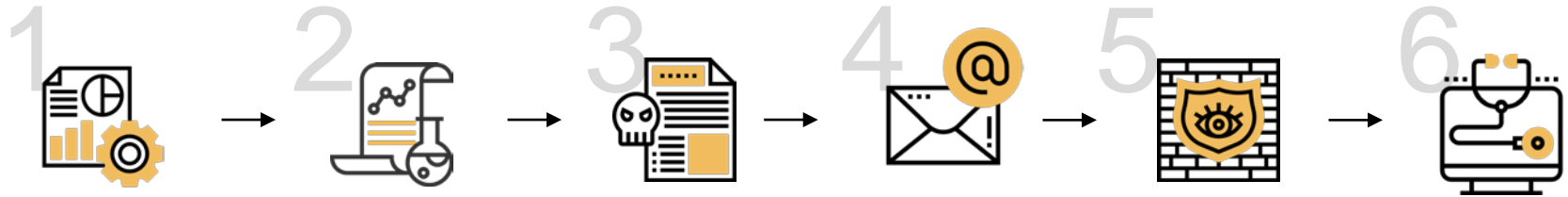


- What is a phishing alert?
  - User receives a suspicious email, so forwards it to [phishing@organization.com](mailto:phishing@organization.com)
- What happens behind the scenes?
  - SOC (security operation center) analyst is assigned to investigate the suspicious email

# Classic Phishing Handling Process

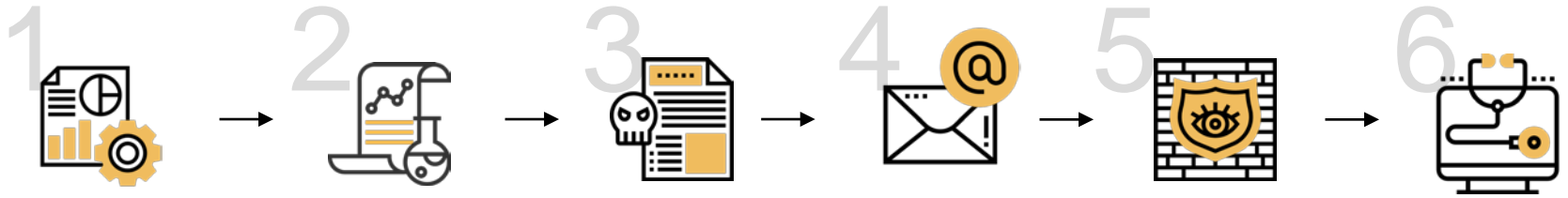


# Classic Phishing Handling Process



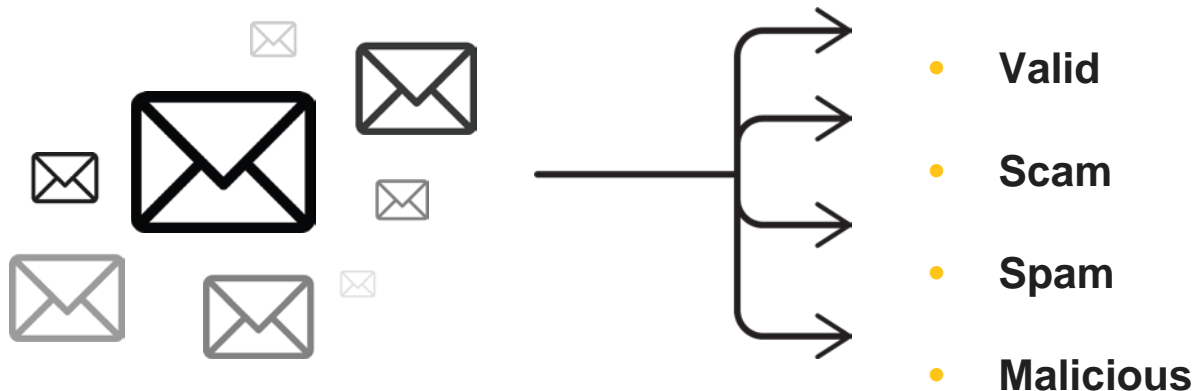
SOC teams waste a  
**significant amount of time**  
investigating these emails

# Goals Definition



- Decrease the amount of time to make a decision
- Help the analyst make a better decision

# Datasets

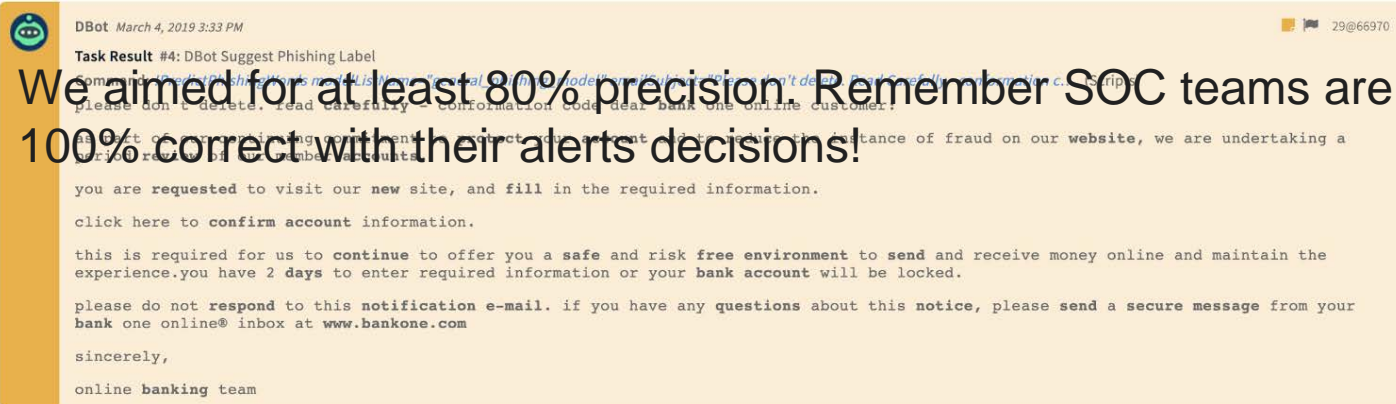


- We collected **~100K labeled emails** from 10 different customers
- All the emails are suspicious emails, other than the most email gateway datasets



# Process Followed to Build the Model

- Research: email headers, IOCs, email body
- Text classification model: FastText (word embedding & NN)
- Explainability - highlighting important words, using Lime  
<https://github.com/marcotcr/lime>



- We aimed for at least 80% precision. Remember SOC teams are not 100% correct with their alerts decisions!

# Example: Customer A

- Categories: Malicious, Other
- Precision & Recall: ~95% for each class

		Predicted		
		Malicious	Other	All
True	Malicious	722	45	767
	Other	18	317	335
	All	740	362	1102

~95%

~95%

# Example: Customer B

- Categories: Malicious, Valid, Spam

**Precision is high, but the coverage is very low**  
**Model confidence is low for malicious emails**

**Probability > 0.85**

Predicted True	Valid	Malicious	Spam	All
Valid	60%	0	9	83
Malicious	2	24	14	40
Spam	6	3	291	300
All	82	27	314	423

Predicted True	Valid	Malicious	Spam	All
Valid	48	0	↓ 75%	50
Malicious	0	6	3	9
Spam	2	1	↓ 25%	259
All	50	7	261	318

# Combine Internal Datasets

- Text normalization using SpaCy
  - Word lemmatization
  - Replace email addresses a const string
  - Replace URLs with a const string
- Fine tuning of a pre-trained model (based on internal datasets) using the customer data

# Example: Customer B

Original confusion matrix

Predicted True	Valid	Malicious	Spam	All
Valid	74	0	9	83
Malicious	2	24	14	40
Spam	6	3	291	300
All	82	27	314	423

Combine other customer data  
Probability > 0.85

Predicted True	Valid	Malicious	Spam	All
Valid	48	5	1	54
Malicious	1	25	2	30
Spam	1	5	218	221
All	50	30	218	303

Coverage

~88%

~83%

~97%

Precision ~96%

~83%

~98%



# Model Deployment

- We support 2 ways of using our phishing models:
  - a. Building a dedicated model based on a specific customer env within Demisto on a permanent env
    - Supervised
    - Semi-Supervised
  - a. Upon request, Demisto DS builds a model and deploys it within Demisto



# Using a Model in Demisto

- Use as part of scoring/severity set
- Close incidents automatically with probability > THRESHOLD
- Handle incidents that were not handled in the past due to low capacity

Command: `!PredictPhishingWords modelListName="general_phishing_model" emailSubject="Please don't delete. Read Carefully - conformation c..."` (Scripts)

## DBot label suggestion

Label	malicious
Probability	0.79

# Using a Model in Demisto



DBot March 4, 2019 3:33 PM

29@66970

**Task Result #4:** DBot Suggest Phishing Label

**Command:** `!PredictPhishingWords modelListName="general_phishing_model" emailSubject="Please don't delete. Read Carefully - conformation c..."` (Scripts)

please don't delete. read carefully - conformation code dear bank one online customer:

as part of our continuing commitment to protect your account and to reduce the instance of fraud on our website, we are undertaking a period review of our member accounts.

you are requested to visit our new site, and fill in the required information.

click here to confirm account information.

this is required for us to continue to offer you a safe and risk free environment to send and receive money online and maintain the experience. you have 2 days to enter required information or your bank account will be locked.

please do not respond to this notification e-mail. if you have any questions about this notice, please send a secure message from your bank one online@ inbox at www.bankone.com

sincerely,

online banking team



# Building a Model in Demisto Platform



You are awesome!

All your tasks are done



Q&A

# *THANK YOU*

Email: [eharush@paloaltonetworks.com](mailto:eharush@paloaltonetworks.com) | Twitter: @PaloAltoNtwks

