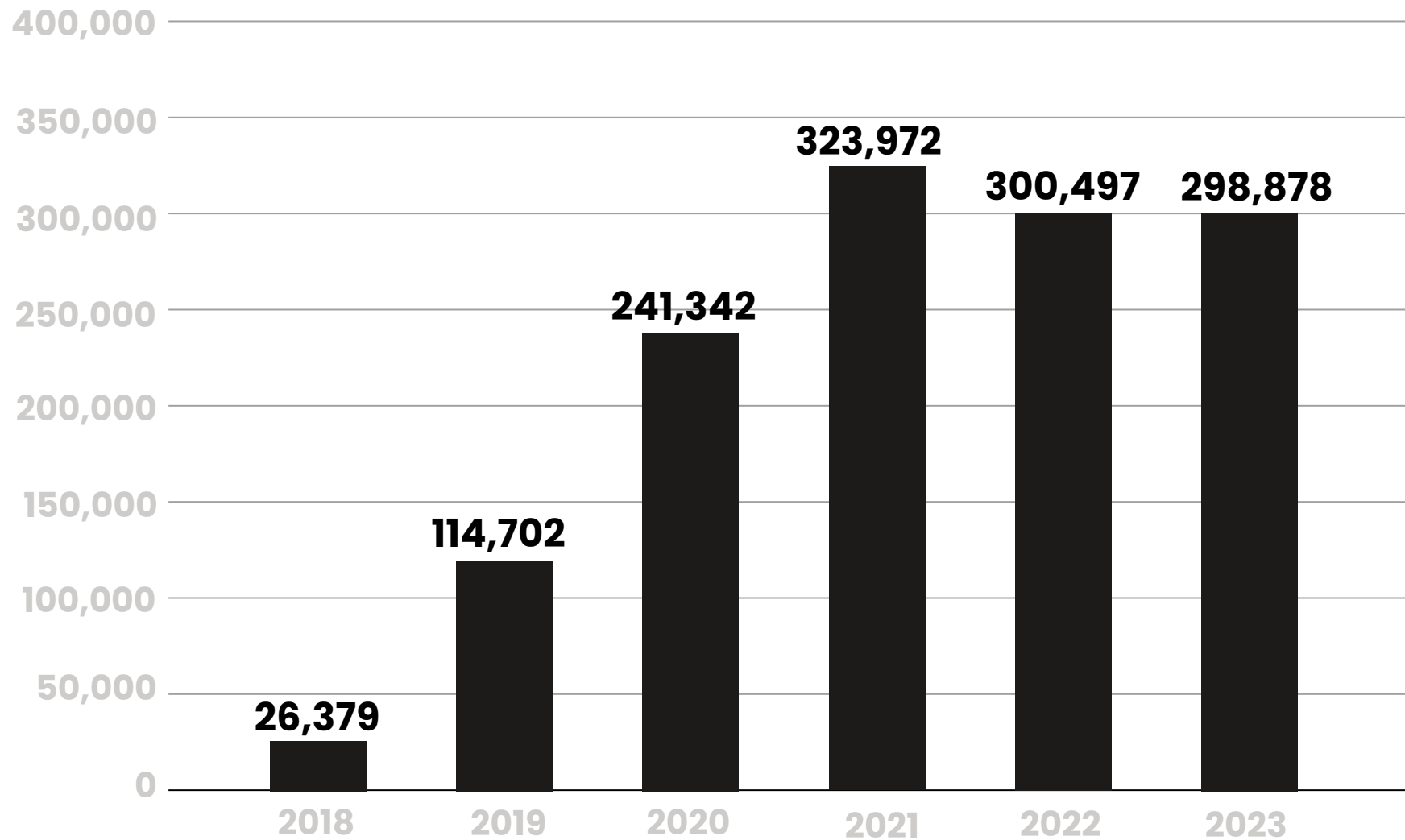


Less Defined Knowledge and More True Alarms: Reference-based Phishing Detection without a Pre-defined Reference List

Ruofan Liu, Yun Lin, Xiwen Teoh,
Gongshen Liu, Zhiyong Huang, Jin Song Dong











Shanghai Jiao Tong University,
National University of Singapore

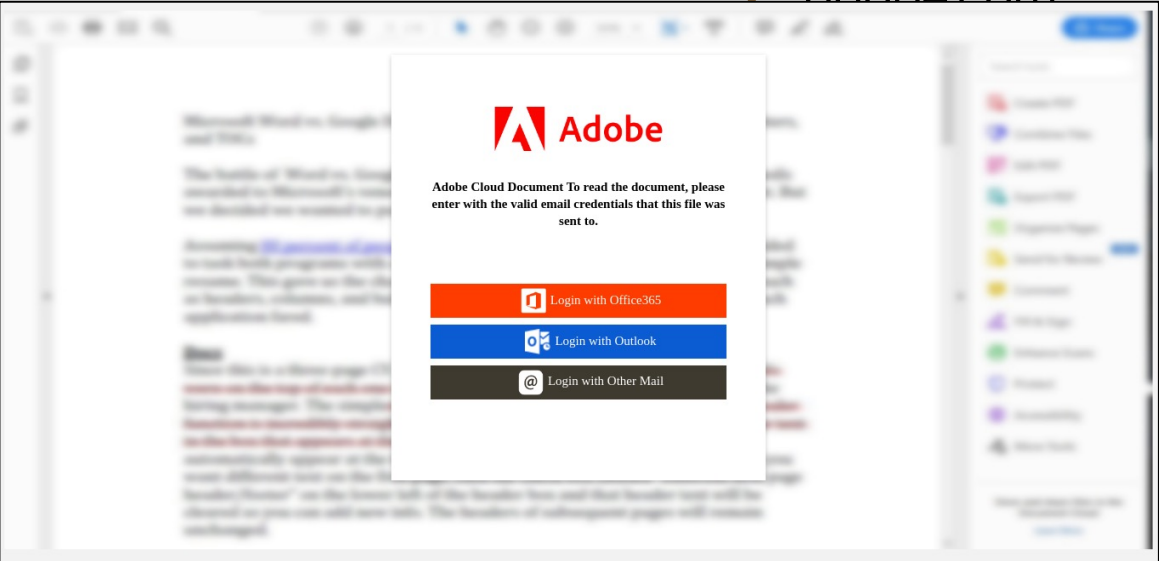




Number of phishing attack victims in the United States [1]

Reference-based Phishing Detection

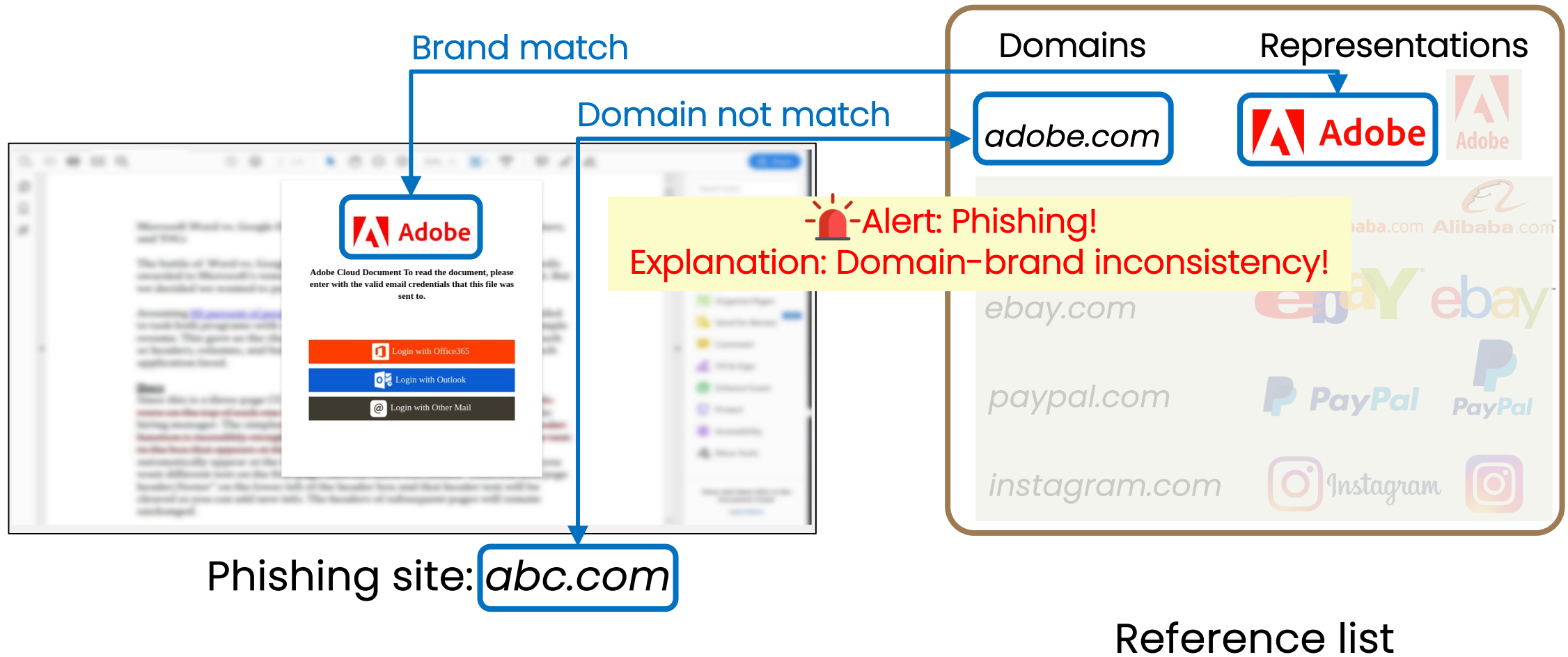
Domains	Representations
adobe.com	 
	 
	 
	 
	 










Phishing site: *abc.com*

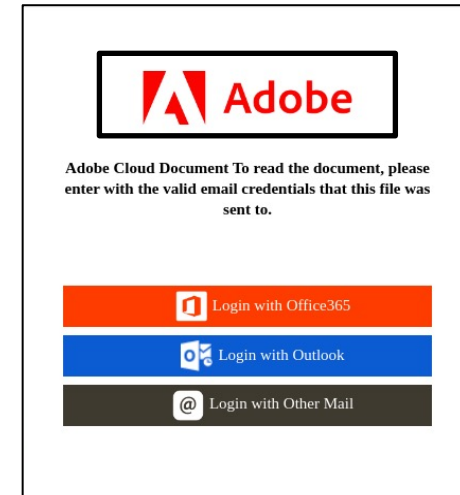
Reference list

Reference-based Phishing Detection



Domains	Representations
adobe.com	 Adobe  Adobe
alibaba.com	 Alibaba.com  Alibaba.com
ebay.com	 eBay  ebay
paypal.com	 PayPal  PayPal
instagram.com	 Instagram 

Reference list



"Look" at the logo

2020

2021

2022

2023

2024

VisualPhishNet
(CCS)











Phishpedia
(USENIX Sec)

PhishIntention
(USENIX Sec)

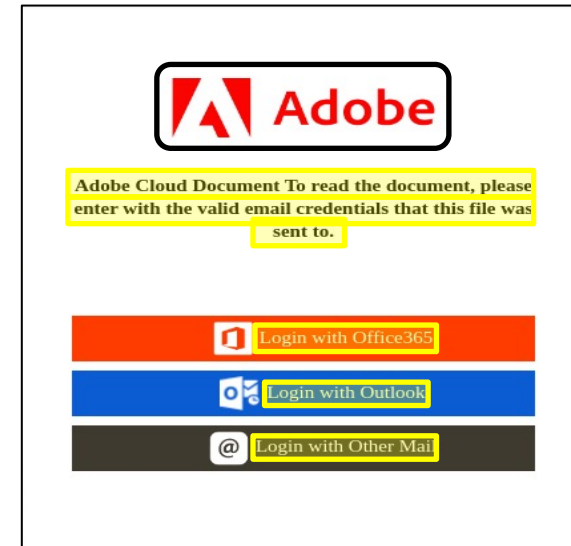
DynaPhish
(USENIX Sec)

NEW
PhishLLM
(Ours)

PhishLLM

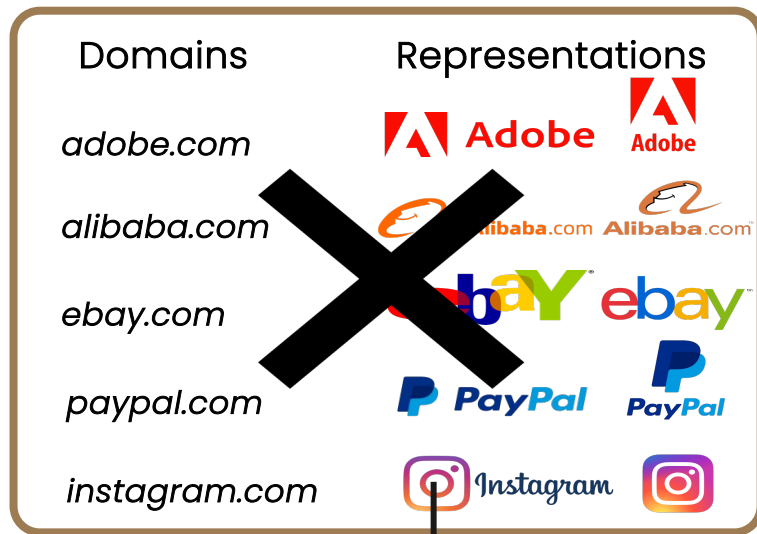
Domains	Representations
adobe.com	 
alibaba.com	 
ebay.com	 
paypal.com	 
instagram.com	 

Reference list free



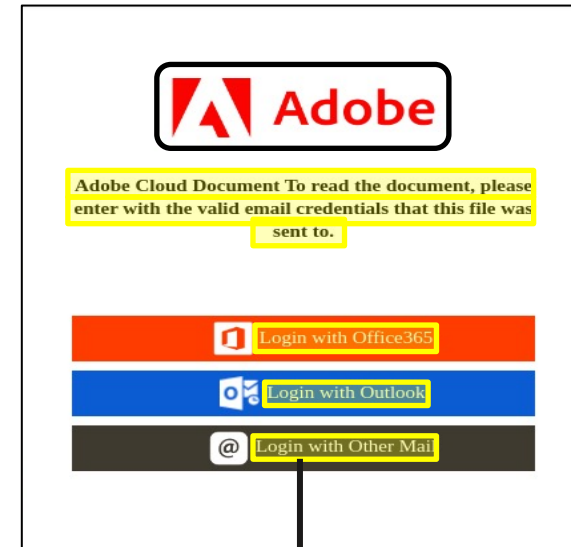
"Look" and "Read"

PhishLLM



Reference list free

LLM is a comprehensive knowledge base



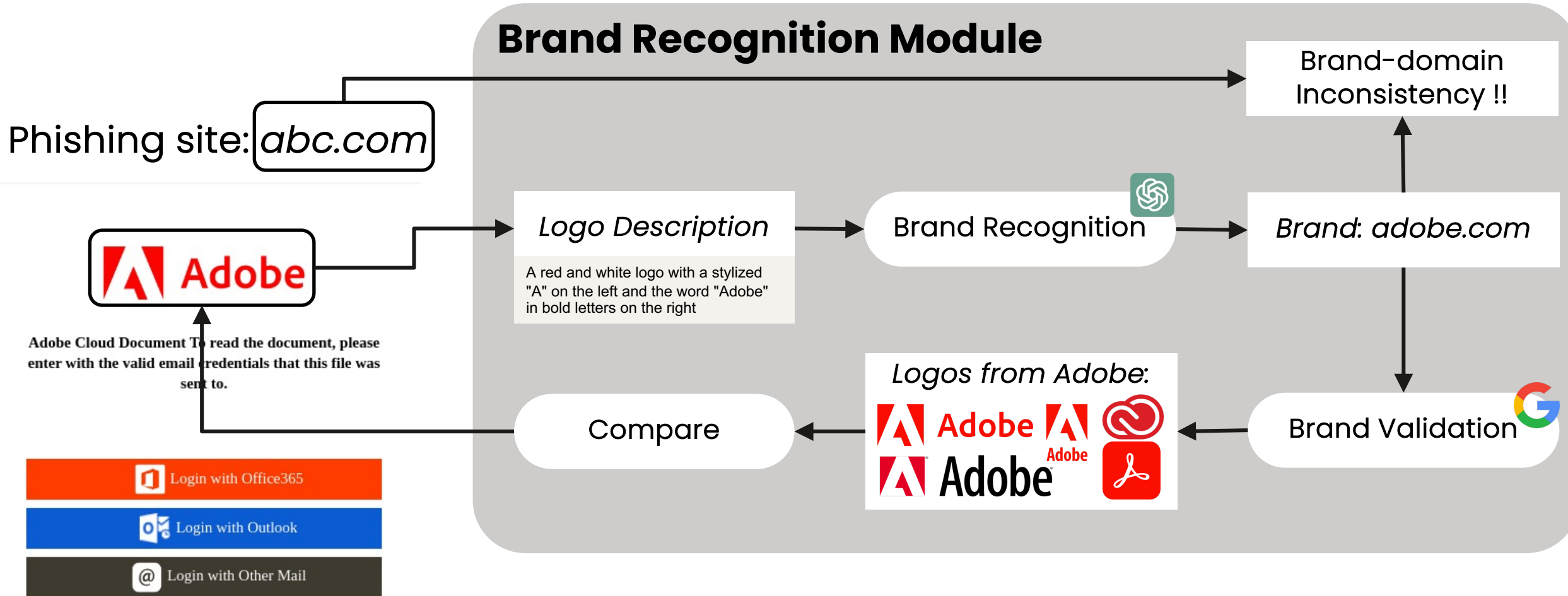
"Look" and "Read"

LLM is good at understanding and reasoning with textual semantics

Reference list free

"Look" and "Read"

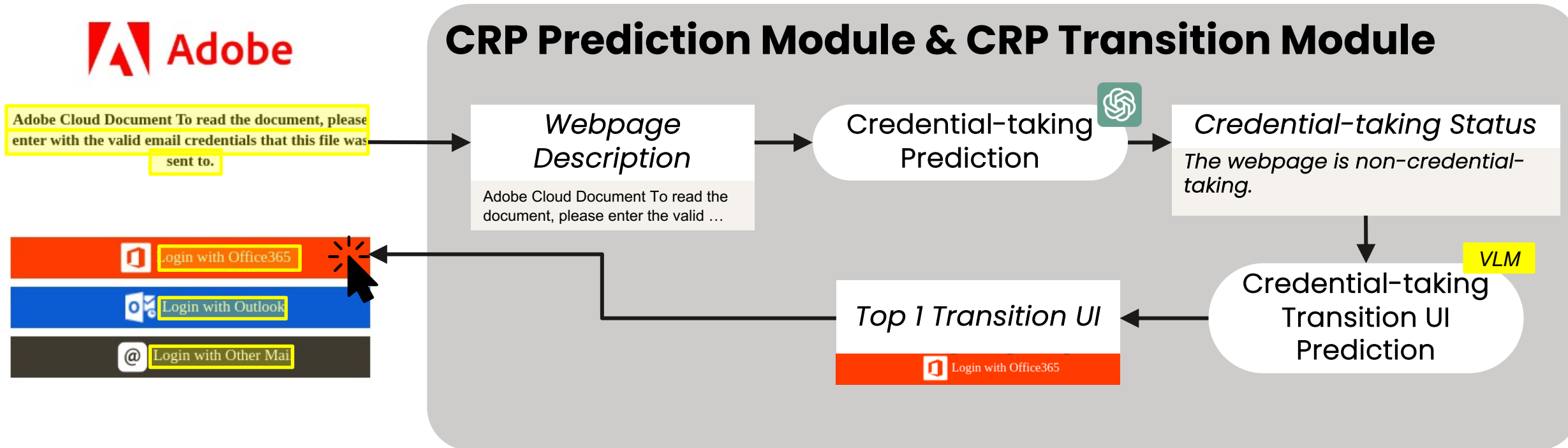
PhishLLM



Reference list free
"Look" and "Read"

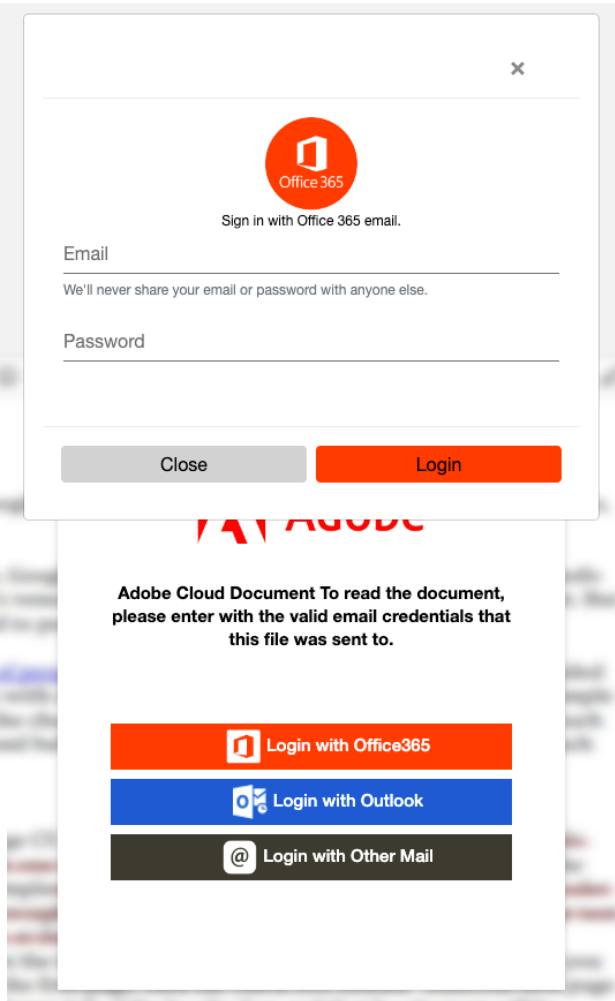
PhishLLM

Phishing site: *abc.com*

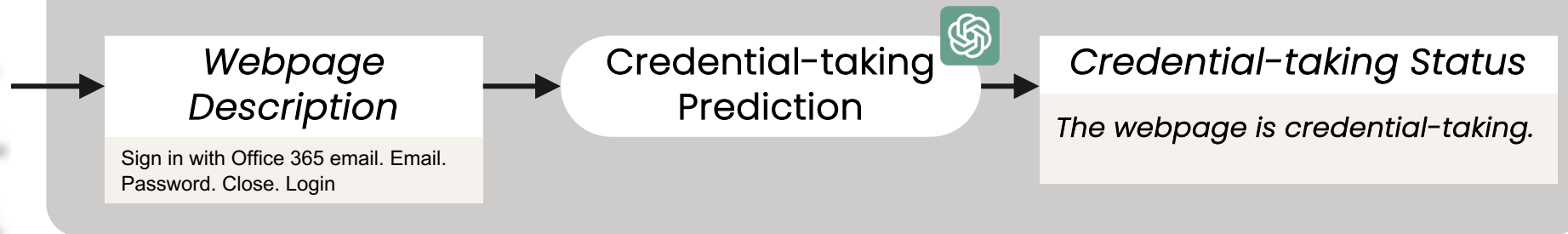


PhishLLM

Phishing site: *abc.com*



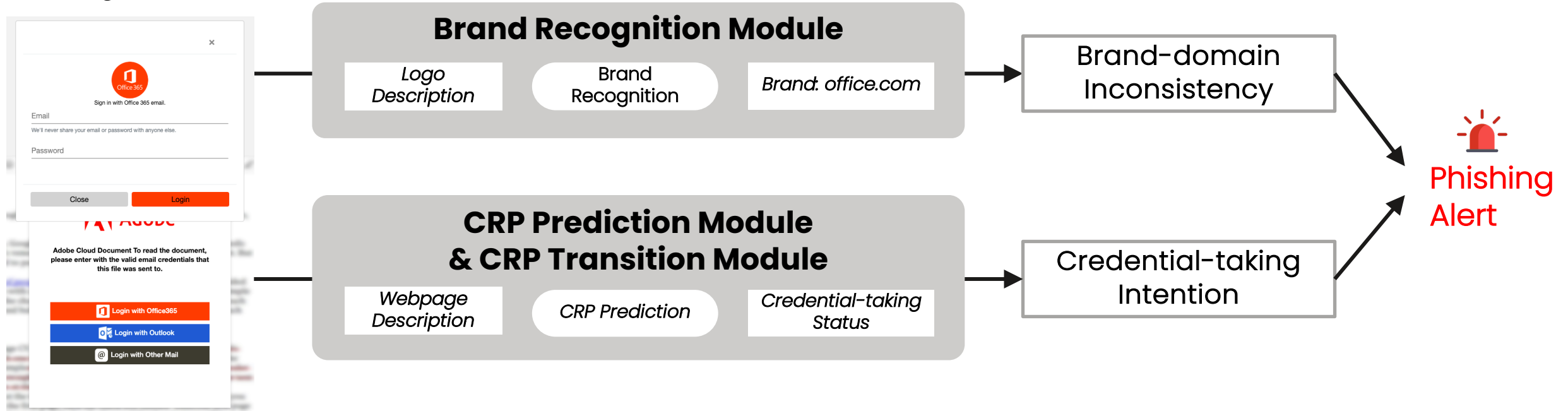
CRP Prediction Module & CRP Transition Module



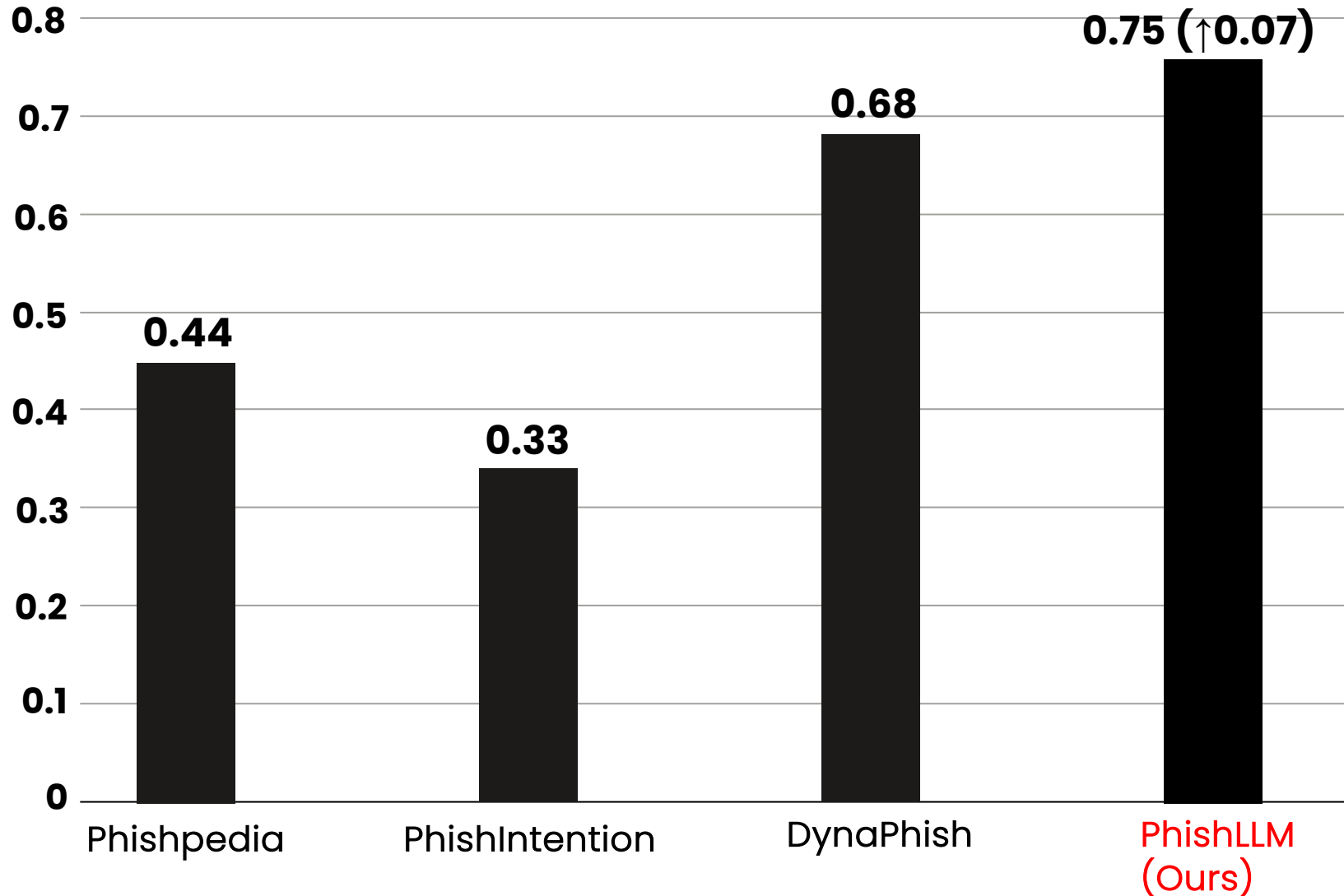
Reference list free
"Look" and "Read"

PhishLLM

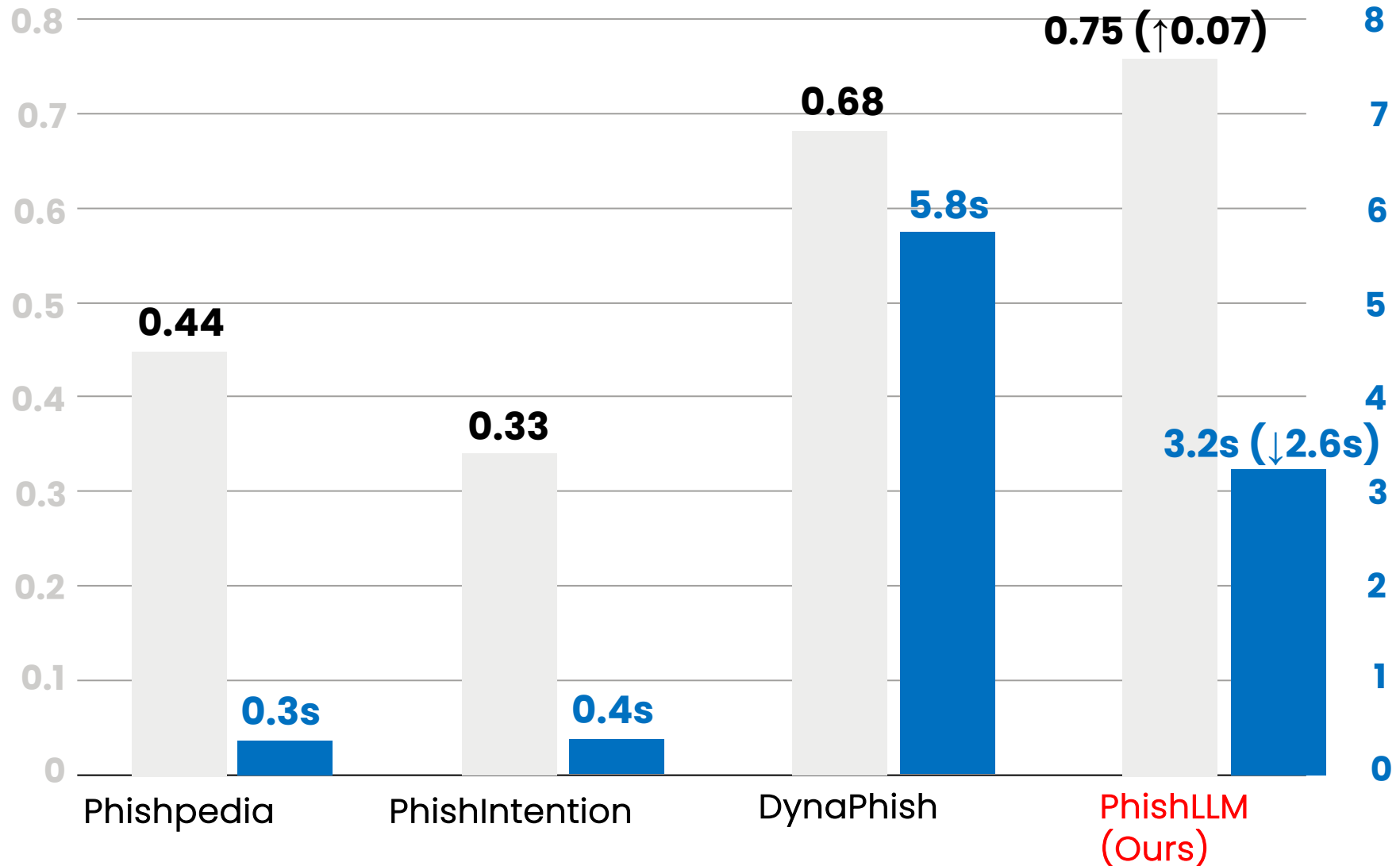
Phishing site: *abc.com*



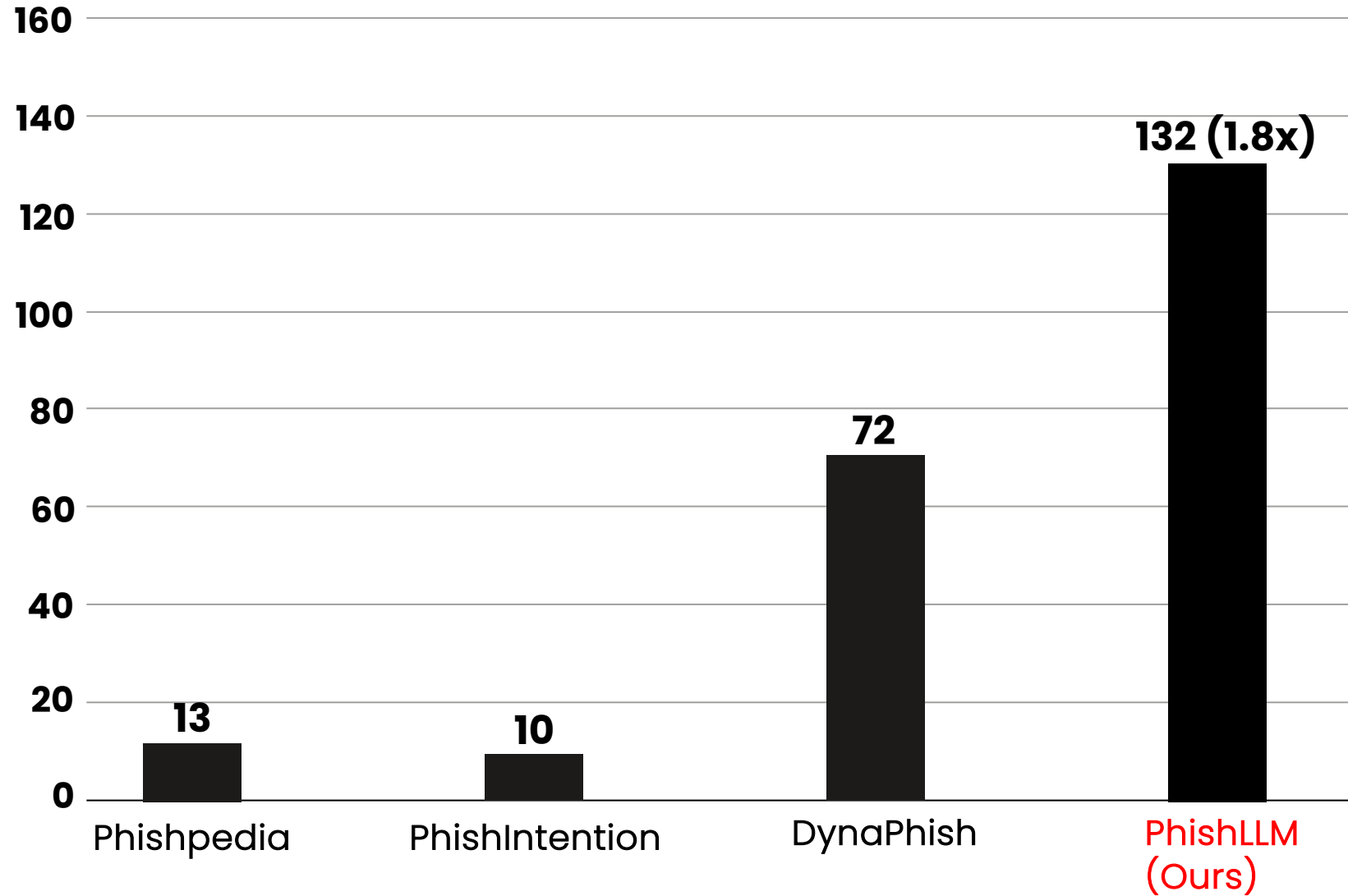
Recall on DynaPD [2] Dataset



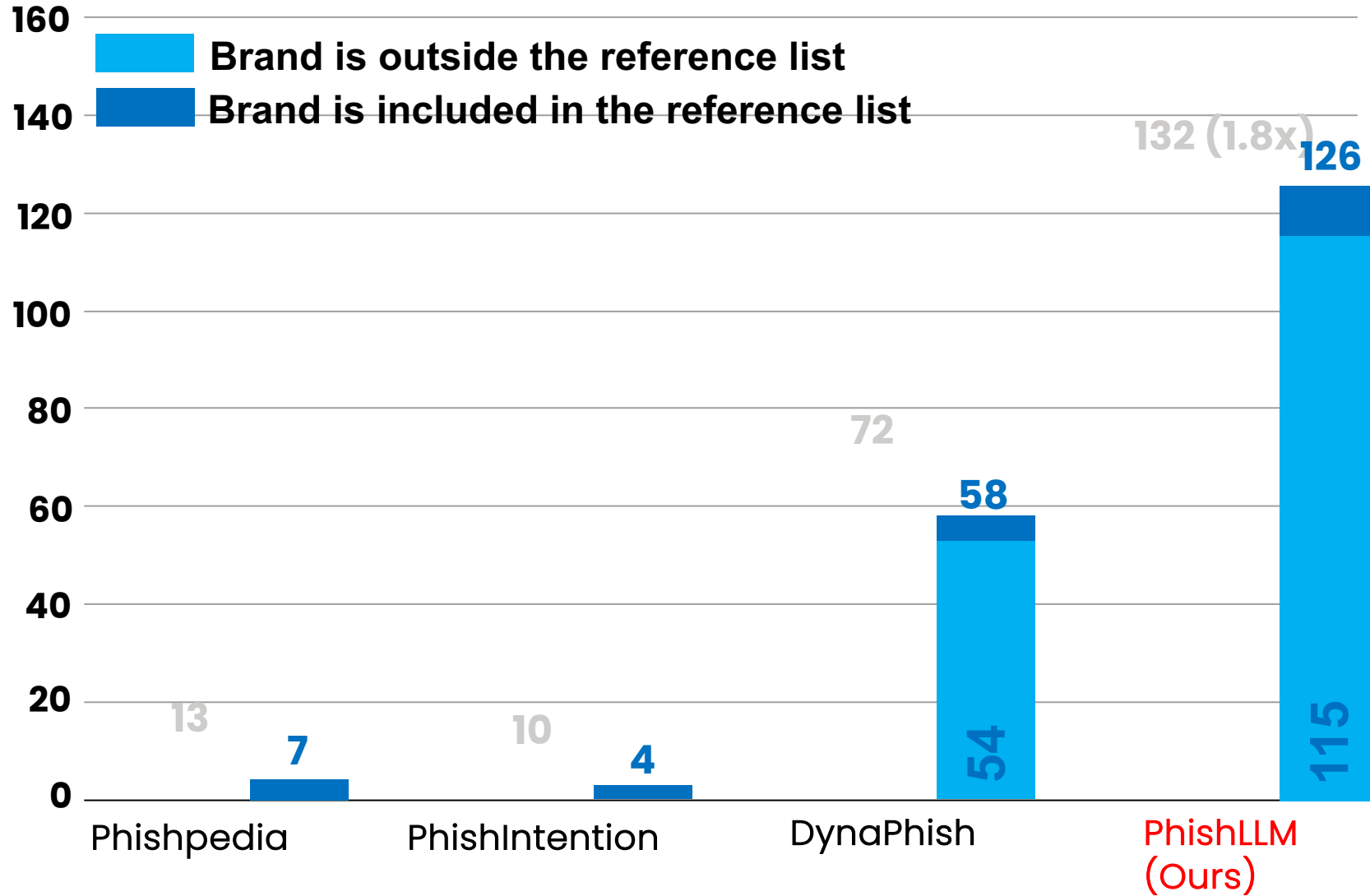
Runtime on DynaPD [2] Dataset



Reported Phishing in the Wild (one-week)



No. Distinct Target Brands Covered by the Reported Phishing



Demo Site: <http://34.204.95.231:6789/>

Adjust the Hyperparameters

Temperature for LLM
(0 for deterministic, 1 for more creative response):

0

Activate brand validation:

off

Number of logo images used for brand validation:

5

Logo matching threshold used for brand validation:

0.7

CRP transition exploration depth limit:

1

Reset

PhishLLM Demo (This demo is for English webpages)

You can select a URL from the dropdown list to test. These URLs are sampled from the OpenPhish feed.

Input the URL here

Submit

Demo Site: <http://34.204.95.231:6789/>

Adjust the Hyperparameters

Temperature for LLM
(0 for deterministic, 1 for more creative response):

0



Activate brand validation:



Number of logo images used
for brand validation:

5



Logo matching threshold used
for brand validation:

0.7



CRP transition exploration
depth limit:

1



Reset

PhishLLM Demo (This demo is for English webpages)

You can select a URL from the dropdown list to test. These URLs are sampled from the OpenPhish feed.

Input the URL here

Submit

Demo Site: <http://34.204.95.231:6789/>

Adjust the Hyperparameters

Temperature for LLM
(0 for deterministic, 1 for more creative response):

0

Activate brand validation:

off

Number of logo images used for brand validation:

5

Logo matching threshold used for brand validation:

0.7

CRP transition exploration depth limit:

1

[Reset](#)

PhishLLM Demo (This demo is for English webpages)

You can select a URL from the dropdown list to test. These URLs are sampled from the OpenPhish feed.

[Submit](#)

Demo Site: <http://34.204.95.231:6789/>

Adjust the Hyperparameters

Temperature for LLM
(0 for deterministic, 1 for more creative response):

0

Activate brand validation:

off

Number of logo images used for brand validation:

5

Logo matching threshold used for brand validation:

0.7

CRP transition exploration depth limit:

1

Reset

PhishLLM Demo (This demo is for English webpages)

You can select a URL from the dropdown list to test. These URLs are sampled from the OpenPhish feed.

<https://ipfs.io/ipfs/bafkreia4pkd2ucwxgsua6iun7q6lih5llarkqgjtfmv2nnrkdhmkpz75m>

⌂ Capturing screenshot

Demo Site: <http://34.204.95.231:6789/>

Adjust the Hyperparameters

Temperature for LLM
(0 for deterministic, 1 for more creative response):

0

Activate brand validation:

Off

Number of logo images used for brand validation:

5

Logo matching threshold used for brand validation:

0.7

CRP transition exploration depth limit:

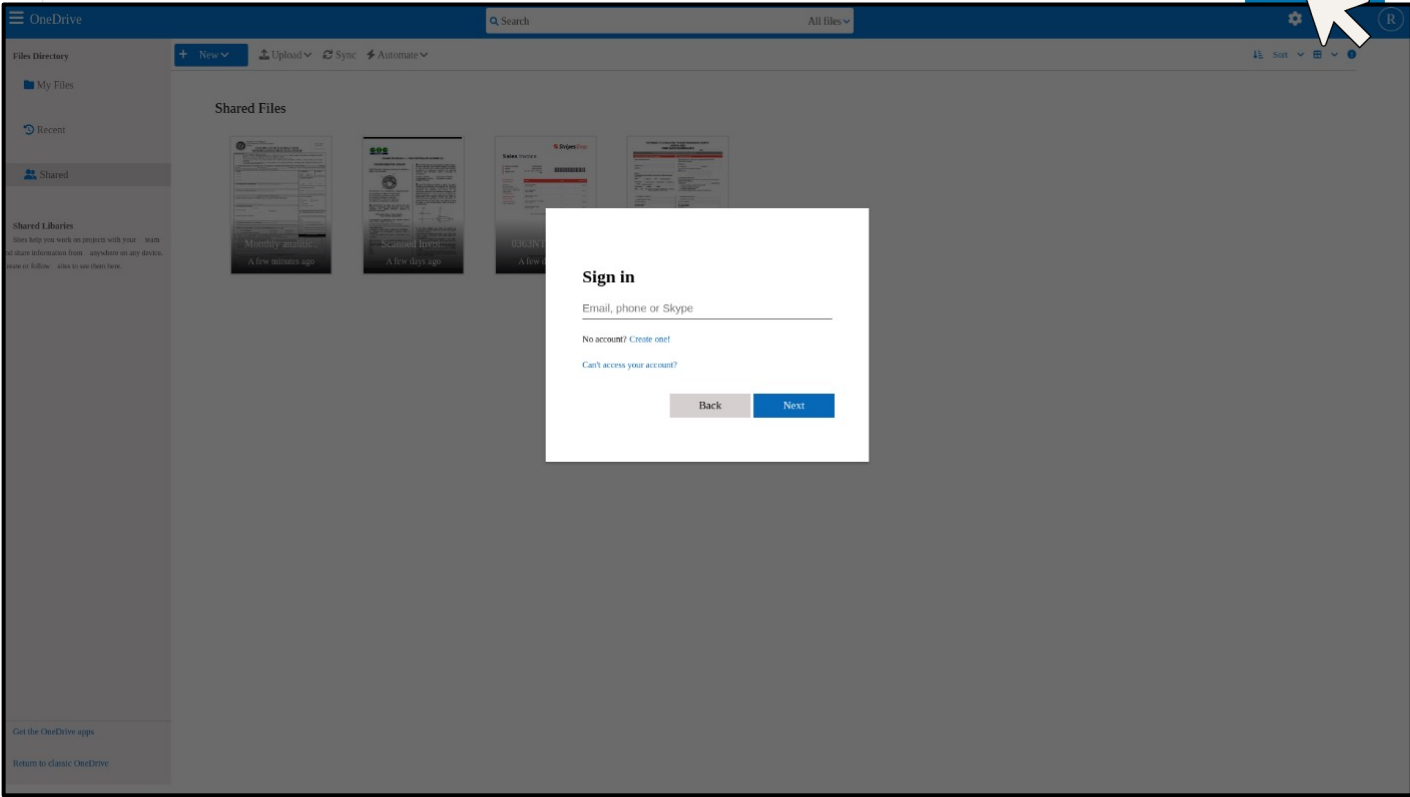
1

[Reset](#)

PhishLLM Demo (This demo is for English webpages)

You can select a URL from the dropdown list to test. These URLs are sampled from the OpenPhish feed.

[Submit](#)



The screenshot shows a OneDrive web interface. A 'Sign in' dialog box is centered on the screen, prompting the user to enter their email, phone, or Skype ID. Below the input field, there are links for 'No account? Create one!' and 'Can't access your account?'. At the bottom of the dialog are 'Back' and 'Next' buttons. The background shows a 'Shared Files' view with several document thumbnails.

[Run PhishLLM](#) [Go Back](#)

Demo Site: <http://34.204.95.231:6789/>

Off

Number of logo images used for brand validation:

5

Logo matching threshold used for brand validation:

0.7

CRP transition exploration depth limit:

1

Reset

Sign in

Email, phone or Skype

No account? [Create one!](#)

[Can't access your account?](#)

Back Next

Run PhishLLM

Go Back

Response
Time taken for initializing webdriver: 3.540012836456299

Response
Time taken for logo detection: 8.043853521347046

Response
Time taken for webpage preprocessing: 17.068352460861206

Generating Logo Description and Webpage Description

Demo Site: <http://34.204.95.231:6789/>

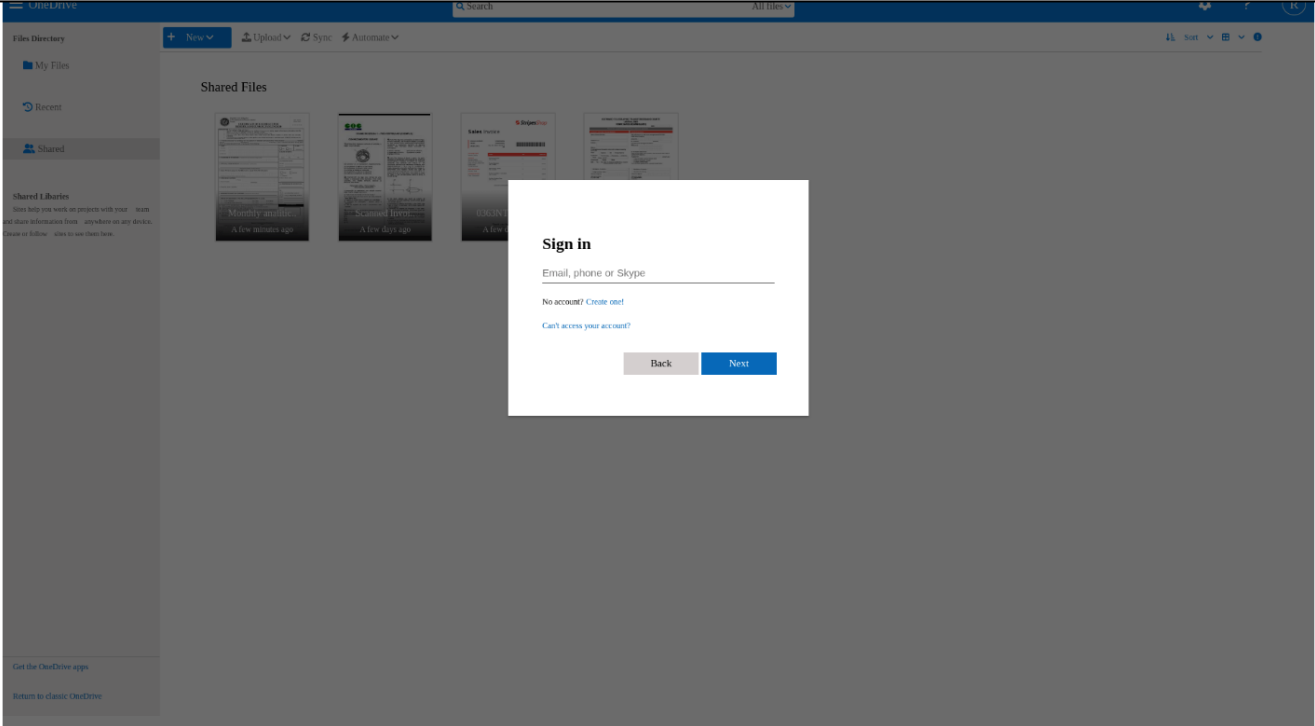
Off

Number of logo images used for brand validation:
5

Logo matching threshold used for brand validation:
0.7

CRP transition exploration depth limit:
1

[Reset](#)



Run PhishLLM

Go Back

Prompt

Given the description on the brand's logo, the logo's OCR text, and the industry sector. Question: What is the brand's domain?

Answer:

Response

Time taken for LLM brand prediction: 0.7974746227264404

Detected brand: **onedrive.live.com**

Brand Recognition

Demo Site: <http://34.204.95.231:6789/>

Off

Number of logo images used for brand validation:

5

Logo matching threshold used for brand validation:

0.7

CRP transition exploration depth limit:

1

Reset

Sign in

Email, phone or Skype

No account? [Create one!](#)

Can't access your account?

Back Next

Run PhishLLM

Go Back

Prompt

Given the HTML webpage text, Question **A. This is a credential-requiring page.** B. This is not a credential-requiring page.

Answer:

Response

Time taken for LLM CRP classification: 1.437190055847168

CRP prediction: There is no confusing token. Then we find the keywords that are related to login: Sign in, Email, phone or Skype. Therefore the answer would be **A.**

CRP Prediction

Demo Site: <http://34.204.95.231:6789/>

Off

Number of logo images used for brand validation:

5

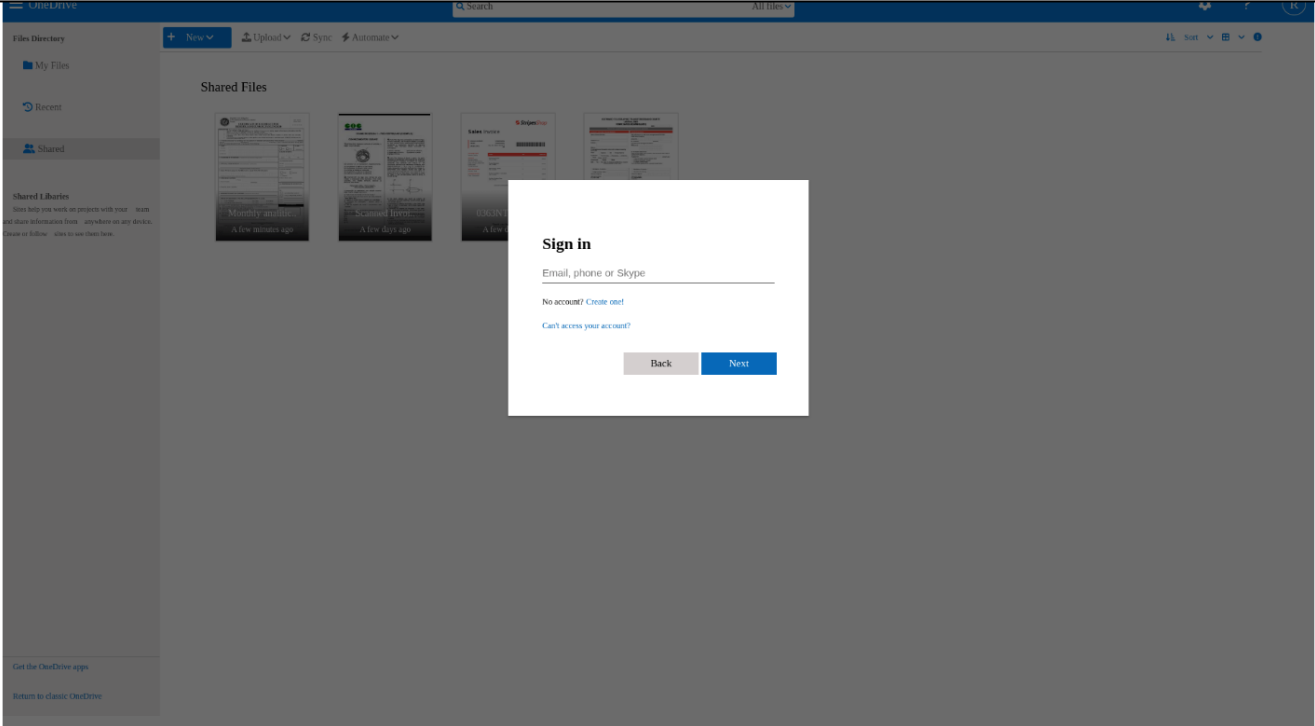
Logo matching threshold used for brand validation:

0.7

CRP transition exploration depth limit:

1

[Reset](#)



The screenshot shows the OneDrive web interface. A 'Sign in' modal is centered on the screen, prompting for an email, phone, or Skype address. Below the input field are links for 'No account? Create one!' and 'Can't access your account?'. At the bottom of the modal are 'Back' and 'Next' buttons. The background shows a 'Shared Files' section with several document thumbnails.

[Run PhishLLM](#) [Go Back](#)

Response

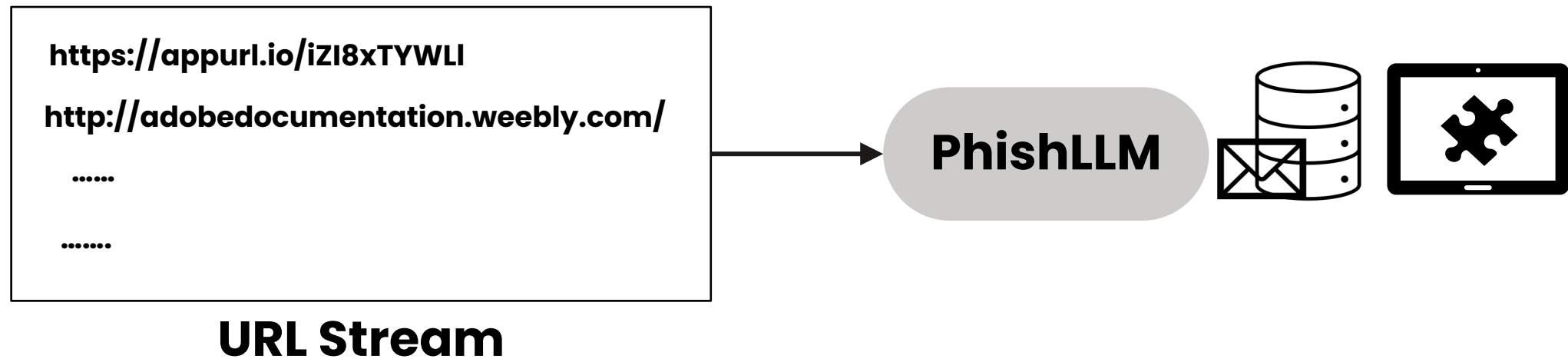
[!] Phishing discovered, phishing target is onedrive.live.com **Final Decision**

PhishLLM has finished running.

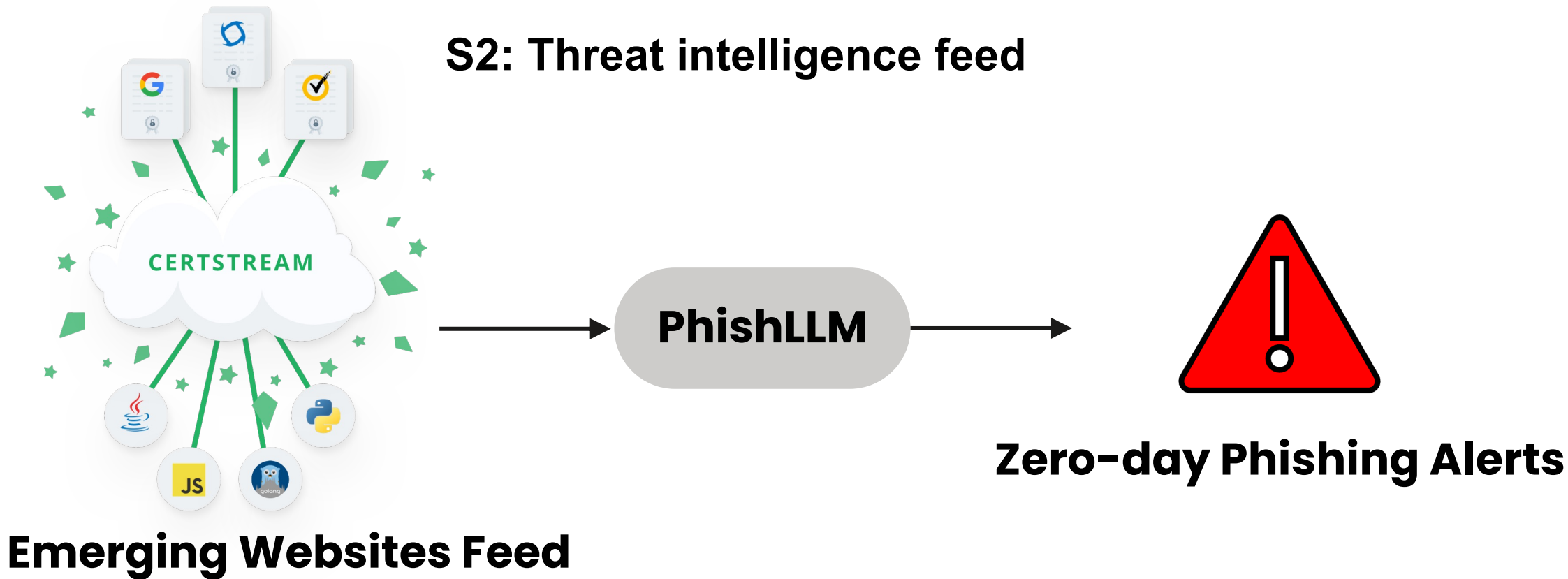
[Go Back](#)

Deployment Scenarios

S1: URL consumption service

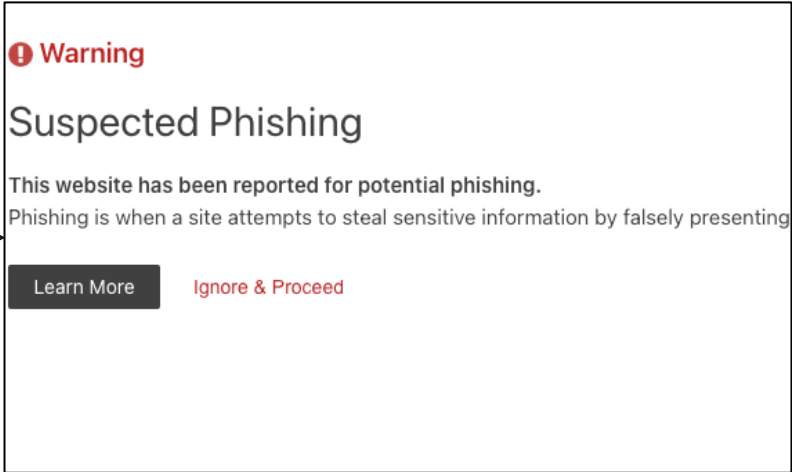


Deployment Scenarios



Deployment Scenarios

S3: Service for Hosting Providers



Web Hosting Providers

PhishLLM

Reference list free

Brand Recognition Module

Logo
Description

Brand
Recognition



Brand: office.com

Brand-domain
Inconsistency

CRP Prediction Module & CRP Transition Module

Webpage
Description

CRP Prediction



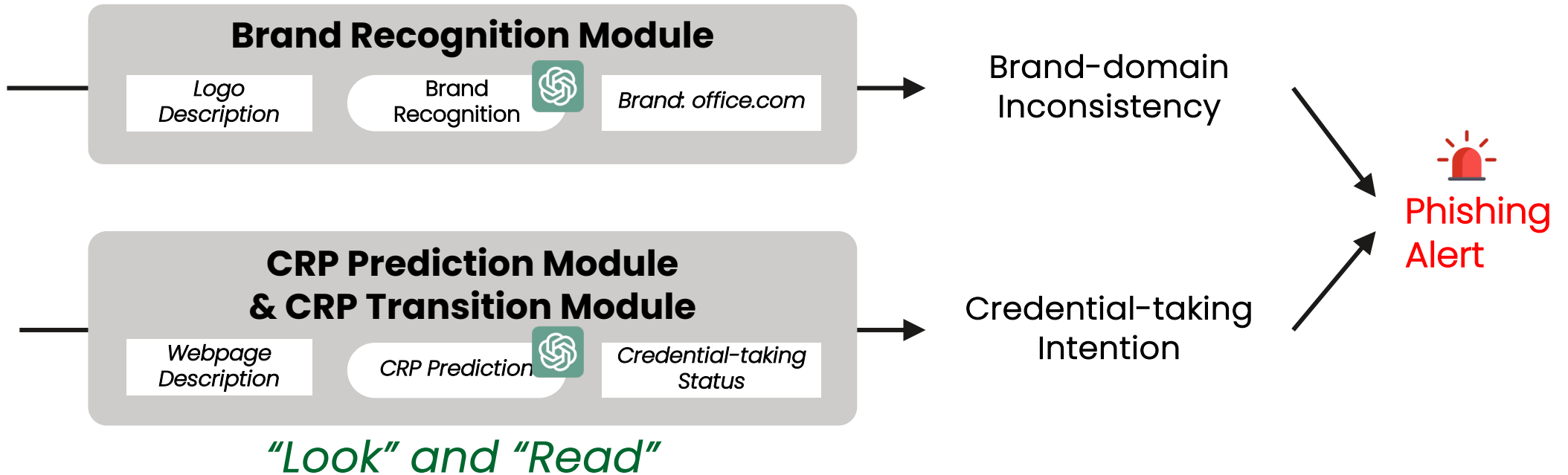
Credential-taking
Status

Credential-taking
Intention

"Look" and "Read"


Phishing
Alert

Phishing site



Interested to Know More?



Code: <https://github.com/code-philia/PhishLLM/>



Website: <https://sites.google.com/view/phishllm/home>



Email: liu.ruofan16@u.nus.edu

