PhishLLM

Reference-based Phishing Detection

- Goal of phishing: have a website that *looks like* a target
- Reference: a map of “look” <-> URL
- Detect impostors and look for inconsistencies
- Problem: where do we get references?

Use LLM for detection

- Identify logo (previous work)
- Process it
  - Image caption
  - OCR
- Send to LLM and ask to come up with domain name
  - Use “minimum-entropy” answer

Validation

- Check whether
    - Logo matches google search “[x.com] logo”
    - Site is indexed
    - Other checks?
- Do these create false positives or negatives on their own?

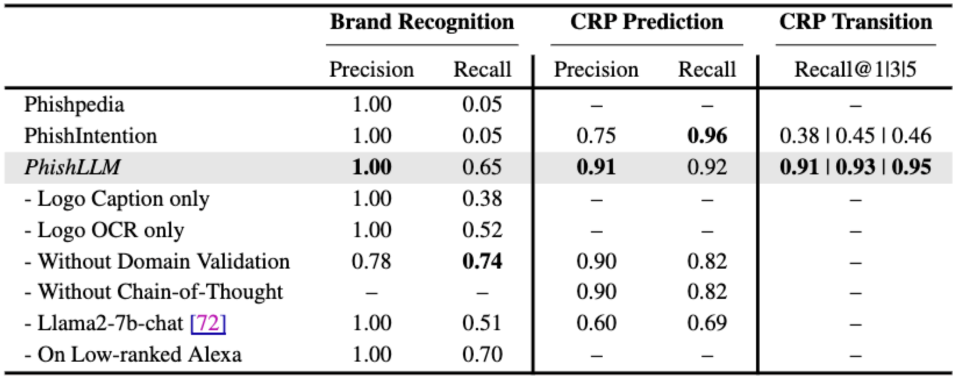
Credential Request Page Prediction

- Text from screenshot
- LLM with chain-of-reasoning
- Click on UI elements likely to lead to credential requests

Evaluation

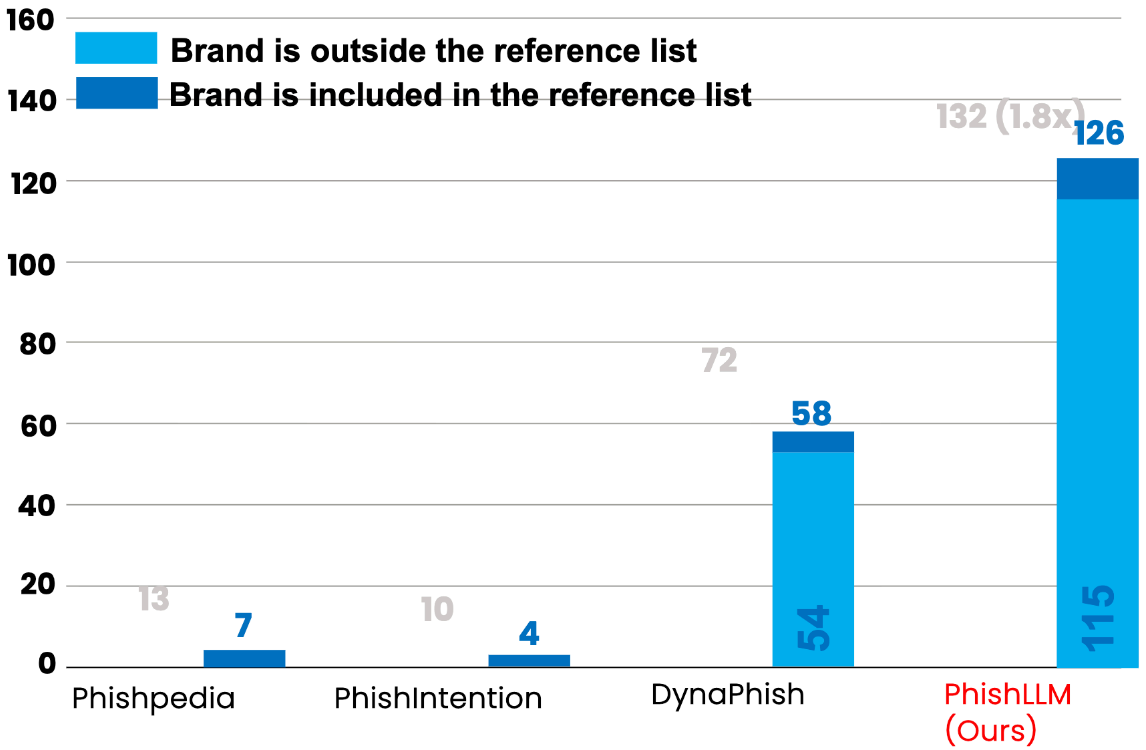
- Data sets
  - Known-phising data sets
  - Known-phishing clean set
  - Certistream set, with expert labeling

Table 4: Component-wise Performance Evaluation.



	Brand Recognition		CRP Prediction		CRP Transition
	Precision	Recall	Precision	Recall	Recall@1|3|5
Phishpedia	1.00	0.05	–	–	–
PhishIntention	1.00	0.05	0.75	**0.96**	0.38 | 0.45 | 0.46
PhishLLM	**1.00**	0.65	**0.91**	0.92	**0.91** | **0.93** | **0.95**
- Logo Caption only	1.00	0.38	–	–	–
- Logo OCR only	1.00	0.52	–	–	–
- Without Domain Validation	0.78	**0.74**	0.90	0.82	–
- Without Chain-of-Thought	–	–	0.90	0.82	–
- Llama2-7b-chat [72]	1.00	0.51	0.60	0.69	–
- On Low-ranked Alexa	1.00	0.70	–	–	–

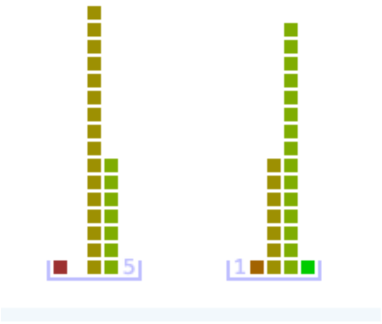
Results


160
140
120
100
80
60
40
20
0
Brand is outside the reference list
Brand is included in the reference list
13
7
10
4
72
58
54
132 (1.8x)
126
115
Phishpedia
PhishIntention
DynaPhish
PhishLLM (Ours)

Vibes

- Novel idea of using LLMs
- Strong empirical results, zero-day detection
- Validation, rather than blind trust in LLM results

- Dependence on outside services
- Low adversarial robustness
- Brittle pipeline
- Limitations



Robustness to Evasion

- Some rudimentary evaluation of prompt injection
  - May have been ok in 2023, outdated now
- Cloaking / other deception
- Non-login pages
- Indexed domain compromise

Practicality

- Are costs reasonable?
- What modality makes sense?
  - Browser plugin
  - Proactive monitoring (eg CT logs)
  - URL monitoring (spam filters)
  - Hosting providers

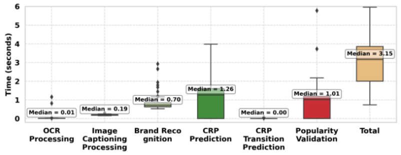
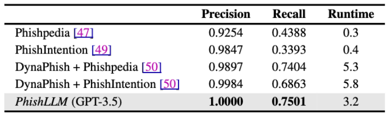


Figure 5: *PhishLLM* Runtime Breakdown

Table 3: Overall Performance



	Precision	**Recall**	**Runtime**
Phishpedia [47]	0.9254	0.4388	0.3
PhishIntention [49]	0.9847	0.3393	0.4
DynaPhish + Phishpedia [50]	0.9897	0.7404	5.3
DynaPhish + PhishIntention [50]	0.9984	0.6863	5.8
PhishLLM (GPT-3.5)	**1.0000**	**0.7501**	3.2

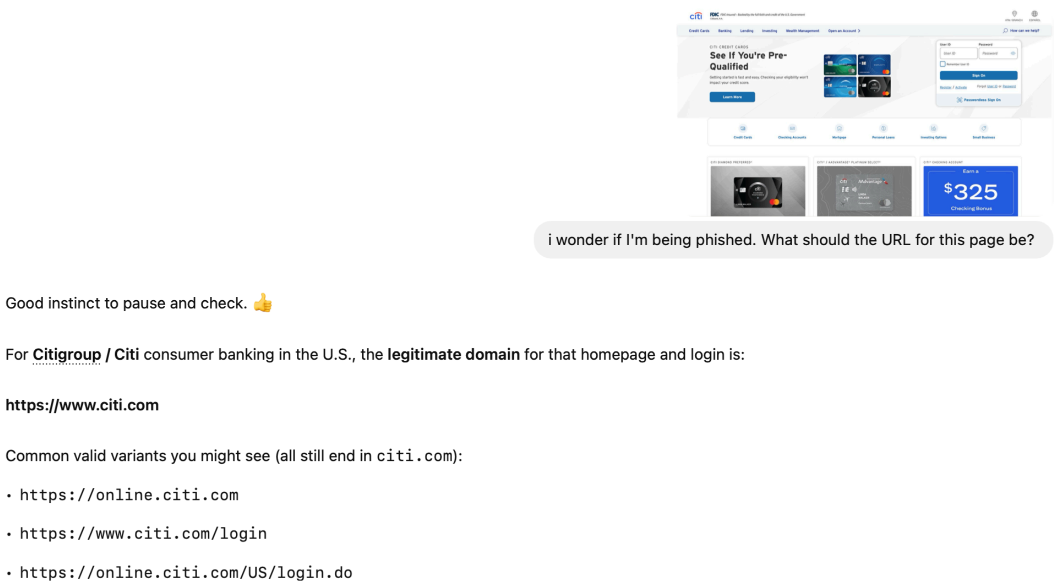
Alternative Evaluation

- What was missing? What should have been added?

Alternative Approaches

- How would we implement some of these ideas better?

How much better would current LLMs be?



Getting started is fast and easy. Checking your eligibility won't impact your credit score.

i wonder if I'm being phished. What should the URL for this page be?

Good instinct to pause and check. 👍

For **Citigroup / Citi** consumer banking in the U.S., the **legitimate domain** for that homepage and login is:

https://www.citi.com

Common valid variants you might see (all still end in `citi.com`):

- `https://online.citi.com`
- `https://www.citi.com/login`
- `https://online.citi.com/US/login.do`

Failure Mode

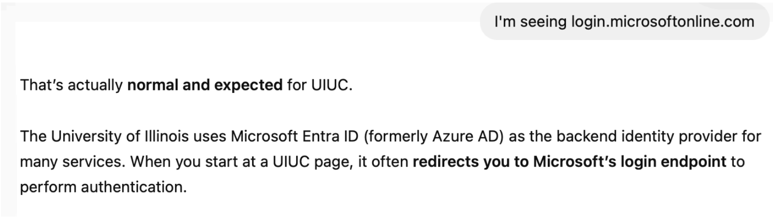
For **University of Illinois Urbana-Champaign** single sign-on, the legitimate login page should be on one of these domains:

✅ **Official UIUC Login URLs**

- **https://identity.uillinois.edu/**
- **https://login.uillinois.edu/**
- You may also see a redirect to something like: https://identity.uillinois.edu/idp/profile/SAML2/Redirect/SSO

All of these should:

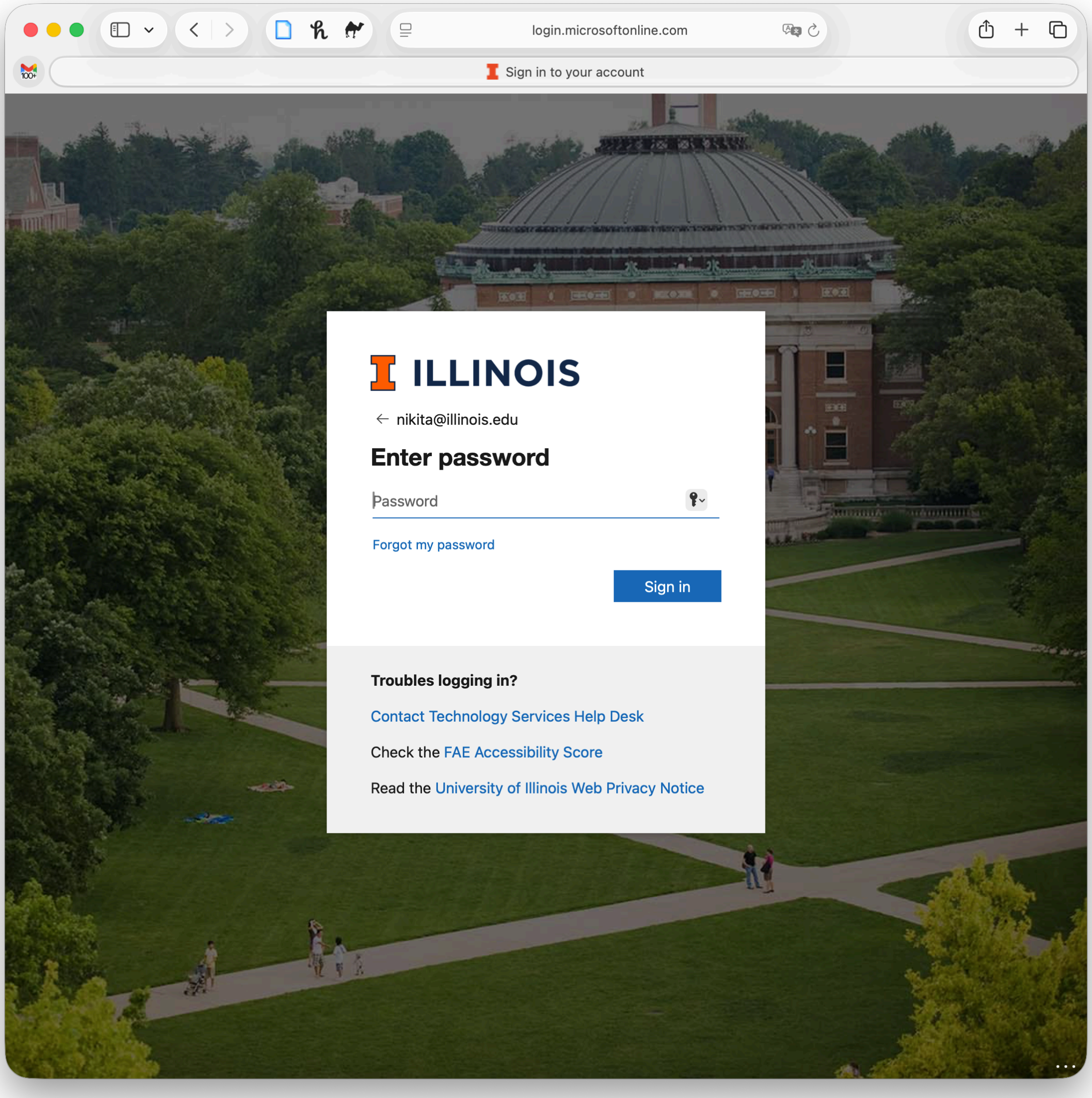
- Use **HTTPS**
- End in **uillinois.edu** (not illinois-login.com, not uiuc-login.net, etc.)
- Show a valid lock icon with a certificate issued to the University of Illinois
-



I'm seeing login.microsoftonline.com

That's actually **normal and expected** for UIUC.

The University of Illinois uses Microsoft Entra ID (formerly Azure AD) as the backend identity provider for many services. When you start at a UIUC page, it often **redirects you to Microsoft's login endpoint** to perform authentication.



login.microsoftonline.com

Sign in to your account

ILLINOIS

← nikita@illinois.edu

Enter password

Password

Forgot my password

Sign in

Troubles logging in?

Contact Technology Services Help Desk

Check the FAE Accessibility Score

Read the University of Illinois Web Privacy Notice