



**Algorithmic and language-based analysis for the
detection of phishing emails**

**Development of a hybrid
evaluation system**

Jannik Hurst

Forschungsfrag

e

Wie lässt sich ein nutzerzentriertes Phishing-Erkennungssystem technisch umsetzen, das verständliche Erklärungen liefert und wie wird es von den Nutzern wahrgenommen?



How can a user-centered phishing detection system be technically implemented to deliver clear, explanatory feedback and how is the resulting prototype perceived by users?



Related Work

PhishURLDetect

Erkennung von Phishing-URLs, die auf betrügerische Websites führen

Utilizing LLMs with Human Feedback Integration

Statt nur zu warnen, bezieht das System den Nutzer aktiv ein, z. B. mit Pop-up-Fragen wie „Kennst du den Absender?“.

Meta GPT-Based Agent

LLM-basierter Agent analysiert E-Mail-Header und -Body getrennt und trifft strukturierte Entscheidungen mit Begründung



Umfassende Überprüfung

→ Kombination aus technischer Analyse und sprachlicher Bewertung

Erklärung

→ Verstehen, warum eine Mail als gefährlich gilt

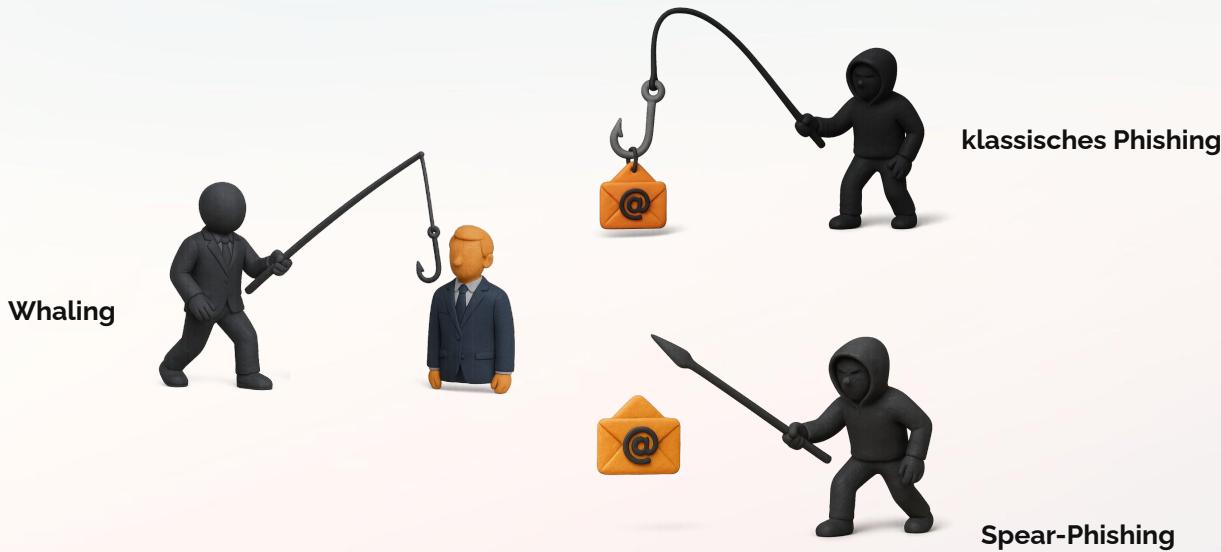
⇒ **Stärken statt nur schützen**



Phishing?

Die Nutzung von **Täuschung**, um eine Person dazu zu bringen, private **Informationen preiszugeben** oder unbeabsichtigt **unautorisierten Zugriff** auf ein Computersystem oder Netzwerk zu ermöglichen.

Arten von Phishing



Beispiel

Einsatz von Dringlichkeit

Marken-Imitation

Gefälschte Links



20. Dezember 2024 um 01:31

Schauen Sie sich diese Mail in Ihrem Browser an



Lieber Kunde jannik.hurst,

WIR KÖNNEN IHR PAKET DERZEIT NICHT LIEFEREN
5841233520000 von unserem Lager an Ihre Adresse.
Aufgrund fehlender Informationen in unserem System,
Bitte beheben Sie dieses Problem innerhalb von (5)
Werktagen, andernfalls müssen wir das Paket an den
Hersteller zurücksenden.

Aufbau

.eml Datei

```
1 Delivered-To: jannik.hurst@gmail.com
2 Received: by 2002:ab2:3904:0:b0:21a:6760:feb8 with SMTP id 14csp1500714lqc;
3 | | | Thu, 19 Dec 2024 16:31:48 -0800 (PST)
4 X-Google-Smtp-Source: AGHT+IGrcp0gqNhr67FbJo+XI5YNjr36HvQt5BJczhRz1FIVvw0kEaw3ql/I
5 X-Received: by 2002:a05:6000:1866:b0:38a:20d9:32e6 with SMTP id ffacd0b85a97d-38a2
6 | | | Thu, 19 Dec 2024 16:31:48 -0800 (PST)
7 ARC-Authentication-Results: i=1; mx.google.com;
8 | | | spf=pass (google.com: domain of return8318@83185.135.69.45us27y5qa84cinlr5.
9 Return-Path: <return8318@83185.135.69.45us27y5qa84cinlr5.135.69.45f8ma4ogilcud9v25
10 Received: from hideoutpoa.com (ip45.ip-5-135-69.eu. [5.135.69.45])
11 | | | by mx.google.com with ESMTPS id 5b1f17b1804b1-4366129b3a1si16168775e9.193.
12 | | | for <jannik.hurst@gmail.com>
13 | | | (version=TLS1 cipher=ECDHE-ECDSA-AES128-SHA bits=128/128);
14 | | | Thu, 19 Dec 2024 16:31:48 -0800 (PST)
15 Received-SPF: pass (google.com: domain of return8318@83185.135.69.45us27y5qa84cinl
16 Authentication-Results: mx.google.com;
17 | | | spf=pass (google.com: domain of return8318@83185.135.69.45us27y5qa84cinlr5.
18 From:=?UTF-8?B?RF8Ewq4=?=<YLPiUIOE@YLPiUIOE.us>
19 To: jannik.hurst@gmail.com
20 Message-ID: <80jt89nvm93tj9tUEZ-80jt89nvm93tj9tUEZ@80jt89nvm93tj9tUEZ.com>
21 Subject: jannik.hurst, DEIN PAKET IST ANGEKOMMEN 📦
22 Content-type: text/html
23 Date: _smtpDate . 278757547
24
25
26
```

Header

Metadaten wie Absender, Server, Authentifizierung

Body

Inhalt der Nachricht, meist als Text oder HTML

Anhang

Dateien, die der Mail beigefügt sind (z. B. PDF, Word, ZIP)

Einsendeweg

e



Web-Portal

<https://mailcheck.help>



Email

start@mailcheck.help

Header-Analyse



SPF, DKIM, DMARC ●

Prüfen, ob der Absender authentisch ist und die E-Mail unterwegs nicht manipuliert wurde.

Metadaten ●

z.B. Weicht der Return-Path von dem Absender ab?

Blacklists ●

Die IP- und Domainadressen werden mit bekannten Spam- und Phishing-Listen abgeglichen

⇒ Nicht möglich wenn .eml Datei nicht vorliegt

```
1 Delivered-To: jannik.hurst@gmail.com
2 Received: by 2002:ab2:3904:0:b0:21a:6760:feb8 with SMTP id l4csp594302lqc;
3     Wed, 18 Dec 2024 04:38:45 -0800 (PST)
4 Return-Path: <20241218123844370eec9e6c9c47eb8127c2d09530p0eu-C2DT0EHK4KR7EQJ
5 Received-SPF: pass (google.com: domain of 20241218123844370eec9e6c9c47eb8127c2d09530p0eu-C2DT0EHK4KR7EQJ
6 Authentication-Results: mx.google.com;
7     dkim=pass header.i=@amazon.de header.s=llktbq2gwxn3x3xrq5ljspgjk2nc5ap
8     dkim=pass header.i=@amazonses.com header.s=uku4taia5b5tsbglxyj6zym32efj7xqv;
9     spf=pass (google.com: domain of 20241218123844370eec9e6c9c47eb8127c2d09530p0eu-C2DT0EHK4KR7EQJ
10    dmarc=pass (p=QUARANTINE sp=QUARANTINE dis=NONE) header.from=amazon.de
11 DKIM-Signature: v=1; a=rsa-sha256; q=dns/txt; c=relaxed/simple;
12     s=uku4taia5b5tsbglxyj6zym32efj7xqv; d=amazonses.com; t=1734525524;
13     h=Date:From:To:Message-ID:Subject:MIME-Version:Content-Type:Feedback-ID;
14     bh=SiJULaOH/5QE3fxBqQ9skY04r9T13Lv++yGuFqL2Epw=;
15     b=TEuA4hGlcLoJj+iup/y05AMfQz60yjaNb+gu1iJKjvPHaKJtLGPLeFuahz6U9No
16     XLdacJfV0eZ0HSoH9gytovmtrqUP+KnZWrbtXaRf34StBJdn/4R1FtQdR9xW4MBQvSh
17     pCEf6F2UcipnxBpM6ITFN6gcC2MiSj8r7nwKV4FM=
18 Date: Wed, 18 Dec 2024 12:38:44 +0000
19 From: "Amazon.de" <order-update@amazon.de>
20 To: jannik.hurst@gmail.com
21 Message-ID: <01020193d9c71aca-23798611-f081-4bb1-96ee-2ffcf618a69d-000000000000>
22 Subject: Zugestellt: deine Amazon.de-Bestellung mit Bestellnr.
23 | 306-0373188-8804310
24 MIME-Version: 1.0
25 Content-Type: multipart/alternative;
26     boundary="-----_Part_615463_517225660.1734525524673"
27 X-AMAZON-MAIL-RELAY-TYPE: notification
28 Bounces-to: 20241218123844370eec9e6c9c47eb8127c2d09530p0eu-C2DT0EHK4KR7EQJ
```

Anhang-Analyse

Verschachtelte E-Mails

.eml-Dateien in Anhängen werden wie separate E-Mails behandelt und vollständig analysiert

Malware-Scan

Alle Anhänge werden auf Schadsoftware geprüft



Body-Analyse

URL-Erkennung

Abgleichen mit White-/Blacklists

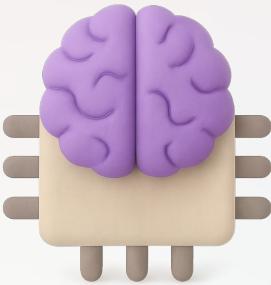
Browser-Emulation

Öffnen von unbekannten URLs, folgen von Weiterleitungen, Screenshot von Zielseite

Überprüfen der Zielseite

Erneute URL-Erkennung, speichern der Meta-Beschreibung





LLM-Auswertung

Eingabe

Analyse-Ergebnisse

Header-, Body- und Attachment-Analyse fließen in Prompt ein

Bereinigter Body

Styling und Leerzeichen werden entfernt

Erklärseiten

Verlinkbare Begriffe für kontextuelle Hinweise

[...]

Hier sind die Ergebnisse der Header-Analyse:

`{header_analysis}`

90% der Spam und Phishing Mails können bereits anhand der Header erkannt werden, da sie versuchen durch einen Firmennamen oder eine bekannte Domain zu täuschen.

ABER: Es gibt auch legitime Mails die DKIM und SPF noch nicht implementiert haben, jedoch sind dies meist kleinere Firmen oder Privatpersonen.

Große Firmen mit viel Traffic haben dies meist schon implementiert.

Hier ist der Inhalt der Mail ohne style tags:

`{email_body}`

[...]



LLM-Auswertung

Ausgabe

Antwort im JSON-Format

Score, Gefahrenliste & Erklärungen

Format je nach Datenlage

Anderes Format bei fehlender Header-Analyse

```
{  
    "title": "Verdächtiger Absender",  
    "trustworthiness": "suspicious",  
    "trustPoints": 2,  
    "threats": [  
        {  
            "title": "Absender Name",  
            "severity": "medium",  
            "description": "Der Absender-Name passt nicht zur E-Mail"  
        },  
        {  
            "title": "Link-Ziel",  
            "severity": "high",  
            "description": "Ein Link führt auf eine als gefährlich erachtete Seite"  
        }  
    "explanation": "Die E-Mail enthält mehrere verdächtige Merkmale, die sich von der ursprünglichen Domain unterscheiden. Der Absendername ist ungewöhnlich und die verlinkte Seite ist möglicherweise Phishing.",  
    "senderName": "Kundenservice Amazon",  
    "senderAddress": "support@amazn-fake.com"  
}
```

Analyse-Ergebnis

Analyseergebnis

Ihre E-Mail Sicherheitsanalyse

Absenderinformationen

Name: OTTO Payments

E-Mail: noreply@info.otto-payments.de

Betreff: 2. Mahnung zu deinem Kauf bei OTTO zur Rechnung vom 11.02.2025

Gesamtpunktzahl

4 von 10

(Inklusive Funktionalität und Design)

Rechnungserinnerung von OTTO Payments – Auffälligkeiten bei der Echtheit

SPP-Prüfung fehlgeschlagen

Die SPP-Prüfung ist fehlgeschlagen. Das bedeutet, dass der Mailserver, von dem die E-Mail gesendet wurde, nicht autorisiert ist, im Namen der Domain info.otto-payments.de E-Mails zu versenden. Dies könnte ein Zeichen für Spoofing sein.

Unklare Meta-Description bei Links

Einige Links in der E-Mail liefern keine sicheren Inhalte.

Datum in der Zukunft

Das Rechnungsdatum liegt in der Zukunft.

Diese E-Mail weist einige Verläufe auf, die als fehlgeschlagene Domäne info.otto-payments.de angezeigt werden. Der Prüfer hat persönliche Informationen angepasst, um die Sicherheit Ihres Browsers aufzuhöhen und die Kundenservice separat, um die

Header Analyse

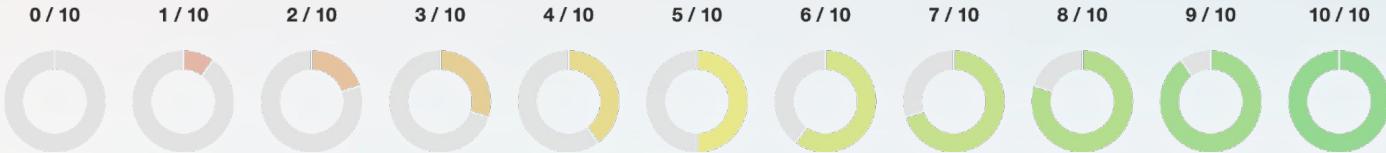
Header und Header-Informationen (Absender) überprüft werden, ob die E-Mail manipuliert wurde.

Prüfung der Absender-Domain
info@payments.de wurde auf null Blacklisten gefunden.

Prüfung des Absender-Mailervers
c184-168-smtp-out.eu-central-1.amazonaws.com wurde auf keiner Blacklist gefunden.

Im Footer der Analyse der Header kann

Sicherheitsbericht



Verdächtige E-Mail im DPD-Design

Absender-E-Mail

Die Absender-E-Mail-Adresse YLPiUIOE@YLPiUIOE.us passt nicht zum Absender-Namen DPD® und deutet auf einen möglichen Betrugsversuch hin. Die Domain YLPiUIOE.us ist unbekannt und nicht mit DPD assoziiert.

Link-Ziele

Die Links in der E-Mail führen zu storage.googleapis.com, was zwar nicht per se schädlich ist, aber in diesem Kontext verdächtig ist, da DPD normalerweise auf die eigene Domain verlinken würde. Die Zielseite hat keine Meta-Description.

Sprache und Inhalt

Die E-Mail enthält eine dringliche Aufforderung, fehlende Informationen zu korrigieren, um die Zustellung eines Pakets zu gewährleisten. Dies ist eine typische Taktik bei Phishing-Versuchen, um Empfänger unter Druck zu setzen. Die Sprache ist fehlerhaft und unprofessionell.

Die E-Mail ist höchstwahrscheinlich ein Phishing-Versuch. Mehrere Indikatoren deuten darauf hin, dass die E-Mail nicht von DPD stammt. Der Absendername stimmt nicht mit der [Domain](#) der E-Mail-Adresse überein, [SPF](#) und [DKIM](#) Validierung sind fehlgeschlagen und die [Weiterleitung](#) führt zu einer fremden Domain. Die Kombination dieser Faktoren deutet stark darauf hin, dass es sich um einen Betrugsversuch handelt. Klicken Sie auf keine Links und geben Sie keine persönlichen Daten preis. Es ist ratsam, die E-Mail unbeantwortet in den Spam-Ordner zu verschieben und gegebenenfalls DPD über den Vorfall zu informieren. Achten Sie zukünftig verstärkt auf solche Merkmale, um Phishing-Versuche frühzeitig zu erkennen. (+1 Punkte)



https://mailcheck.help/explain/domain



Domain

Die Domain ist der Teil einer E-Mail nach dem @-Zeichen. Sie gibt Auskunft darüber, von welchem Server die E-Mail versendet wurde. Sie kann auf eine Blacklist gesetzt sein, wenn sie für Spam oder Phishing missbraucht wurde. Es ist daher wichtig, die Domain zu überprüfen, um die Authentizität der E-Mail zu gewährleisten.

Nutzerstudie



Rein qualitativ

Wie verständlich sind die Berichte?

Lerneffekt bei künftigen Phishing-Mails?

 Jetzt Feedback geben



Nutzerstudie

Aufbau und Ablauf

Einreichung über Webportal oder Mail

Feedback direkt unter dem Bericht möglich

15 Antworten von 11 Teilnehmern ($\bar{\varnothing} 31.73$ J.)

Offene Textfelder und Skalenfragen (0-10)

Nutzerstudie

Ergebnisse

Kriterium	Bewertung, Ø
Erklärungen	6.8/10
Darstellung	7.6/10
Unterstützung bei Risikoeinschätzung	5.8/10
Vertrauen in den Bericht	9/10

Future Work

Erklärungen an Nutzerwissen anpassen

Interaktiv statt statisch

Größere Studie zur Wirksamkeit





start@mailcheck.help



mailcheck.help

