RESEARCH DOCUMENT

Stephen Foley
C00282842
Supervisor:
Hisain Elshaafi

Table of Contents

1. Abstract

Phishing remains one of the most common and successful initial access vectors used by attackers, with organisations facing increasing volumes of email, SMS, QR-code and voicebased phishing campaigns that often bypass traditional technical controls. This research examines modern phishing techniques and related social engineering methods, with a particular focus on how psychological factors, such as authority, urgency, and social proof, influence user susceptibility to these attacks. The document analyses existing phishing simulation and awareness tools to identify weaknesses, including high cost for SMEs, limited support for emerging vectors like quishing and vishing, poor user engagement, and inadequate reporting and training integration.

To address these gaps, the project designs and implements a phishing simulation platform aimed at small organisations, providing configurable email QR-code-based campaigns, user and group management, interaction tracking, and basic reporting to highlight high-risk users who require additional training. The system is built using Python, Django, Django REST Framework, Celery, Redis and PostgreSQL, with security and privacy controls such as HTTPS-only communication, role-based access control, and anonymised storage of interaction data. Ethical considerations, including informed consent, GDPR-compliant data handling, and non-punitive feedback, guide both the system design and the planned lab evaluation with consenting participants, which will be used to assess the effectiveness of the prototype in improving user awareness of modern phishing attacks.

2. Introduction

Phishing is one of the most successful attack techniques employed by threat actors, and according to a 2024 report by Huntress, 90% of cyberattacks begin with phishing. An estimated 3.4 billion phishing emails are sent daily, and in 2022, 48% of all emails sent were phishing emails [Huntress, 2024. Threat actors are constantly improving their methods, every day makes falling for phishing much easier. According to IBM's Cost of a Data Breach report, the average cost of a phishing attack is $4.48 million USD [IBM, 2025]. DeepStrike says that 54% of all ransomware attacks begin with phishing as the initial access vector [DeepStrike, 2025]. Threat actors are combining multiple phishing methods, such as vishing/deepfakes, quishing and smishing, to make their attacks seem much more convincing to their victims. A rise in vishing and deepfake attack vectors can be seen due to the rise of Generative AI use.

Phishing is no longer just limited to email inboxes. Threat actors are targeting common corporate collaboration platforms, including Microsoft Teams and Slack. They are also evolving their methods to use SMS, QR Codes and phone calls to deliver their malicious payload and trick their victims into entering credentials. Due to the increasing sophistication of attack vectors, organisations can no longer rely solely on technical controls such as firewalls and email filters. They also need to ensure that all users are appropriately trained and up to date on the most current attack vectors. Organisations often attempt to train their users through awareness training and phishing simulations, but many of the existing tools that do this are expensive, use outdated attack vectors and offer a poor user experience overall.

2.1 Objectives

To achieve this project's aim, my project has the following objectives:

I. Conduct research on Phishing, Modern Phishing Attack Vectors, Social Engineering and Ethical Concerns.
II. Find weaknesses in existing phishing tools.
III. Design system architecture for a phishing simulation platform for small organisations.
IV. Implement a basic prototype that supports:
    - Email phishing simulations.
    - Creating users and groups.
    - Generate Reports.
V. Improve on existing features in the prototype and work out any kinks. VI. Add more features and test the prototype again.
VII. Find test groups and obtain their consent to participate in a phishing simulation.

2.2 Scope and Limitations

The scope of this project is limited to the design and implementation of a functioning phishing simulation tool. The project prototype focuses on core functionality rather than design for large enterprise scalability, integration with corporate email systems and advanced analytics.

The real-world deployment of this tool is outside the current scope due to ethical, legal, and time constraints. So I will be carrying out and testing the tool in a lab environment with a small group of consenting participants.

3. Phishing

3.1 What is phishing

Phishing refers to threat actors attempting to trick users into divulging sensitive credentials such as usernames, passwords, credentials and credit/debit card information, for financial gain by selling the stolen information or convincing victims to transfer money [Cloudflare, 2024]. The term phishing is a play on words of actual fishing, where threat actors try to get users to bite their bait.

Classic phishing attacks typically involve emails that attempt to impersonate legitimate trusted organisations such as Google, Microsoft, Banks, PayPal, IT internally in a company, etc.

3.2 Modern techniques of phishing

Phishing

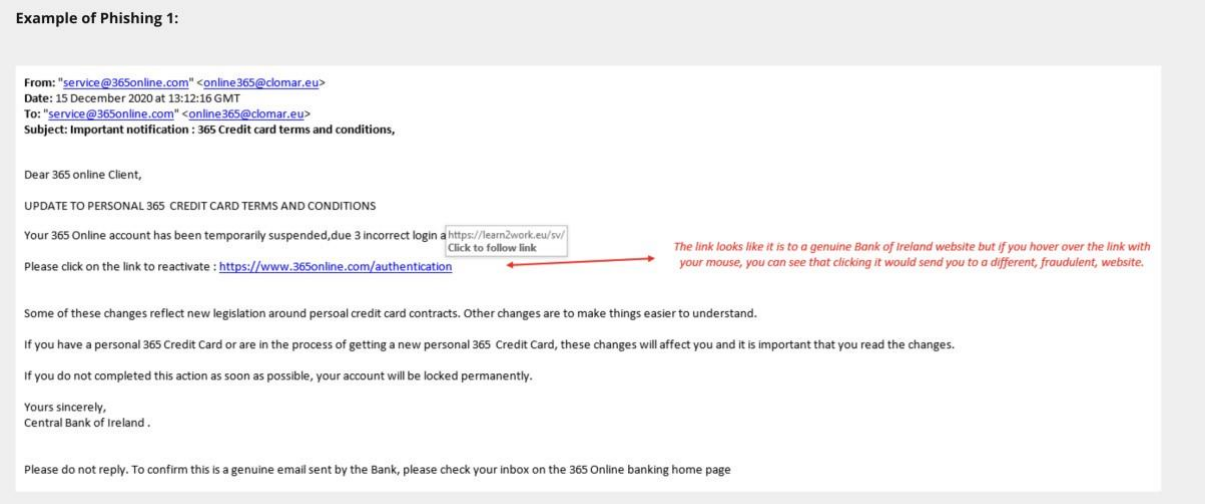

Figure 1: Phishing example 1 [Bank of Ireland, 2025]

Quishing

This attack vector involves threat actors trying to convince their victims into scanning QR codes on their phone and then redirecting them to a malicious website and getting them to enter sensitive information or download malware. Threat actors typically embed the QR codes into phishing emails.[Cloudflare, 2025].

What makes quishing so effective is that most email security services tend to overlook QR codes as they treat them as harmless images. Research has shown that current machine learning phishing detection models struggle with QR code attacks due to embedded links or text-based URLs that traditional scanners can analyse. A 2022 case study involving Deutsche Bank demonstrates the danger of these attacks. This attack utilised QR codes disguised as security verification notices, enabling threat actors to bypass email filters entirely and target the bank's corporate clients [Trivedi et al., 2025].

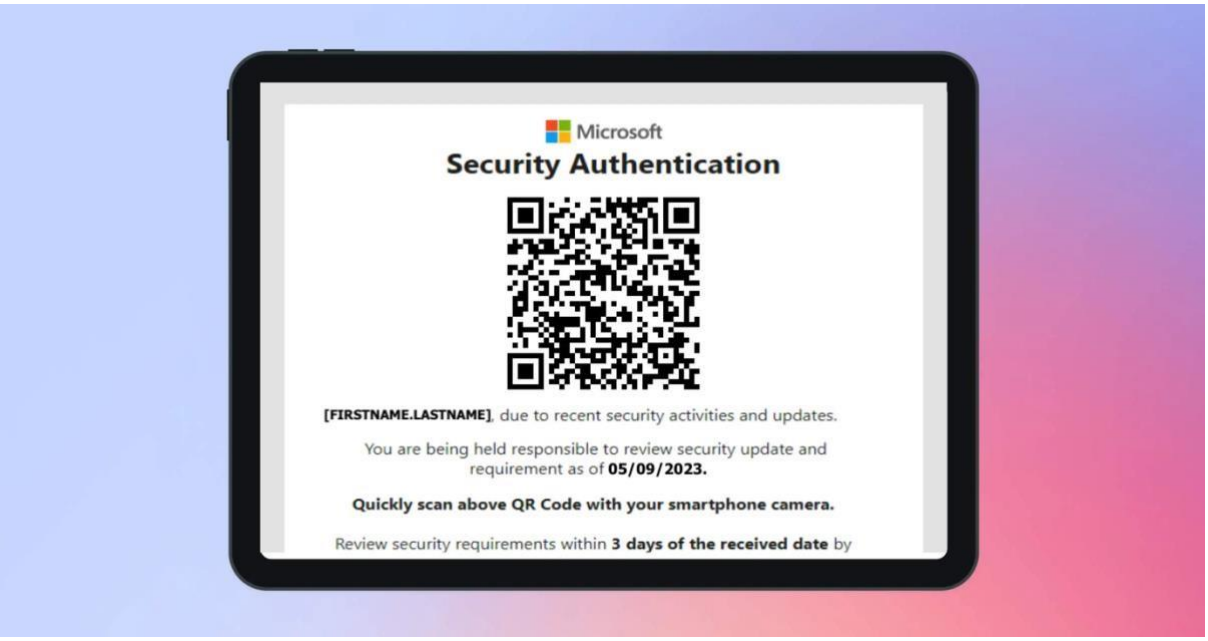

Figure 2: Quishing example (Hoxhunt, 2025)

AI-generated Phishing Emails

Thanks to the rise of Generative AI and Large Language Models (LLMs), the world of phishing emails has undergone significant changes. Threat actors are now able to craft grammatically perfect and customised phishing emails in seconds [Baker & Cartier, 2025]. These new AI-generated emails do not exhibit any of the telltale signs of phishing, such as poor grammar, incorrect phrasing or any other inconsistencies.

Research has shown that AI-assisted generated emails enable attackers to personalise emails to suit the tone, language, and cultural references, matching individual targets or users in a specific country or region. This means that phishing attacks are much more challenging to detect.

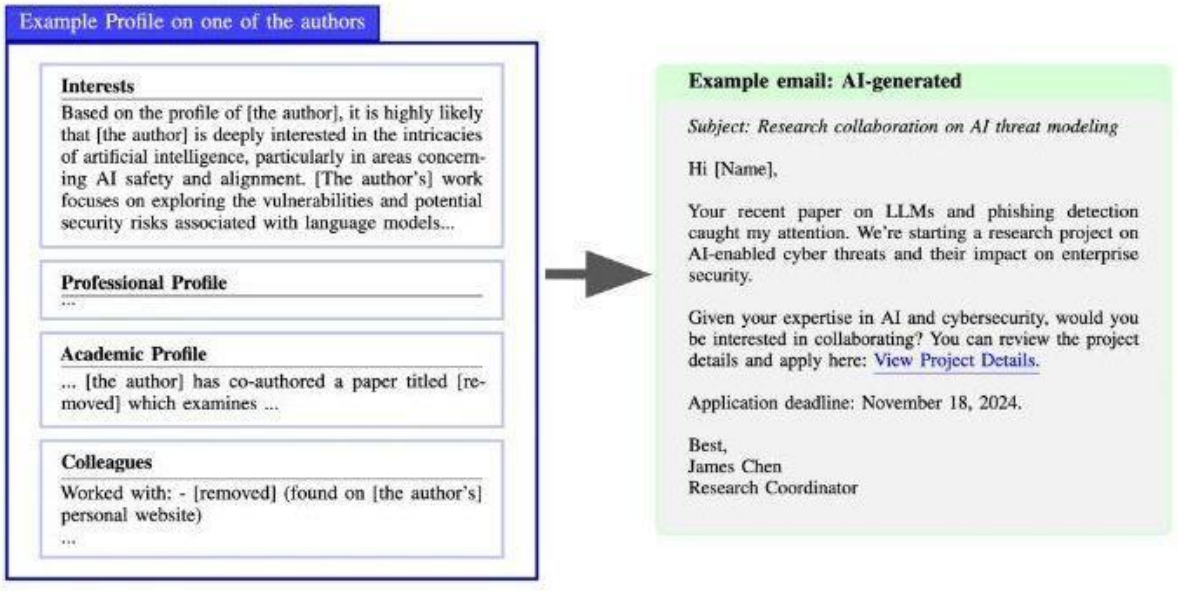

Figure 3: AI Phishing email [Malwarebytes, 2025]

Smishing

This technique involves the threat actors sending malicious text messages to users and trying to convince them to open a link and enter sensitive credentials/or brings them to a fake payment portal [Central Bank of Ireland, 2014]. These types of attacks can be more effective

than traditional phishing emails, as people tend to put more trust in their mobile phone security over their computer and as a result are less suspicious of SMS messages.

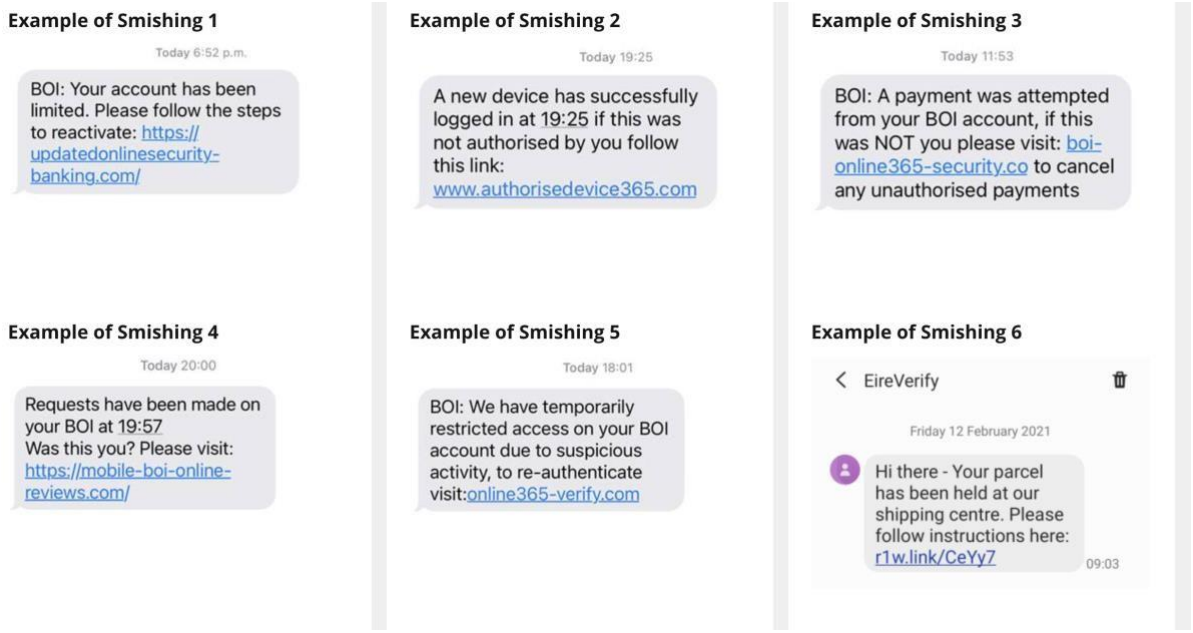

Figure 4: Smishing Examples [Bank of Ireland, 2025]

Spear-phishing and Whalephishing

Spear-phishing involves threat actors researching their targets in detail before they craft their extremely personalised attack vectors. Whalephishing refers to targeting high-value people such as CEOs, CFOs or other C-suite executives [Fortinet, 2025]. These attacks generally leverage OSINT gathered from social media, company websites and any other posts or articles publicly available.

Deepfakes/Vishing

Deepfakes are generally AI-generated videos that try to clone someone you know or trust to convince you to give out credentials or transfer them money. Vishing is a technique where threat actors ring their targets/victims and try to convince them to hand over information. Thanks to AI, threat actors are no longer able to clone voices and mimic someone you know. These two methods can be combined so that the targets believe they are speaking to a real person [Sakthivel A, 2024].

Phishing-as-a-Service (PhaS) kits

These are ideal for less skilled threat actors as they provide hosting and infrastructure. There are kits that are designed to steal MFA tokens by acting as reverse proxies to steal session tokens or one-time passwords [Sakthivel A, 2024].

4. Social Engineering

4.1 What is social engineering

According to NIST, “The act of deceiving an individual into revealing sensitive information, obtaining unauthorised access, or committing fraud by associating with the individual to gain confidence and trust.” [NIST, 2020]

Phishing is one of the most common social engineering techniques.

4.2 Psychological Aspects

Social engineering attacks work because they exploit fundamental psychological principles that are deeply embedded in all human consciousness. An understanding of these principles is essential for threat actors and defenders.

Authority and Obedience

The principle of authority is one of the strongest principles in social engineering. Research done by Stanley Milgram in the 1960s demonstrated that people tend to obey people of authority even if it means doing something harmful. [Milgram, 1963]. In the context of social engineering, threat actors impersonate people of authority in their company, such as C-suite executives, IT Staff or law enforcement.

When a victim receives an email or phone call from someone pretending to be the CEO of their company, and they are demanding urgent action, the victim is much more likely to comply than if the request came from an unknown person or unknown team member.

Urgency and Time Pressure

Threat actors deliberately create artificial time constraints to try and force their victims into a sense of urgency and force them to act without thinking. When a victim thinks they only have a limited time to respond, they are less likely to engage in critical thinking and actually evaluate the request that is being made. Research in cognitive psychology demonstrates that time pressure impairs explicitly the ability to detect deception and inconsistencies [Montanez et al, 2020]. Business Email Compromise (BEC) attacks use urgency tactics to attempt to compromise their targets. This creates an illusion of needing to act immediately and often overrides the victim’s ability to verify the legitimacy of the request.

Social Proof and Conformity

Social proof is a psychological phenomenon where people believe that actions performed by multiple other people are likely correct. Threat actors leverage this by fabricating scenarios where numerous people appear to have already complied with the same request. When they make it seem that most of the victims’ colleagues have already clicked the link and entered credentials, then the victims are more likely to assume that providing information is safe to do so. [Cognisys, 2024]

Trust and Likeability

Attackers attempt to build up trust with their victims before attempting their attack. This can happen in many ways, such as through emails, phone calls, and social media interaction. The threat actors build up emotional connections with their victims, which makes they much more likely to fulfil the threat actors request [Passeri, 2022].

Fear

Threat actors exploit fear by sending loads of fake alarms or false threats to their victims. Common tactics include scareware, such as fake pop-ups or emails saying their account was accessed. The threat actors then say that if the victim does not act fast, then sensitive information will be released [Marchand, 2025]

Individual Differences in Susceptibility

Research has identified personality and experiential factors that affect susceptibility to social engineering attacks. People with lower risk perception are more vulnerable to certain phishing methods [Wright & Marett, 2010]. Age, expertise and some other factors also play significant roles. Elderly people who are less familiar with technology are more likely to be victims rather than younger people.

4.3 Techniques of Social Engineering

- **Baiting/Clickfixing:**

  This is where attackers attempt to draw in their victims by promising them something that appeals to them. Common baiting tactics include:

  - Offering free software, games, movies, etc.
  - Offering Cash or Gift cards.

  The attackers rely on the curious nature of their victims to overcome their rational thinking for the chance of gaining something [Unit 42, 2025].

  Common Clickfixing methods:

  - Fake Captcha.
  - Fake Pop-up saying you need to update software.

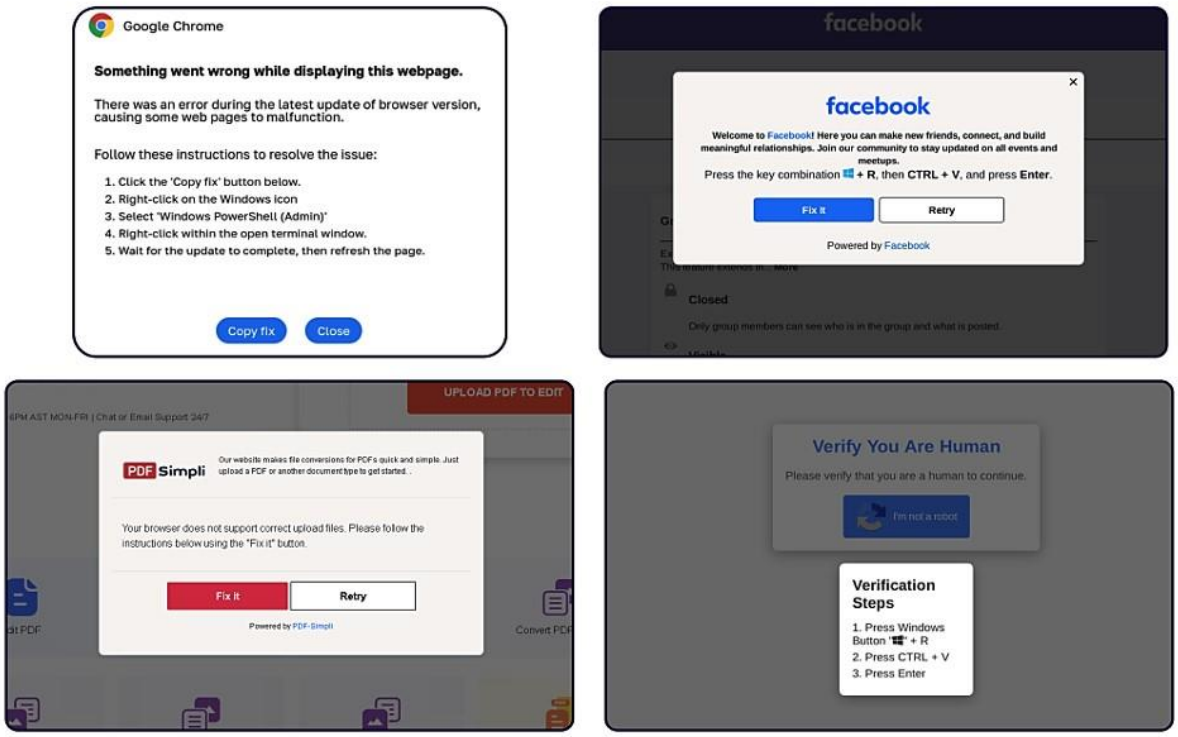

Figure 5: ClickFix tactic used by malicious websites impersonating Google Chrome, Facebook, PDFSimpli, and reCAPTCHA [HSS, 2024]

- **Scareware**:
  - Threat actors bombard a victim with fake threats or false alarms, hoping they will react without thinking, resulting in the victim clicking on the malicious payload. Examples of this include pop-ups claiming your computer is infected or stating that someone has signed in to your account [Unit42, 2025].

Figure 6 Example Scareware pop-up [Avast, 2022]

- **Pretexting**:

- Pretexting involves threat actors using fabricated stories to gain a victim's trust and manipulate them into sharing sensitive information, sending money, or downloading malware [Holdsworth and Kosinski, 2024]. One example is an attacker impersonating an IT support member, pretending you need to update your system.

Figure 7: Pretexting attack techniques [Fortinet, 2023]

- **Phishing**:
  - This is where threat actors make a payload to try and act on the victim's sense of urgency and curiosity, and hope they click on a malicious link or provide sensitive credentials.

- **Multi-Vector Attacks:**
  - Modern social engineering campaigns often use multiple techniques together to get the best results. A simple phishing email may be followed up with vishing, where both attacks use OSINT to create convincing pretext stories. This leads to more people falling victim to threat actors. [Fortinet, 2025]

- **Tools:**
  - Attackers can use pre-built tools to help them perform social engineering attacks. One tool is SeToolkit. It is an open-source tool that provides numerous templates for testing social engineering attacks. [TrustedSec, 2025].

5. Technologies I Will Use

5.1 Programming Languages

- Python
  - This is going to be the primary language of my project. It has many built-in libraries that I will utilise, such as Django, which is for web development. It also has additional libraries for QR code generation, email sending and logging

- SQL
  - PostgreSQL will be the primary database.
  - Django's Object-Relation Mapper (ORM) will act as a secondary database. Django's ORM automatically handles parameterised queries, which reduces SQL injection vulnerabilities compared to raw SQL.

- JavaScript

- Provides advanced features to allow me to enhance the user interface, such as form validation, etc.

- HTML & CSS
  - These two languages will allow me to build the web interface, landing pages and templates for my simulator.

5.2 Frameworks and Libraries:

- **Django:**
  This is a lightweight web framework that provides the following:

  URL Routing: Django's URL dispatcher maps HTTP requests to appropriate functions or class-based views.

  Models (ORM): Django's Object-Relational Mapper automatically handles SQL parameterisation and prevents injection attacks. ORM supports :
  - Query building with method chaining.
  - Automatic database migrations.
  - Query optimisation and prefetching.

  Admin interface:
  Django provides an automatically generated admin interface for managing database records.

  Template Engine:
  The Django template language enables dynamic HTML generation with variable substitution, conditional logic, and template inheritance, which all reduce code duplication.

  Authentication And Authorisation:
  Built-in systems for user authentication and authorisation are in place by default.

  Django REST Framework (DRF):

  Django's REST framework allows users to extend Django with tools for building RESTful APIs.

- **PostgreSQL16:**

  This has features such as:
  - Campaign storage: Campaign metadata, scheduling.
  - Interaction Tracking: Email opens, Links, clicks, reporting events.
  - User Management: Profiles, Grouping, calculated risk scores.
  - Security: SSL required connections.

- **Celery & Redis:**

  Campaign Delivery features:
  - Email sending for 1000+ users at once.
  - Automatic retry logic.
  - Scheduled campaign capability.
  - Task monitoring and failure alerts.
  - Message queueing for emails.
  - Session storage for admin authentication.
  - Rate limiting for API protection.

6. Ethical Issues

6.1 Informed Consent vs Deception:

There is an ethical dilemma between informed consent and a realistic simulation in creating a phishing simulator. Informed consent requires that all participants understand the purpose, nature, and potential consequences of participating in a phishing simulation before it occurs. However, this transparency creates a paradox of ineffectiveness in testing, as employees who know they are being tested possess an inherent advantage. In contrast, real-world threat actors provide no such warning. This creates opposing pressures between the need for authentic threat assessment and respect for employee autonomy and well-being [Hatfield, 2019].

Benefits of Informed Consent

Research demonstrates that transparency regarding security simulations reduces anxiety and shame compared to environments where employees feel under surveillance without warning [CREST, 2024]. Informed consent strengthens GDPR Article 6 (1)(f) which states "processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child." [European Data Protection Board, 2024]. By designing a transparent testing framework it helps build trust and reduces the likelihood of employees failing to report phishing attempts. Blame-oriented cultures reduce organisational detection rates for genuine threats as they make employees feel silly for falling for a fake phishing attempt, which makes them less likely to report a real phishing attack [CREST, 2024].

A study of 1,300+ employees across 20 European organisations found that continuous phishing simulations combined with immediate feedback achieved a 52% reduction in phishing susceptibility within six to eight months, with approximately 70% of employees who fell for an initial simulation not repeating unsafe actions in subsequent exposures [Tóth et al., 2024].

Drawbacks and Mitigation

The primary disadvantage is testing bias, employees who know they may be targeted exhibit higher vigilance than they would against real attacks. To attempt to mitigate this, some strategies include randomising simulation timing and frequency, changing attack vectors, conducting different campaigns for different employee groups, and providing blinded reporting that emphasises aggregated metrics [Tóth et al., 2024].

6.2 Best Practices and Ethical Framework

Transparency and Communication

Organisations must clearly communicate that simulations are educational tools designed for collective protection, not punitive mechanisms. CREST research demonstrates that organisations adopting a culture of blame regarding phishing staff face unintended consequences, including elevated anxiety, reduced reporting, and decreased productivity [CREST, 2024]. Consent forms should specify the nature of simulations, channels of deployment, data handling procedures, and the non-punitive nature of the exercise. The NIH Office of Science Policy guidance emphasises that consent documents should be written at an 8th-grade reading level with clear headers, short sentences, and plain language descriptions [NIH Office of Science Policy, 2024].

Non-Punitive Feedback

Phishing simulations must prioritise education over punishment. Research directly comparing punitive versus educational approaches demonstrates that mandatory training and punishment approaches produced significant adverse outcomes: participants viewed them as unfair, both increased state anxiety, mandatory training inhibited productivity, whilst brief, contextual training was seen as fair and reduced phishing susceptibility without productivity loss [CREST, 2024].

6.3 Legal and Regulatory Considerations

Organisations have the legal right to conduct security awareness testing, including phishing simulations, as part of legitimate business operations and risk management [Hatfield, 2019]. However, consent and notification requirements vary by jurisdiction; some require explicit consent before surveillance or testing can occur [European Data Protection Board, 2024]. In some jurisdictions, spoofing email domains or simulating phone calls (vishing) may be prohibited by telecommunications laws. [Hatfield, 2019] Beyond GDPR, other privacy frameworks (CCPA in California, PIPEDA in Canada, UK Data Protection Act 2018) impose requirements on employee data handling [European Data Protection Board, 2024]. Simulations should not disproportionately target, or burden, protected groups based on age, gender, disability, or other protected characteristics [Hatfield, 2019].

For large-scale security testing programmes, independent ethical review can strengthen legitimacy and effectiveness. [NIH Office of Science Policy, 2024] Reviewers can assess risk-benefit proportionality, ensure consent adequacy, recommend protocol modifications, and establish ongoing monitoring for adverse effects. [CREST, 2024]

Maintain comprehensive documentation, including the programme charter, consent forms, legitimate interest assessment, campaign records, training records, adverse incident reports, and regular review minutes. Such documentation demonstrates good faith and responsible security management. [European Data Protection Board, 2024]

7. Security Implications

- Phishing simulations allow admins to improve the awareness of all users in a safe environment.
- Metrics collected by the tool can help identify some employees that could benefit from some phishing training.
- If admins decide to test users continuously, it creates more security-aware employees, which in turn provides a safer organisation.
- If my tool gets misconfigured and exposed, threat actors abuse the platform.
- Implement authentication to ensure that only people I grant access to can use the tool.
- Ensure I follow secure coding practices.

8. Existing Tool Analysis

8.1 GoPhish

Detailed weaknesses:

- Scalability: Can only handle 50 emails per minute. So, for a company with 500 users, it would take 10 minutes.
- No Async Processing: Synchronous email delivery blocks UI during large campaigns.
- Template Limitations: Only supports basic HTML templates.
- Report Limitations: CSV reports only.

8.2 KnowBe4

Detailed weaknesses:

- Expensive: It is between €40 and €120 per user a month, which is extremely expensive for large organisations.
- Template Fatigue: 2.5m simulations reveal that after six campaigns, 34% recognise the simulation. [Hoxhunt 2025]
- Limited Customisation: You must use pre-built templates unless you are willing to pay more.

8.3 Microsoft Defender Phishing Simulation Tool

Detailed weaknesses:

- Ecosystem Users Only: Must have an existing license with Microsoft, even to have the option for this add-on.
- Additional License Required: An additional €30 per user per month.
- 5 Templates Only: Credential Harvesting, MFA fatigue, Malicious Office Docs.
- No Detailed Analytics: Only basic fail/pass counts provided.

A report done by sosafe showed that only 24.5% threat of compromise reduction vs 66% lab claims [sosafe, 2025]. This is mainly due to unrealistic templates and delayed training.

Summary and Conclusions

This project investigates the awareness of users in terms of modern phishing techniques and social engineering. It highlights how advanced threat vectors are getting with the use of Vishing and Quishing. Thanks to these new attack vectors, I believe that phishing is going to get harder and harder to detect, which will result in more attacks.

This is where a tool like mine comes into play. This tool is aimed at organisation admins to help ensure that all employees are aware of the most up-to-date attack vectors that are currently out there. My tool will initially support configurable phishing campaigns on a small scale. It will collect metrics and generate reports that administrators can provide to the rest of the organisation to facilitate further training.

I have also considered the ethical standpoint and other security issues, where I ensure that consent is given by every user, I test this in practice, and that they are informed of what to expect.

References

Attacks, C. (2024). HC3: Sector Alert. [online] Available at: https://www.hhs.gov/sites/default/files/clickfix-attacks-sector-alert-tlpclear.pdf [Accessed 2 Dec. 2025].

Baker, E. and Cartier, M. (2025). Phishing Trends Report (Updated for 2025). [online] Hoxhunt.com. Available at: https://hoxhunt.com/guide/phishing-trends-report#part-iphishing-trends-amp-statistics [Accessed 2 Dec. 2025].

Centralbank.ie. (2014). What is smishing | Central Bank of Ireland. [online] Available at: https://www.centralbank.ie/consumer-hub/explainers/what-is-smishing [Accessed 2 Dec. 2025].

Cloudflare (2024). What is a phishing attack? | Cloudflare UK. [online] Cloudflare. Available at: https://www.cloudflare.com/en-gb/learning/access-management/phishing-attack/ [Accessed 2 Dec. 2025].

Cloudflare (2025). What is quishing. [online] Available at: https://www.cloudflare.com/engb/learning/security/what-is-quishing/ [Accessed 2 Dec. 2025].

Danielson, L. (2024). Statistics on Phishing Attacks that Target Businesses | Huntress. [online] Huntress.com. Available at: https://www.huntress.com/phishing-guide/phishingattack-statistics [Accessed 2 Dec. 2025].

DeepStrike (2025). Phishing Statistics 2025: How AI, Behaviour, and First-Party Data Are Redefining Cyber Defence. [online] DeepStrike. Available at: https://deepstrike.io/blog/Phishing-Statistics-2025 [Accessed 2 Dec. 2025].

Fortinet (2025). What Is Social Engineering? Preventing Social Engineering Attacks. [online] Fortinet. Available at: https://www.fortinet.com/resources/cyberglossary/social-engineering [Accessed 2 Dec. 2025].

Holdsworth, J. and Kosinski, M. (2024). Pretexting. [online] ibm.com. Available at: https://www.ibm.com/think/topics/pretexting [Accessed 2 Dec. 2025].

IBM (2025). Cost of a data breach report 2025. [online] IBM. Available at: https://www.ibm.com/reports/data-breach [Accessed 2 Dec. 2025].

Marchand, C. (2025). The Psychology of Social Engineering. [online] Coalitioninc.com. Available at: https://www.coalitioninc.com/blog/security-labs/the-psychology-of-socialengineering [Accessed 2 Dec. 2025].

NIST (2020). social engineering - Glossary | CSRC. [online] csrc.nist.gov. Available at: https://csrc.nist.gov/glossary/term/social_engineering [Accessed 2 Dec. 2025].

Passeri, P. (2022). How Threat Actors Weaponise Your Trust. [online] Infosecurity Magazine. Available at: https://www.infosecurity-magazine.com/blogs/threat-actorsweaponize-trust/ [Accessed 2 Dec. 2025].

Sakthivel, A. (2024). Threat Spotlight: Phishing techniques to look out for in 2025. [online] Barrcuda Blog. Available at: https://blog.barracuda.com/2024/12/04/threat-spotlightphishing-techniques-2025 [Accessed 2 Dec. 2025].

sosafe (2025). Effectiveness of Phishing Simulations. [online] SoSafe. Available at: https://sosafe-awareness.com/blog/real-world-data-effectiveness-phishing-simulations/ [Accessed 2 Dec. 2025].

Trivedi, A., Jangal, K. and Gupta, R. (2025). Phishing Detection in Advanced QR Code Attacks: Challenges and AI-Driven Solutions. International Journal for Research in Applied Science and Engineering Technology, [online] 13(1), pp.479–482. doi:https://doi.org/10.22214/ijraset.2025.66306.

TrustedSec (2024). The Social Engineering Toolkit (SET). [online] TrustedSec. Available at: https://trustedsec.com/resources/tools/the-social-engineer-toolkit-set [Accessed 2 Dec. 2025].

Unit 42 (2025). 2025 Unit 42 Global Incident Response Report: Social Engineering Edition. [online] Unit 42. Available at: https://unit42.paloaltonetworks.com/2025-unit-42-globalincident-response-report-social-engineering-edition/ [Accessed 2 Dec. 2025].

Zensec (2025). 2025 phishing statistics: The alarming rise in attacks - Zensec. [online] Zensec. Available at: https://zensec.co.uk/blog/2025-phishing-statistics-the-alarming-rise-inattacks/ [Accessed 2 Dec. 2025].

Canham, M., Posey, C., Strickland, D., Constantino, M. (2021). Phishing for Long Tails: Examining Organisational Repeat Clickers and Protective Stewards. *SAGE Open*, 11(1), 21582440219906561. [online] Available at: https://journals.sagepub.com/doi/full/10.1177/2158244021990656 [Accessed 9 Dec. 2025].

CREST. (2024). *Phishing your staff: A double-edged sword? Guide to ethical simulated phishing*. Centre for Research & Education on Security and Trust. [online] Available at: https://crestresearch.ac.uk/resources/phishing-your-staff/ [Accessed 9 Dec. 2025].

European Data Protection Board. (2024). *Guidelines 1/2024 on processing of personal data based on legitimate interest (Article 6(1)(f) GDPR)*. Publications Office of the European

Union. [online] Available at: https://www.edpb.europa.eu/system/files/202410/edpb_guidelines_202401_legitimateinterest_en.pdf [Accessed 9 Dec. 2025].

Hatfield, J.M. (2019). Virtuous human hacking: The ethics of social engineering in penetration-testing. *Computers & Security*, 83, 354-366. [online] Available at: https://www.sciencedirect.com/science/article/abs/pii/S016740481831174X [Accessed 9 Dec. 2025].

NIH Office of Science Policy. (2024). *Informed Consent for Research Using Digital Health Technologies: Points to Consider & Sample Language*. U.S. Department of Health and Human Services. [online] Available at: https://osp.od.nih.gov/wpcontent/uploads/2024/05/DigitalHealthResource_Final.pdf [Accessed 9 Dec. 2025].

Tóth, R., Dubniczky, R.A., Limonova, O., & Tihanyi, N. (2024). Sustaining Cyber Awareness: The Long-Term Impact of Continuous Phishing Training and Emotional Triggers. *arXiv Preprint*, 2510.27298v1. [online] Available at: https://arxiv.org/html/2510.27298v1 [Accessed 9 Dec. 2025].

Appendix

Table of Figures