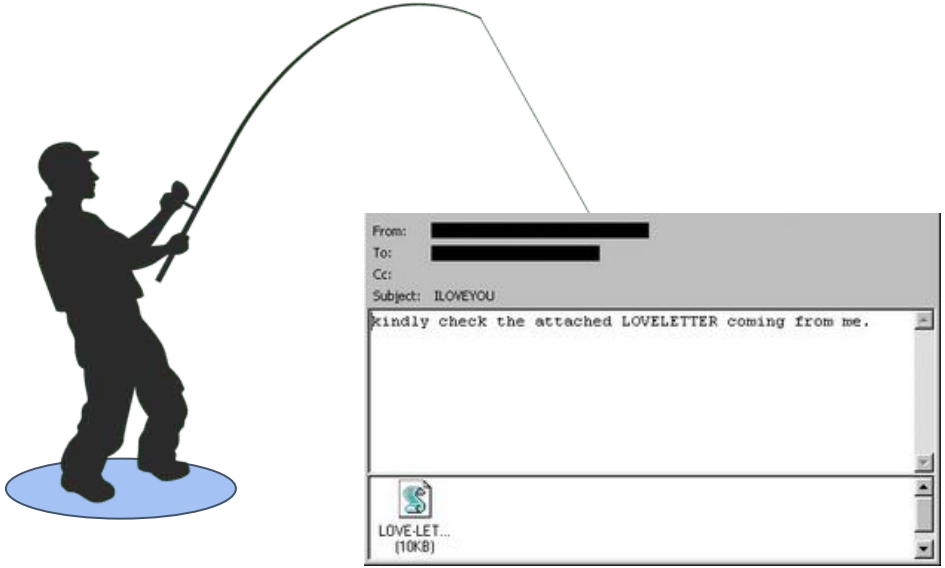
GonePhishing

An Interactive Cyber Security Simulation Sandbox

"Here to make sure you know what you're reeling in."



Meet the team

Zandr

Team Lead & QA/QT

Tonio

UI & Email Generation

Ethan

Attack Specialist

Tyler

Database Lead

Aaron

Data & Statistics

Heaven

Scoring and Feedback

Speaker 1: Zandr (Team Lead)

Project Aim and Scope

1

Our aim

To design and build a prototype phishing simulation sandbox where staff practise spotting phishing emails and receive a score and instant feedback.

2

Scope

A simulated inbox with AI written phishing, spoofing and link redirection scenarios, plus scoring, feedback and reports.

3

Success

One scenario that a player can complete from start to finish, with their score, feedback and results saved and shown in a report.

Our aim builds on learn by doing training, Kumaraguru et al. [2] found this effective in user studies.

Speaker 1: Zandr (Team Lead)

Project Objectives

1 Simulated inbox

Build a realistic inbox with a reading pane and sender details where players Report or mark Safe.

2 AI scenarios

Generate AI written emails covering phishing, spoofing and link redirection, with hidden warning signs.

3 Secure data

Set up a MySQL database with personal logins, hashed passwords and results stored per user.

4 Scoring and feedback

Create a 0 to 100 behavioural score and instant feedback explaining why each action was risky or safe.

5 Reporting

Turn stored results into bar and line charts that show mistakes and progress over time.

6 Test and review

Get one scenario working end to end, test it with users and use the feedback to plan next steps.

Speaker 1: Zandr (Team Lead)

Project Vision and Target Audience

1

Our vision

Traditional security training is a video or slide deck that people forget instantly. GonePhishing lets staff practise spotting phishing emails in a safe, simulated environment.

2

Audience

Organisations with remote or hybrid staff, such as MSPs, small businesses and universities.

According to Vishwanath et al. [1], email habits wear down the effect of training, so people need regular practice.

Speaker 1: Zandr (Team Lead)

Existing Solutions and Our Unique Value

1

What exists

Platforms such as KnowBe4 already send fake phishing emails, so the idea is proven.

2

The gap

Results often focus on whether someone clicked, with less detail on why a choice was good or bad.

3

Our value

GonePhishing gives a score for every action, instant advice and AI written scenarios that suit each user.

Kumaraguru et al. [2] built PhishGuru and a game called Anti Phishing Phil, which their user studies show that this style of training works.

Speaker 1: Zandr (Team Lead)

Resources and Collaboration Tools

1

VS Code

Our shared code editor for the JavaScript backend and interface.

2

GitHub

Holds our code, branches and pull requests. The team leader reviews and tests before merging.

3

Teams

Frequent updates, weekly meetings and upload notices so no one is left waiting.

4

Google Docs

Docs and Slides let multiple people edit at the same time.

Speaker 2: Tonio (UI Lead)

Simulated Email Interface

Familiar layout

- It’s imperative that it is built to look like a real work email, requiring an inbox list, a reading pane, and clear sender details.
- Users are given a score, level and streak to track progress, and achievements are rewarded to build a positive association with successfully spotting a fake email

Natural reactions

- Players open each email and decide, whether to Report it as phishing or that is is safe
- A realistic environment makes players react as they would at their desk.
- This helps to create mind muscle memory, as it will help create cognitive associations with their simulated inbox and real inbox

Williams et al. [3] found that cues such as authority made people more likely to click, so our inbox copies real workplace email.

Speaker 2: Tonio (UI Lead)

Onboarding Tutorials and AI Scenarios

1

Welcome

A short guided tour greets new players.

Shows how the inbox works before scoring starts.

2

Practise

Users try one guided example first.

Helpful hints teach them the controls.

3

Generate

AI API calls write a new email each round.

No two sessions are ever the same.

4

Adapt

The AI builds a new scenario every time.

Players learn to spot many attack types.

According to Yamin et al. [4], large language models can generate dynamic, adaptable cyber security exercise scenarios.

Speaker 3: Ethan (Attack Specialist)

Modern Attack Research

1

Phishing

Fake emails that push people to click a link or share a password.

Urgent: confirm your password

Sign in

2

Spoofing

Forged sender names and headers that make a message look trusted.

Real: hr@company.com

Fake: hr@company-verify.com

3

Link redirection

Links that look safe but lead to a fake page, which we will simulate safely.

Looks like: portal.company.com

Goes to: company-verify.com

Alkhalil et al. [5] map the stages of a phishing attack, which helps us build realistic simulated attacks.

Speaker 3: Ethan (Attack Specialist)

Why These Three Attack Vectors?

Grounding our simulated attack methodology in empirical data from modern threat research.

1 Phishing

Prevalence & Impact

Primary Initial Access Vector

Over 90% of organizational breaches start with a phishing email. Targets human psychological vulnerabilities rather than technical flaws.

Simulation Goal:

Test baseline human vigilance and link identification.

2 Spoofing

Prevalence & Impact

High Trust Exploitation

Header forgery bypassing basic user scrutiny. Spoofed executive identity accounts for $2.7B+ annual losses in BEC attacks.

Simulation Goal:

Validate technical domain security (SPF/DKIM) & authority response.

3 Link Redirection

Prevalence & Impact

Deceptive URL Routing

Links that appear legitimate (e.g. *portal.company.com*) but route users through open redirects to malicious landing pages (e.g. *company-verify.com*).

Simulation Goal:

Train users to inspect destination URLs and hover over links safely.

Research Foundation: *Alkhalil et al. (2021) "Phishing Attacks: Recent Comprehensive Study and a New Anatomy" maps attack phases to model realistic, stage-by-stage simulations.*

Speaker 3: Ethan (Attack Specialist)

Threat Realism and UI Coordination

1

Grounded in real cases

Wright et al. [6] sent simulated phishing emails to employees during normal work. We use realistic work situations to choose which attacks to include.

Real work situation → Attack chosen → **Email scenario**

2

Working with Tonio

Ethan reviews the generated emails so that Tonio can place small warning signs in each email, such as an odd sender address, a mismatched link or urgent wording.

Ethan reviews the emails → **Tonio adds warning signs** → **Users spot the clues**

Hidden clues let users learn the signs of a scam by looking first; instant feedback then confirms what they found or missed.

Speaker 4: Tyler (Database Lead)

Database Selection and User Login

1

MySQL Database - DBeaver or MySQL Workbench (for DB management)

A relational database that organises information into connected tables, that supports SQL queries to retrieve, filter and analyse results.This can be integrated with the JavaScript backend planned for GonePhishing, and can be also be developed locally before the system is deployed online.

2

Secure login

Each employee has a personal account, so every score belongs to one person and can be tracked over time.

Codd [7] introduced the relational model of linked tables, which keeps each person's results separate and easy to search.

Speaker 4: Tyler (Database Lead)

Data Provisioning and Privacy Compliance

1

Tables for Aaron

Results are saved per email, per user and per date so graphs can be built from them.

2

Encryption

Passwords are encrypted (hashed) and personal data anonymised when stored and when sent.

3

Privacy rules

We collect only what we need and follow data protection rules such as GDPR.

Tikkinen-Piri et al. [8] explain what the GDPR means for companies that collect personal data, so we collect only what we need.

Speaker 5: Aaron (Data Lead)

Automated Statistics and Visual Reporting

1 Collect
Results are retrieved from the database after each phishing simulation.

2 Count
Calculate scores, link-click rates, reporting rates and correct decisions.

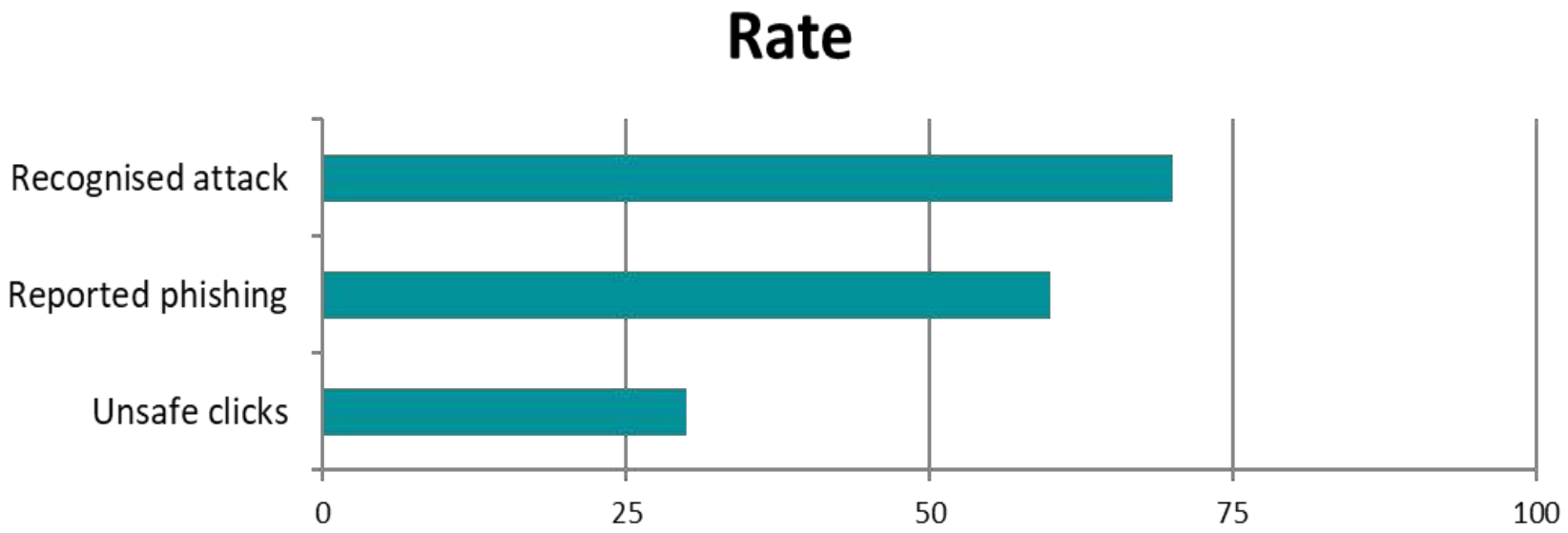
3 Graph
Generate bar charts comparing outcomes and line graphs showing progress.

4 Review
Managers identify common mistakes, strengths and areas needing training.

Cleveland and McGill [9] tested how accurately people read different charts, so we use simple bar and line charts.

Speaker 5: Aaron (Data Lead)

Illustrative company outcomes



Bar graphs compare various results better than Line graphs, as they are more suitable when comparing changes over repeated trials, whereas pie graphs don't work as well when comparing multiple values.

Speaker 5: Aaron (Data Lead)

Data Warehousing and Trend Tracking

1 Example

A user scores 5/10 on their first phishing test and 9/10 on a later test.

Store both results and use a line graph to show improvement.

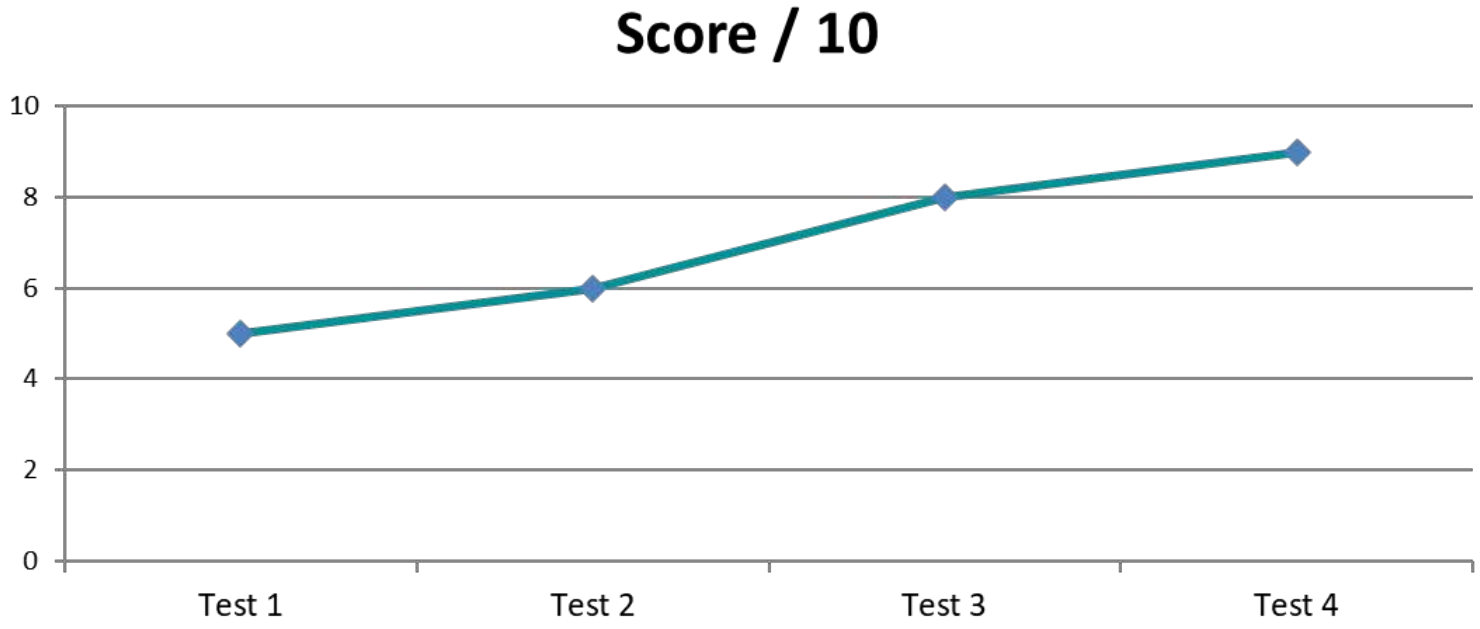
2 Trends over months

Compare monthly results to see whether the organisation is improving.

Track click rates, reporting rates and common phishing mistakes.

Speaker 5: Aaron (Data Lead)

Data Warehousing and Trend Tracking



Speaker 6: Heaven (Scoring and Feedback Lead)

Behavioural Scoring and Assessment System

1

The problem

Measuring clicks alone doesn't show everything someone did. For example, someone might click and then report an email.

2

Our proposed solution

GonePhishing will use a 0-100 scoring system that considers actions such as reporting, deleting, clicking and submitting credentials.

3

Limitations

More serious actions receive lower scores. Reporting a phishing email after clicking or submitting credentials is recognised, but it does not erase the earlier risk.

Proposed scoring examples (Illustrative)

Click → Report	Submit credentials → Report
25/100	**10/100**

Dawkins and Jacobs [10] explain that phishing-email difficulty provides important context when interpreting simulation results.

Speaker 6: Heaven (Scoring and Feedback Lead)

Feedback and Continuous Improvement

1 Immediate feedback

After each simulated email, users receive their score, an explanation of why their actions were safe or risky, and practical advice on how to improve

2 Session Review

Users can review individual decisions, identify mistakes and receive recommendations based on their most serious errors.

3 Progress tracking

Previous and current scores are compared to show changes in simulation performance over time.

4 Platform Improvement

An optional survey collects ratings and suggestions about email realism, feedback clarity and learning value.

Research: Ho et al. [11] found limited benefits from embedded phishing training, with only 24% of training sessions completed. This highlights the importance of user engagement in our proposed feedback system.

Speaker 1: Zandr (Team Leader)

Costs, Risks and Roadmap

1

Costs and legal

We plan for hosting, database storage and AI usage fees. Data is encrypted and kept to a minimum, and users know it is training.

2

Risks

Users may tire of the game, feel falsely safe, or face slow AI replies. We will vary scenarios and test often.

3

Roadmap

Next, build one working phishing scenario. Then test with users and add more scenarios and reports.

GonePhishing turns passive training into active practice, so staff learn to spot phishing before it costs their company resources.

List of References

[1] A. Vishwanath, B. Harrison, and Y. J. Ng, “Suspicion, cognition, and automaticity model of phishing susceptibility,” Commun. Res., 2016, doi: 10.1177/0093650215627483.
Cited by: Zandr

[2] P. Kumaraguru, S. Sheng, A. Acquisti, L. F. Cranor, and J. Hong, “Teaching Johnny not to fall for phish,” ACM Trans. Internet Technol., vol. 10, no. 2, Art. 7, 2010, doi: 10.1145/1754393.1754396.
Cited by: Zandr

[3] E. J. Williams, J. Hinds, and A. N. Joinson, “Exploring susceptibility to phishing in the workplace,” Int. J. Hum. Comput. Stud., vol. 120, pp. 1–13, Dec. 2018, doi: 10.1016/j.ijhcs.2018.06.004.
Cited by: Tonio

[4] M. M. Yamin, E. Hashmi, M. Ullah, and B. Katt, “Applications of LLMs for generating cyber security exercise scenarios,” IEEE Access, vol. 12, pp. 143806–143822, 2024, doi: 10.1109/ACCESS.2024.3468914.
Cited by: Tonio

[5] Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, “Phishing attacks: A recent comprehensive study and a new anatomy,” Front. Comput. Sci., vol. 3, Art. 563060, Mar. 2021, doi: 10.3389/fcomp.2021.563060.
Cited by: Ethan

[6] R. T. Wright, S. L. Johnson, and B. Kitchens, “Phishing susceptibility in context: A multilevel information processing perspective on deception detection,” MIS Q., vol. 47, no. 2, pp. 803–832, 2023. [Online]. Available: https://aisel.aisnet.org/misq/vol47/iss2/13
Cited by: Ethan

List of References, Continued

[7] E. F. Codd, “A relational model of data for large shared data banks,” Commun. ACM, vol. 13, no. 6, pp. 377–387, Jun. 1970, doi: 10.1145/362384.362685.
Cited by: Tyler

[8] C. Tikkinen-Piri, A. Rohunen, and J. Markkula, “EU General Data Protection Regulation: Changes and implications for personal data collecting companies,” Comput. Law Secur. Rev., vol. 34, no. 1, pp. 134–153, 2018, doi: 10.1016/j.clsr.2017.05.015.
Cited by: Tyler

[9] W. S. Cleveland and R. McGill, “Graphical perception: Theory, experimentation, and application to the development of graphical methods,” J. Amer. Statist. Assoc., vol. 79, no. 387, pp. 531–554, Sep. 1984, doi: 10.1080/01621459.1984.10478080.
Cited by: Aaron

[10] S. Dawkins and J. Jacobs, “NIST Phish Scale user guide,” Nat. Inst. Standards Technol., Gaithersburg, MD, USA, NIST Tech. Note 2276, 2023, doi: 10.6028/NIST.TN.2276.

Cited by: Heaven

[11] G. Ho et al., “Understanding the Efficacy of Phishing Training in Practice,” in *2025 IEEE Symposium on Security and Privacy (SP)*, pp. 37–54, 2025, doi: 10.1109/SP61157.2025.00076.

Cited by: Heaven

List of References, Continued

[12] P. Wlekły, "Aesthetics and Usability of Statistics Data Visualisation through Charts: An Eyetracking Study as a Tool for Chart Analysis," European Research Studies Journal, vol. XXVII, no. 4, 2024.
Cited by: Aaron

[13] Chart.js Team, Chart.js Documentation: Getting Started, API, and Chart Types, version 4.5.1, 2025. [Online]. Available: https://www.chartjs.org/docs/latest/.
Cited by: Aaron