# Overview

This document will go over several ways we can gather click data from our phishing emails. The goal is to collect analytics in a professional, ethical, and private way. Without the use of obvious tracking links like bit.ly and storing the minimum amount of PII required.

# Options

## Option 1: Use Google Analytics 4

### Overview

We send click data from our backend to Google Analytics 4 (GA4) using its Measurement Protocol API. This lets us log when someone clicks a phishing link and view the results in GA4’s dashboard.

### Pros

* Reliable and free.
* Gives detailed analytics automatically.
* Supports anonymous event tracking.

### Cons

* Data is stored by Google, not us.
* Requires an internet connection to send events.
* Setup involves configuring GA4’s measurement protocol.

## Option 2: Build our own in-house tracker

### Overview

We could track clicks ourselves and have them go straight to our backend and store them in our own database.

### Pros

* Full Control over all data.
* No outside services needed.
* No API setup with Google.

### Cons

* We’d have to build all analytics and graphs ourselves.
* No built in visualization tools.
* Requires database setup and maintenance.

## Option 3: Hybrid

### Overview

We can combine the best of both worlds and use GA4 for tracking, but also pull GA4 data into our own database for display on our dashboard. GA4 would record all link clicks and our backend would use the GA4 data API to fetch click data. Our backend then saves that data into our own database and we could display custom charts and summaries using the stored data.

### Pros

* Professional setup
* Keeps a local copy of analytics for more customized displays
* Still benefits from GA4’s reliability and privacy controls

### Cons

* Slightly more setup complexity
* We’ll need to handle API authentication with GA4
* Adds some backend logic for syncing and storage

# Privacy and Ethics

* Minimal Data Collection: We will only collect the employee’s name and click data (like which simulated phishing link was clicked and when)
* Consent Required: In a real deployment, employees must consent to having their name tracked, for the Demo we will use placeholder or mock names.
* No Sensitive Information: We will not collect or store Personal emails, passwords, IP Addresses, or other sensitive data.
* Transparency: Documentation will make clear what data is collected, why it’s collected, and how it will be used.

# References

GA4 Measurement Protocol. Google Developers. Retrieved from: <https://developers.google.com/analytics/devguides/collection/protocol/ga4>

GA4 Data API (v1). Google Developers. Retrieved from: <https://developers.google.com/analytics/devguides/reporting/data/v1>

Product Privacy Notice. KnowBe4. Retrieved from: <https://www.knowbe4.com/legal/product-privacy-notice>

Handling of User Data During Phishing Tests. KnowBe4 Support. Retrieved from: <https://support.knowbe4.com/hc/en-us/articles/205074487-What-Happens-to-User-Data-That-Is-Entered-During-a-Phishing-Test>

ChatGPT (GPT-5). Explanation of GDPR compliance and analytics integration for phishing simulation projects. OpenAI. Retrieved from: <https://chat.openai.com/>