Collecting information about the people visiting your site is crucial - if you want to know who is visiting your site. In the case of this blog, I would like to know if anyone is reading my posts. A lot of great solutions already exist. The most popular one is *Google Analytics*. Their 'free' offering will work perfectly for most people and is used by most sites you visit.

I don't like the business model of *Google Analytics*. It exists to serve advertisers - so it tries to collect as much information as possible. Also, not all of that information is exposed to the entity running the site. For example, there is no way to export all data collected by *Google Analytics*. That means that at some point you will have to opt into their *Google Analytics 360* plan - which offers deeper insights but is way to expensive for most people/companies.

So I decided to create my own "Google Analytics" - **with a focus on privacy**.

## But wait… have you considered other options?

"Google Analytics" is far from the only analytics platform out there. Some of the ones I have considered are:

* Heap Analytic
* Amplitude
* Simple Analytics

Just to be clear, I do not think this solution will be better than any of those. The main purpose of this solution is killing time and having something to write about.

## So what is an analytics platform?

To me, an analytics platform is **a system for aggregating events**. In that defenition there are two main words that need explaination - **event** and **aggregation**.

Let's start by looking at what I mean by an **event**. An event in this context can be any action the user is doing on the site - loading the page, clicking a button, hovering the mouse over an element… You get the point. In code, that event could look something like:

These events will be generated while the user is visiting the site throught a tracking script. This tracking script will send events back to the server. The server will recieve an event that looks something like this:

The amount of these events varey. Some trackers track every mouse movement, while other trackers only track when a user enters a site. If there are a lot of these events there is no longer a point to looking at single events. We need to get the bigger picture.

This is where aggregations come in. Let's say we have 1000 of the events above. Possible insights could be:

* What browsers are the users using?
* How many pages does the average user visit?
* How long does the average user stay on the site?

To answer these questions we need to do aggregations. Aggregations in simple terms is calculating something based on a set of events. If you are familiar with functional programming, most aggregations can be formulated as a reduce step. Something like:

Doing these aggregations and letting the user query them is the main job of an analytics platform.

## I understand, but where does Gallop come in?

Gallop is not a analytics platform, or at least that is not the goal. The main goal of Gallop is to create a simple API where user submitted events can be processed and insights generated.

I want to be able to build a dashboard on top of that GraphQL API that can show staticsics of the people visiting my site.

So events generated when people visit the different pages on my site will be collected and aggregated with an indexer. This indexer will put the events into an index where they can be easily queried.

I will write more blog posts about the different aspects of the projects. The next one will be about the GraphQL API.

## Sounds cool. What will the finnished project look like?

Except from the fact that the project probably will never be finnished.

1. There will be a dashboard on this site where users can see how many people have visited the site. There will also be analytics on a per article level. Hopefully this can be added to the statis site generor.