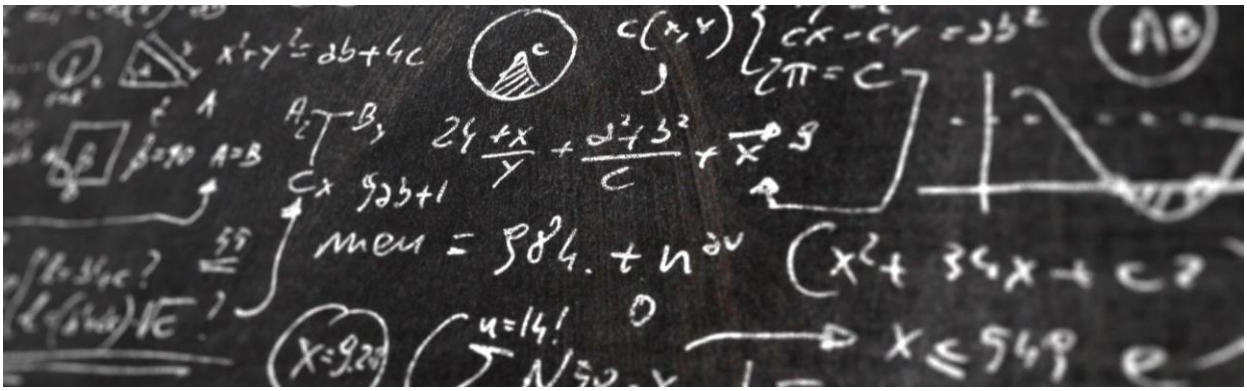
SRE: FBI, Dr's, and Scientists



Why do we study history? To learn from it and hopefully not repeat the same mistakes – at least not in the same fashion. Once an incident has happened - it's in the past, it is history. The incident lifecycle goes like this: something happened, fix it, make sure it doesn't happen again. Making sure it does not happen again is the job of an SRE and it has an official name: **RCA/Postmortem**.

The goal of an RCA is to get to the bottom of an issue (identification) and the postmortem is to dissect that analysis and act on it. To do that we need to answer the question: Why did it happen? What can we do to make sure it does not happen again?

Before we go any further, this process is meant to be blameless. **The engineered system is always at fault.** How did the system allow Eric to do that? Not, why did Eric do that? The goal is to solve the issue with engineering (and maybe at times training) because it cannot be assumed every person has encountered every situation ...which is why the system must continuously mature.

Let's look at 3 professions that investigate and some of their methodologies:

1. **Scientists – Scientific Method**

Observe the situation in question, Ask Questions, research the relevant area, make a hypothesis, test it with an experiment, analyze the data and make a conclusion.

2. **Doctors – Differential Diagnosis**

Examine clinical history, take a physical exam, perform diagnostic testing, send to a referral consultation, communication of the diagnosis, begin treatment.

3. **Police – Criminal Investigation**

Collect evidence, take witness statements, examine trend data (global), use intuition, interrogation of suspects, analysis of evidence, creation of theories, to the courts....

What can we learn from these examples? They all perform some degree of Observation, Testing, Analysis, and Conclusion Finalization.

- We need to take notes on the incident, analyze baseline data, speak to people that were on the incident call and gather their observations.
- Involve subject matter experts to dig deep into the analysis and architecture.

- Examine macro patterns and industry data of best practices and where we are deviating.
- Experiment replication of the issue
- Perform diagnostic testing
- Acting on the extrapolations of our analysis.

We must focus it through an SRE lense:

- Product – Can people be trained differently?
- Development and Automation – Can we redesign the program to avoid this issue or autocorrect it?
- Capacity Management and FinOps – Does this make financial sense to invest in this correction or refactor opposed to a new feature? Did this happen because of physical limitation?
- Testing and Releasing – Could we have tested more and caught this? Could we have released it differently and caught it?
- Incident Response and Prevention – How did we respond?
- Monitoring and Observability – How can we catch the issue programmatically and sooner?