Postmortem Culture
## Learning from Failure

YURY NIÑO

DevOps Engineer
**Chaos Engineering Advocate**

Digital Labs
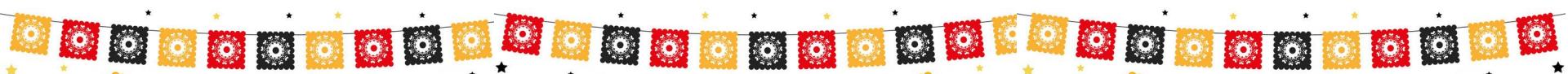
Garagoa is a town located in the **Boyacá** Department in **Colombia**.
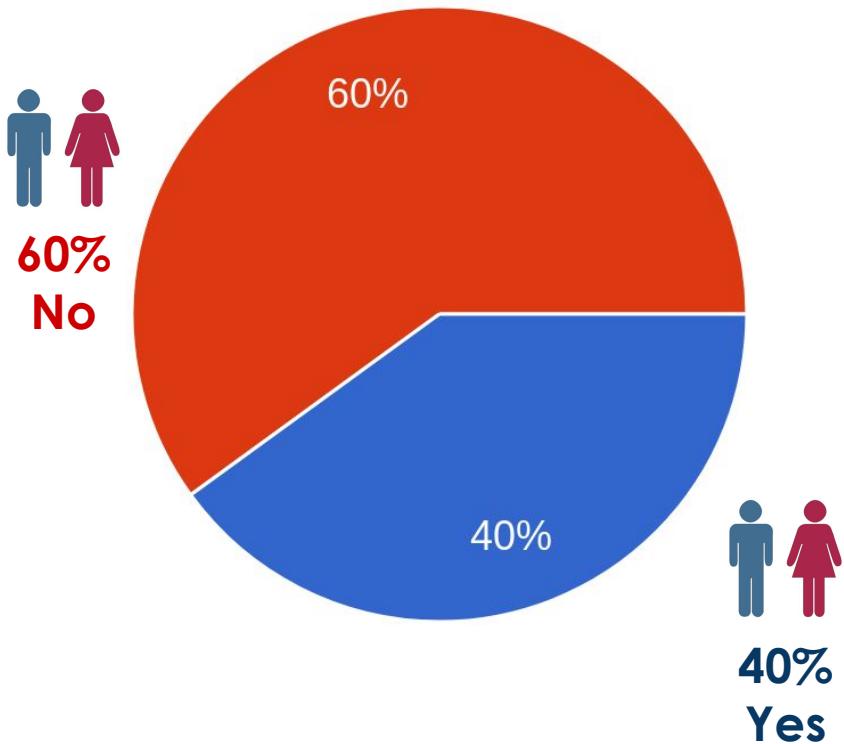
Each December 16th people in **Garagoa** celebrate the end of the year with a **postmortem** ceremony called **the dead of sadness.**

GARAGOA
COLOMBIA

Have you written a Postmortem?

60%

40%

60%
No

40%
Yes

18 Yes

27 No
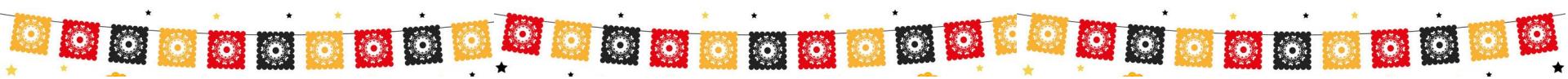
A survey of
45 Software Engineers,
showed that **Postmortems
are not a common
practice.**

Agenda

💀 About **postmortems.**

💀 Why don't we write postmortems?

💀 Blameful culture.

💀 **Chaos Engineering**.

💀 **Chaos Gamedays**.

💀 Automating Gamedays & Postmortems.

💀 **Gaveta** by Digital Labs.
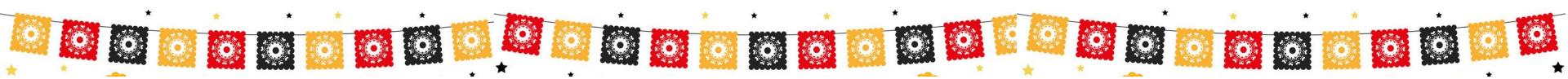
A **postmortem** is a written record of an incident, its impact, the actions taken to mitigate it, the root cause, and the follow-up actions to prevent the incident.

A **postmortem** is an artifact with a detailed description of exactly what went wrong in an incident.

**What went wrong, and how do we learn from it?**

Feb 10, 2017 - GitLab

Postmortem of database outage of January 31

Postmortem on the database outage of January 31 2017 with the lessons we learned.

← Back to company

On January 31st 2017, we experienced a major service outage for one of our products, the online service GitLab.com. The outage was caused by an accidental removal of data from our primary database server.

Bartek Polanczyk
@SzybkiSasza

Follow

#aws #s3 is down! @cloudfront is down! The world is burning!
1:09 PM - 28 Feb 2017
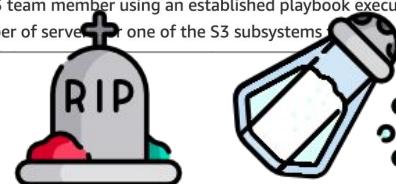
198 151

aws

# Summary of the Amazon S3 Service Disruption in the Northern Virginia (US-EAST-1) Region
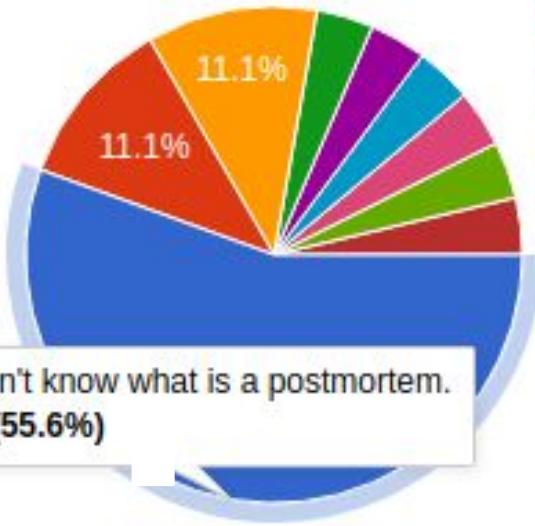
We'd like to give you some additional information about the service disruption that occurred in the Northern Virginia (US-EAST-1) Region on the morning of February 28th, 2017. The Amazon Simple Storage Service (S3) team was debugging an issue causing the S3 billing system to progress more slowly than expected. At 9:37AM PST, an authorized S3 team member using an established playbook executed a command which was intended to remove a small number of servers for one of the S3 subsystems that is used by the S3 billing process.

If the **Postmortems** are
a best practice, why don't we
do it?

Because

I don't know what is a postmortem.

I am a developer, so I don't have to participate in these activities.

My team has not been involved in an i...

I haven't found the time to do it

My team had never been created doc...

Por el momento no he generado ning...

The teams I've been involved don't ha...

The culture in my current job not invol...

I haven't formalize the incident manag...

I don't know what is a postmortem.
15 (55.6%)

Adrian Hornsby @adhorn · 8h

I have some serious questions for you Twitter. What currently prevents the wide adoption of chaos engineering in your organization?
#SRE #AWS #Serverless #Software #ChaosEngineeing

Hard to start. ✓	24%
Enough chaos in prod!	44%
Cultural issues (blame)	22%
Other - comment	10%

337 votes · 15 hours left

💬 17   🔁 29   ♡ 25   ⬆

John Allspaw @allspaw

If you believe "blameless postmortem" means "no accountability" - you haven't been listening.
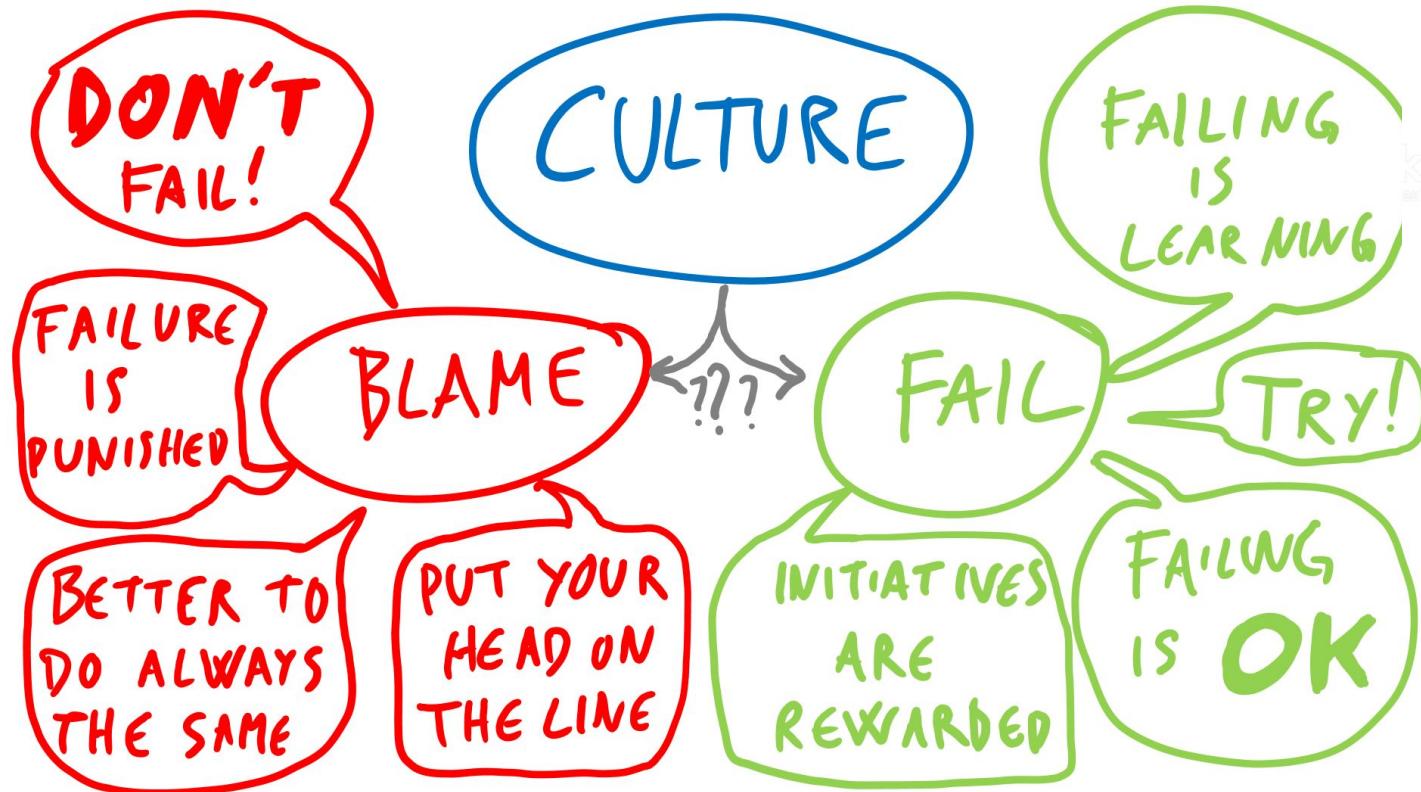
10:25 AM · Nov 17, 2017 · Twitter for iPhone

DEFINED

The 5 Whys: A method of getting to the root cause of a problem by repeatedly asking "why?" Each "why" question is a response to the previous answer. The technique was developed by Sakichi Toyoda, founder of Toyota.

The Mercury News

THURSDAY, OCTOBER 21, 2019

Outages. Fires. Who's to blame?

SHAME

RIP

Root cause: dylanfour@ ignored the automation setup and ran the cluster turnup logic manually, which triggered an existing bug.

How to change a blameful culture?

Chaos Engineering

It is the discipline of **experimenting** failures in **production** in order to **reveal** their weakness and to **build confidence** in their **resilience** capability.

https://principlesofchaos.org/
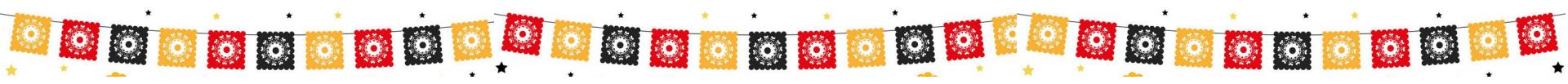
Chaos GameDays

They are **events** to conduct **chaos** experiments against a system to validate or invalidate **hypothesis** about a system's **resiliency**.

They are an ideal way to ease into **Chaos Engineering.**

Brian Lee, Jason Doffing

Chaos Gamedays

Master of Disaster declares **start of incident** and attack!!!

First On-Call member sees, triages, and tries to mitigate the impact.
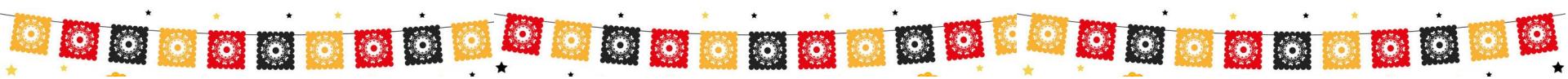
Team understands, analyzes and solves.

Postmortem
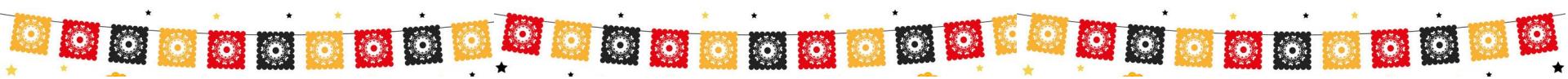
What does it mean
in the practice?

Planning a Gameday

- 💀 Create an **agenda**.
- 💀 Define **users** & **roles**.
- 💀 Send communications**.**
- 💀 Design an experiment.
- 💀 Provision:
  - ○ HW/SW
  - ○ Chaos attackers
  - ○ Observability

Planning a Gameday

Resource Downloads

Get prepared for your GameDay.

⤓ GameDay Agenda	⤓ GameDay Execution Template	⤓ GameDay Recording Template
⤓ Objec... Communication...	⤓ Plan... ...f...	⤓ P... ...Checklist
⤓ GameDay fir... next ste...		

ChaosIQ

Chaos Engineering Game Day Plan Template

Summary

| Lead Chaos Engineer		Date	
Chaos Start Time		Ticket #	
Chaos End Time		Chaos Duration	
Time to Detect Chaos (via observability tools)		Time to Automatically Remediate/ Mitigate Chaos	

Chaos Day Countdown: 90 Days

90 Days may sound like a long time, but getting everyone involved requires giving a large amount of lead time. This is especially important if you'd like your CTO and CEO to attend.

Digital Labs

She is Gaveta!

Master of Disaster

First On Call

Team

Gaveta

Go   Terraform

Planner

Communicator

Chaos Environment Provisioner

Chaos Runner

aws

DATADOG   New Relic

Documentor
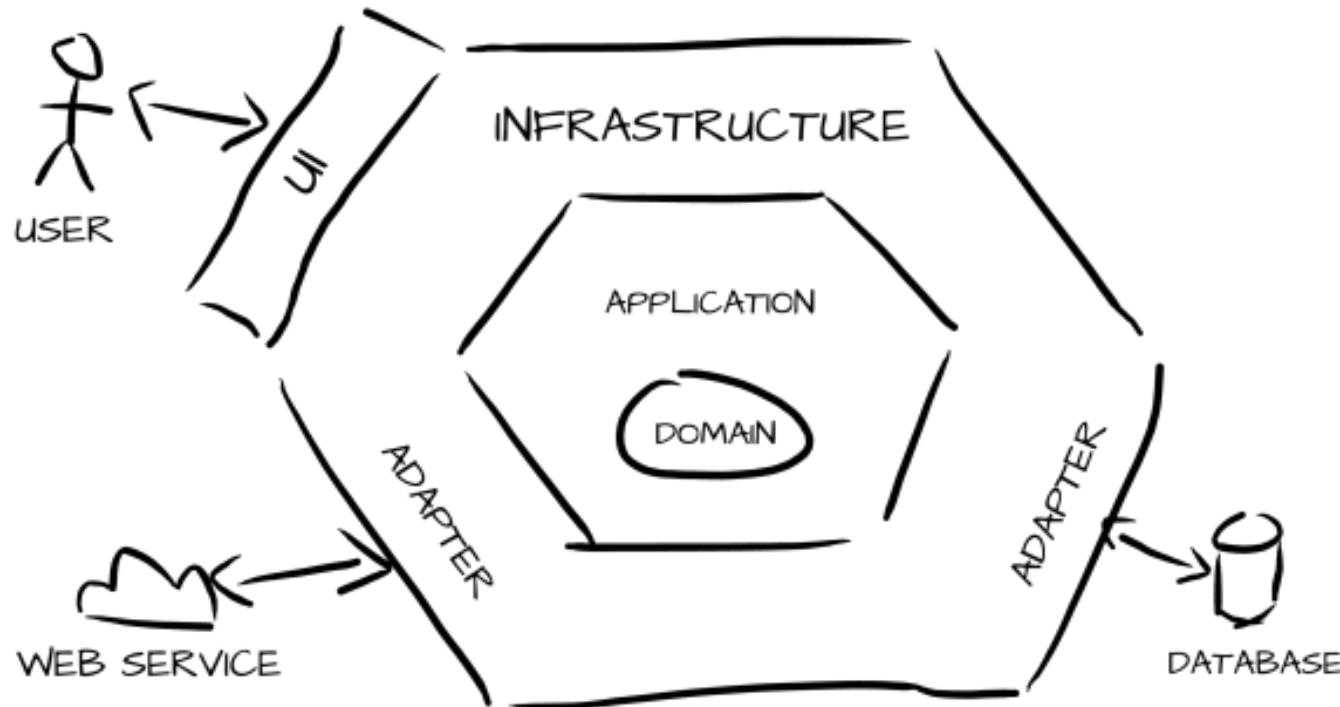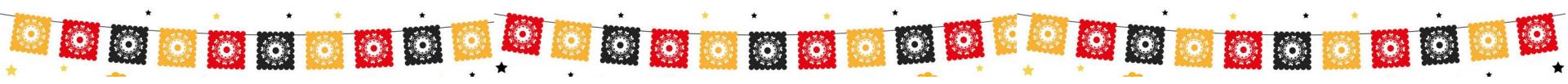
Backlogger
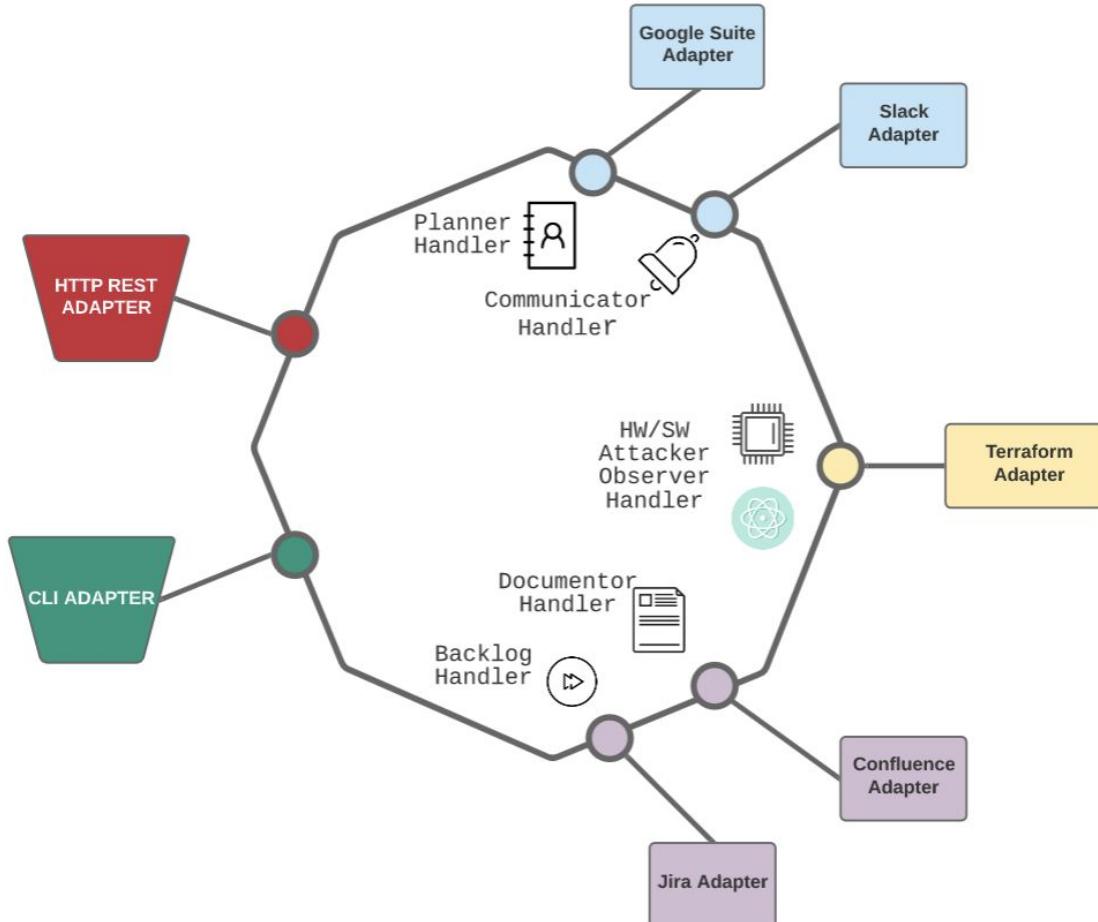
Your System
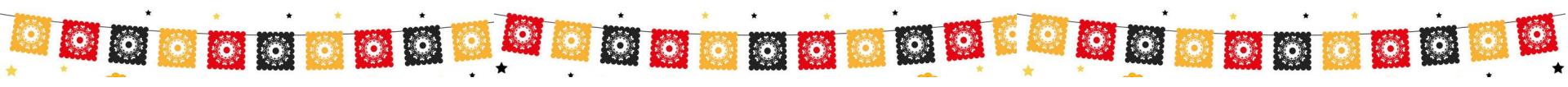
Gaveta uses a hexagonal architecture

Gaveta uses a hexagonal architecture

GAVETA

Gameday June 10th 2020

🔒 Generate Tokens

PLAN 80% 📋	**PROVISION** 50%	**RUN** June 10th 💣	**POSTMORTEM** Actions 📗

Postmortem

Overview: Include a short sentence or two summarizing the contributing factors.

**What Happened:** Include a short description of what happened.
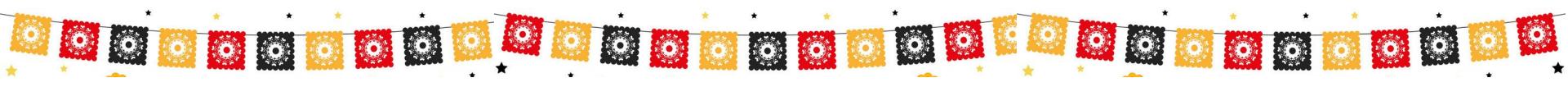
Contributing Factors: Include a description of any conditions that contributed to the issue. If there were any actions taken that exacerbated the issue, also include them here with the intention of learning from any mistakes made during the resolution process.

Resolution: Include a description of what solved the problem. If there was a temporary fix in place, describe that along with the long-term solution.

Impact: Be very specific here, include exact numbers.

Actions

Id	Action
1101	Implement an aditional resilience stratey in the communication between transactions to payments.
1102	Check the load balancer in aws, because the main cluster is not having a proper behaviour.
1103	Document the values of the hystrix configuration.

GAMEDAYS

⊕ Create

☰ List

CONFIGURATION

🔧 Adapters

Incident_ 2020-01-01

Created by Yury Niño Roa
Last updated Jan 07, 2020

`impact` `high`

`duration` `10 min`

> "On the afternoon of November 21st, we got a problem regarding credentials.

👷 **Owner:** Carlos Cortés

🤲 **Impact**

Item	Value
Incident Duration	10 min
Resolution Time	100 min
Accounts Affected	62
Users Affected	15

🎛️ What Happened
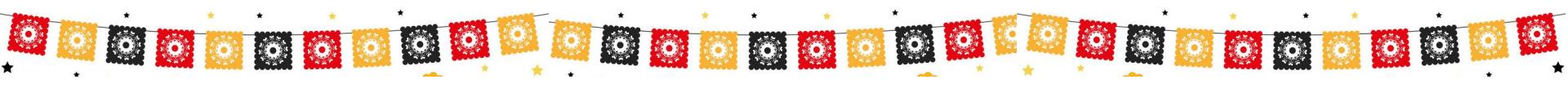
ChaosIQ — O'REILLY

Chaos Engineering Game Day Plan Template
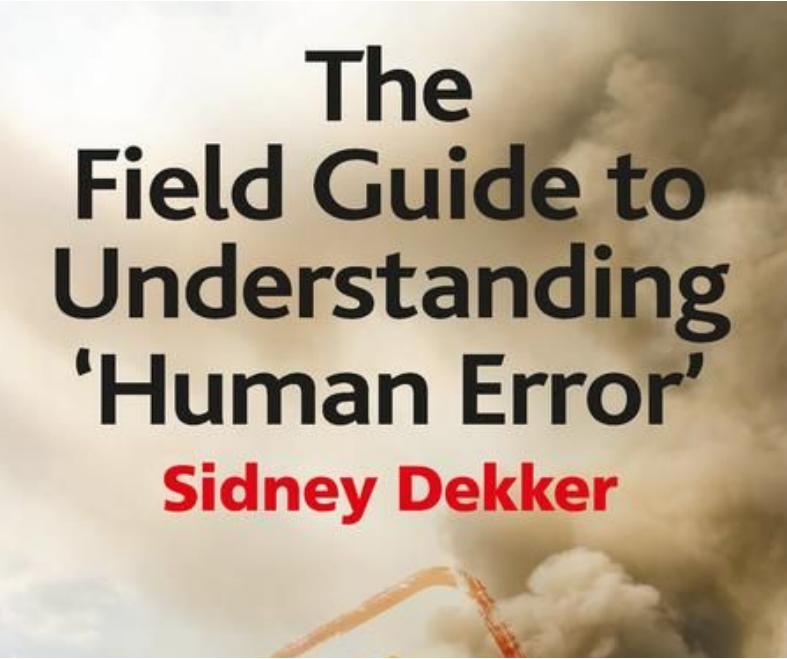
Summary

| Lead Chaos Engineer		Date	
Chaos Start Time		Ticket #	
Chaos End Time		Chaos Duration	
Time to Detect Chaos (via observability tools)		Time to Automatically Remediate/ Mitigate Chaos	
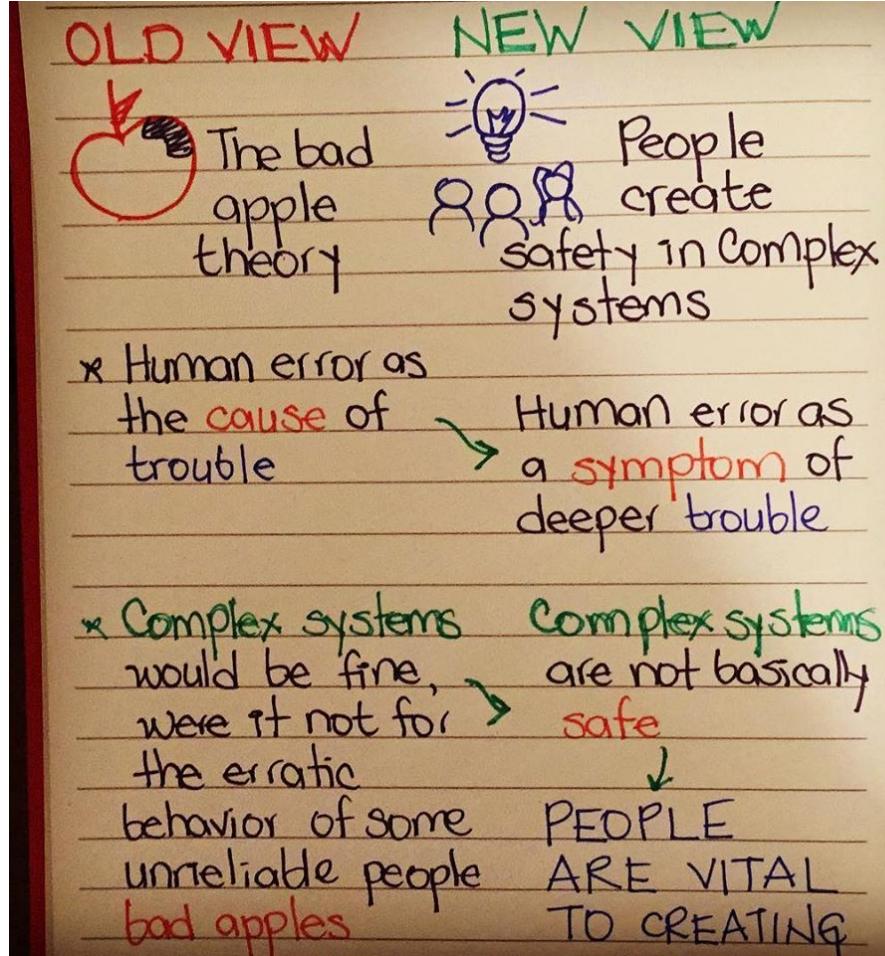
# The Field Guide to Understanding 'Human Error'

Sidney Dekker

*The best way to promote a **postmortem** culture is adopting a new view, a view focused in the **syntoms**, no in the **causes** ...*

OLD VIEW | NEW VIEW

The bad apple theory

People create safety in complex systems

x Human error as the cause of trouble → Human error as a symptom of deeper trouble

x Complex systems would be fine, were it not for the erratic behavior of some unreliable people bad apples → Complex systems are not basically safe ↓ PEOPLE ARE VITAL TO CREATING

Thanks for coming!!!

@yurynino