

# Caffeinated Quantum Squadron Sprint ① Stakeholder Update

February 16, 2023

# INTRODUCTION (1/2)

**TWITTER** is one of the fastest growing social media sites, allowing people to connect and share their thoughts with the world through a series of *tweets*.

With millions of active users, Twitter has inevitably come across the pernicious and insidious nature of hate speech on its platform.

**TWITTER INVITES CHALLENGE. CHALLENGE INCITES CONFLICT. CONFLICT BREEDS CATASTROPHE.**

## Demographics

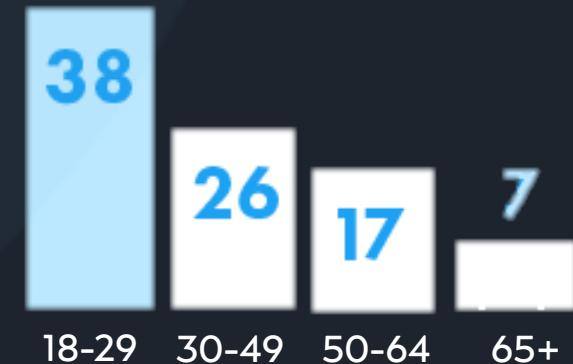


66%



34%

Age (%)



# INTRODUCTION (2/2)

KEVIN HART

## wasted



Kevin Hart  @KevinHart4real · 1 Jun 2009

I just saw the biggest **gay** guy ever! This **dude** looked like hulk hogan with heels on! I can't lie I got scared!!!!!!

15

25

22



## THE KEVIN HART SITUATION

In December of 2018, Kevin Hart was detested when a series of his old homophobic tweets resurfaced.

As a result, the comedian publicly stepped down from opportunities like hosting the 2019 Academy Awards.

## WHAT CAN WE DO?

### MEAN TWEET DETECTOR

Although Twitter allows all voices to be heard, there are no effective measures to monitor hate speech.

We intend to build a **MEAN TWEET DETECTION pipeline as a combative solution.**

# PIPELINE PROGRESSION

1

## DOCKERIZE AIRFLOW

Eliminate all kinds of environment setup errors.

2

## DATA COLLECTION

Gather data relating to our objective of classifying mean tweets (store in CSV).

3

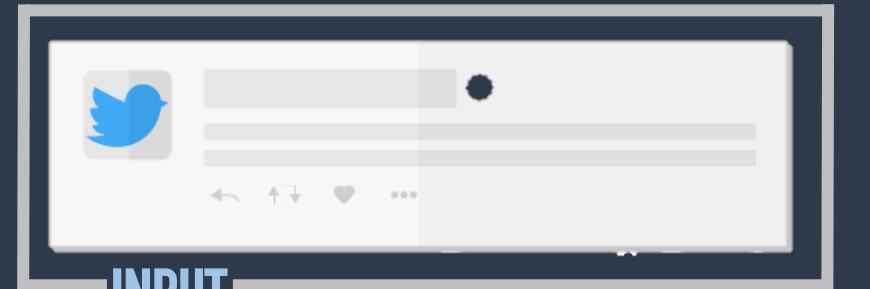
## DATA PREPARATION

Cleanse the data to limit errors and inaccuracies for efficient data analysis.

4

## MODEL TRAINING

Validate the performance of the detection model, so it can be deployed for end-users.



INPUT

MEAN TWEET DETECTOR  
PIPELINE



MEAN OR NOT MEAN?!

# AIRFLOW IN A DOCKER CONTAINER

**Creating docker containers for Apache Airflow allows for the creation of a consistent isolated environment that can be deployed across multiple workstations.**

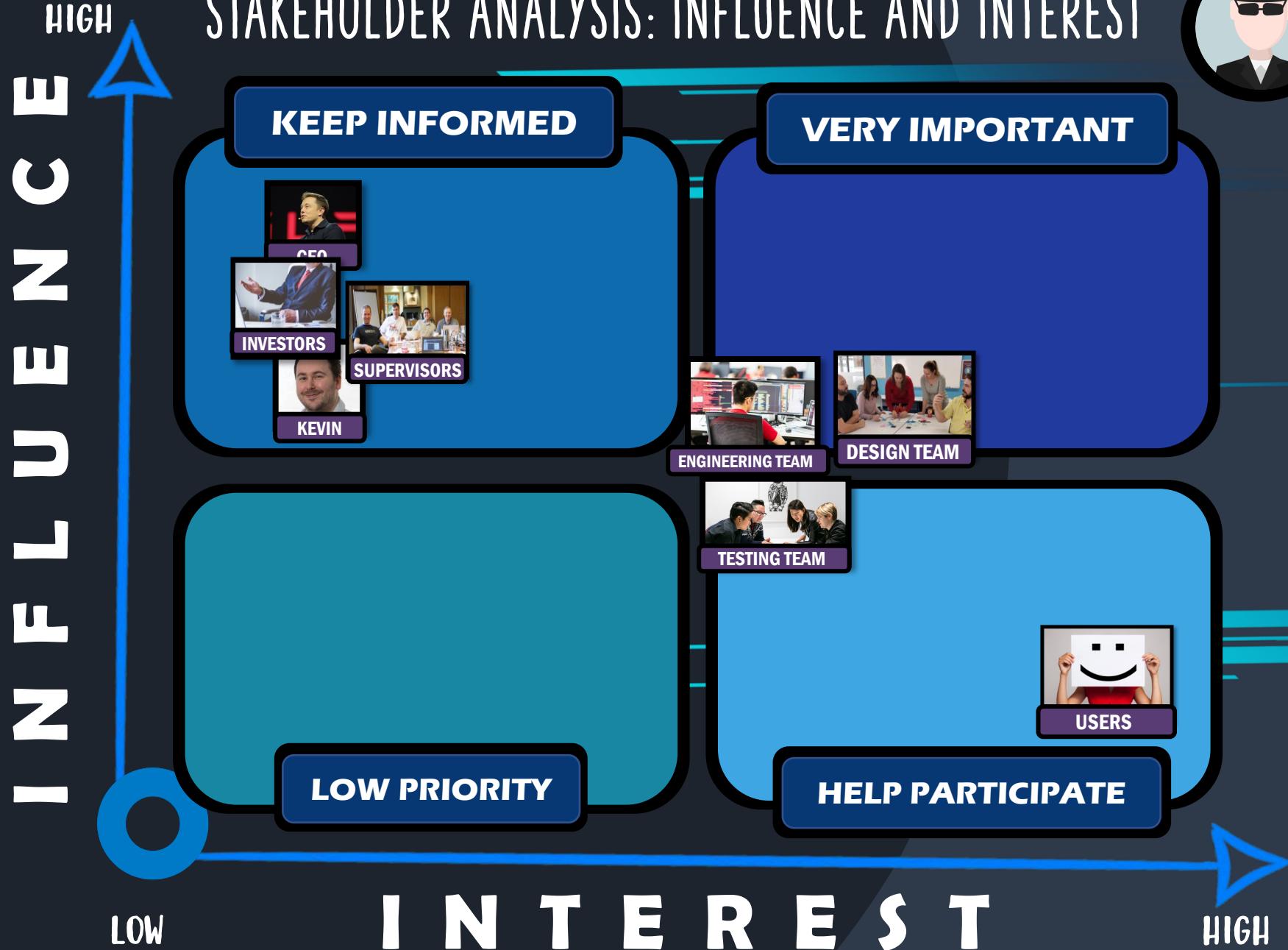
**Creation of containers ultimately induce faster deployments and increased cost-effectiveness.**

## OUR DOCKER-COMPOSE.YML:

<https://github.com/tyeborg/mean-tweet-pipeline>



# STAKEHOLDER ANALYSIS: INFLUENCE AND INTEREST



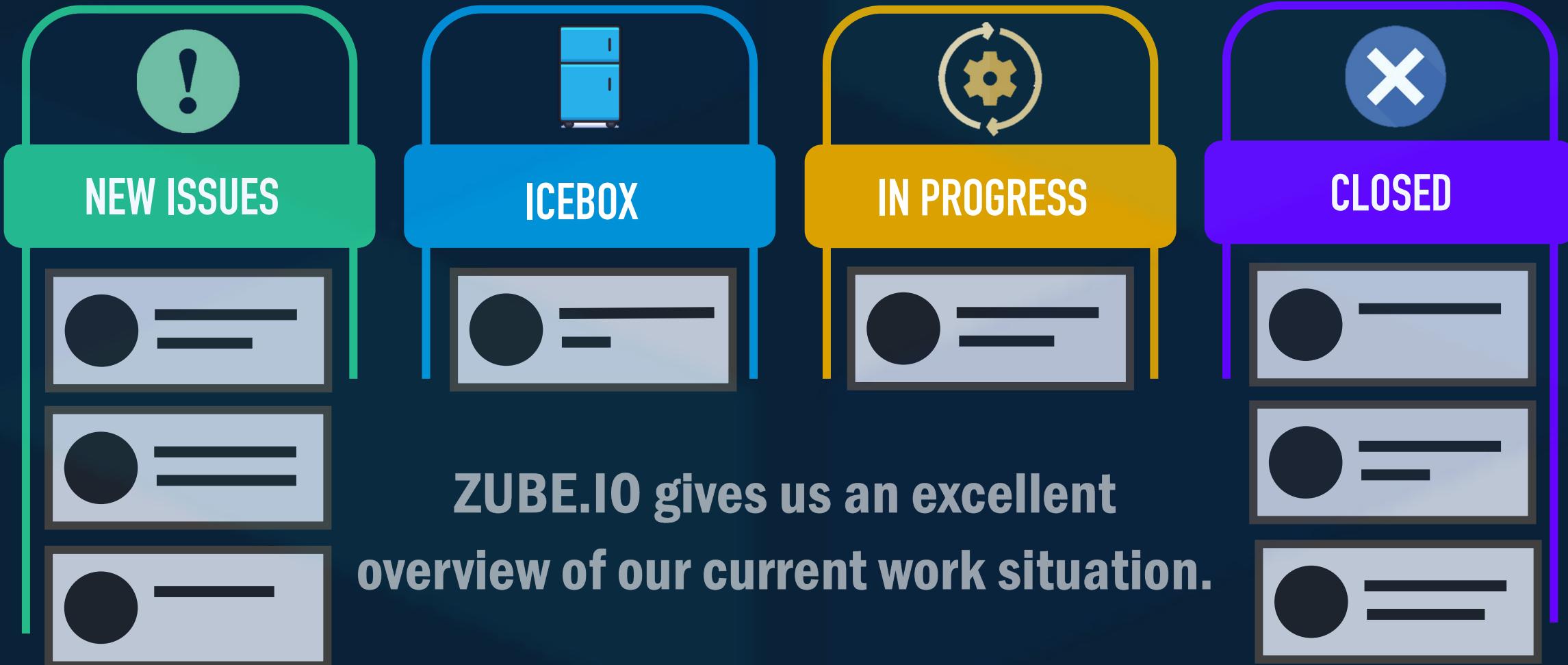


# THE PRODUCT BACKLOG



ID	USER STORY	PRIORITY	SPRINT	STATUS
01	As <b>DATA ENGINEERS</b> , we MUST set up the project workflow, so we can ensure that future developments will run as smooth as possible.	MUST HAVE	SPRINT 01	-----
02	As <b>DATA ENGINEERS</b> , we MUST set up task management as well as evaluate requirements in user stories and use cases, so we can each perform individual tasks optimally and efficiently in an organized manner.	MUST HAVE	SPRINT 01	-----
03	As <b>DATA ENGINEERS</b> , we MUST create a development environment using Docker, so we can start building DAGs in Apache Airflow.	MUST HAVE	SPRINT 01	-----
04	As <b>DATA ENGINEERS</b> , we MUST be able to consume a data source that contains information about the tweet sentiment, so that we can predict the severity of a particular tweet.	MUST HAVE	SPRINT 02	-----
05	As <b>DATA ENGINEERS</b> , we SHOULD clean our data, so we can ensure that we have data of the highest quality for increased data analysis.	SHOULD HAVE	SPRINT 02	-----
06	As <b>DATA ENGINEERS</b> , we MUST utilize prepared data to train a model to find patterns and make predictions, so we can distinguish the severity of a tweet based off its text .	MUST HAVE	SPRINT 02	-----
07	As <b>DATA ENGINEERS</b> , we MUST embed appropriate tests to our pipeline, so we can ensure our data goes through the proper quality checks, before being deployed to the end data consumers.	MUST HAVE	SPRINT 03	-----
08	As <b>DATA ENGINEERS</b> , we SHOULD perform hyperparameter tuning upon our model, so we can increase its performance.	SHOULD HAVE	SPRINT 03	-----
09	As <b>DATA ENGINEERS</b> , we SHOULD effectively deploy our pipeline into production, so we can start utilizing it for its intended use of detecting severity of tweets.	SHOULD HAVE	SPRINT 04	-----

# TASK MANAGEMENT WITH ZUBE.IO



# CODE OF CONDUCT

## The Caffeinated Quantum Squadron Code of Conduct

To establish and maintain the trust, respect, and collaboration within the Caffeinated Quantum Squadron, I do hereby pledge to abide by the policies of the team throughout the completion of the project. I give my word of honor that I will observe the following policies:

### I. BEHAVIOR - All members within the team are to be treated with respect, kindness, and equality.

As a team, we all depend on each other to produce the best work we can achieve. However, an environment where people feel uncomfortable or threatened is not a productive or creative one. Therefore, each member is expected to uphold the responsibility to encourage and support one another during the timeframe of the intended task.

### II. ORGANIZATION - Attend scheduled meetings regularly and on time.

Meetings will occur in the 2<sup>nd</sup> Floor Conference Room inside the Sir David Bell Building (University of Roehampton Campus) on Thursday's from 2:00 PM to 4:00 PM each week. Members are required to meet these arrangements unless they are hindered by an unexpected/unfortunate predicament.

### III. INTEGRITY - Maintain honesty and clear communication within the environment.

Communication is an integral component to project management; hence all members are encouraged to inform one another about any discrepancies in order to resolve workplace conflict in an efficient manner. Additionally, in no circumstance shall a member within the team collude with any members from outside parties about information concerning the project. If there is beyond reasonable doubt of suspected plagiarism, extreme measures will be enacted for breaching academic honesty, which may result in a failing grade.

In an event that a member violates any rules detailed within the pillars of Behavior, Organization, and Integrity, disciplinary action will be enforced. To ensure the presence of fairness and governance when dealing with disobedience, each member will be allocated a total of three warnings.

- 1<sup>st</sup> Warning - Minor problems will be dealt with by informal advice and counselling.
- 2<sup>nd</sup> Warning - Warn that action under the 3<sup>rd</sup> Warning will be considered if there is no satisfactory improvement. Additionally dealt with during an informal discussion.
- 3<sup>rd</sup> Warning - Penalties from the Module Delivery Team will be implemented.

\* Google Docs will serve as the platform to record a counselling session if a warning is issued.

\* The type of warning may vary depending on the severity of the situation deemed applicable by a group majority (excluding the member on trial).

Examples of scenarios that justify the issuance of a warning:

- Failure to attend meetings (especially without a valid reason or excuse)
- Inability to complete assigned tasks within a given deadline
- Exhibition of poor, ineffective communication amongst teammates
- Plagiarism

Signing this document exemplifies a thorough review of the Caffeinated Quantum Squadron Code of Conduct and an adherence to its policies of Behavior, Organization, and Integrity. Please be aware that any violations of the Caffeinated Quantum Squadron Code of Conduct may lead to disciplinary action, up to and including the involvement of the Module Delivery Team, as determined to be appropriate by a group majority.

  
(Yaksh Bhatt)

  
(Sidharth Jain)

  
(Barbara Kobak)

  
(Tyler Supersad)

The code describes how we are expected to behave towards members of our group. Failing to comply with the code may be misconduct.

The code is reinforced by three important pillars:



BEHAVIOR



ORGANIZATION



INTEGRITY



YAKSH BHATT



SIDHARTH JAIN



BARBARA KOBAK



TYLER SUPERSAD

**THANK YOU FOR  
LISTENING!**