

Sound Impact Analysis

Giulia Pulvirenti

Analisi in tempo reale dell'impatto
dei suoni virali su YT Shorts

</> TAP Project

ViSTORM



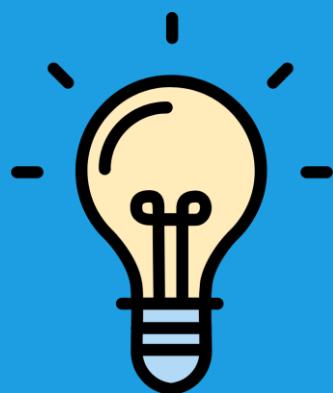


AGENDA

4 steps

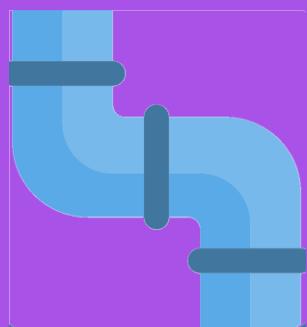
1

L'idea e la
nascita del
progetto



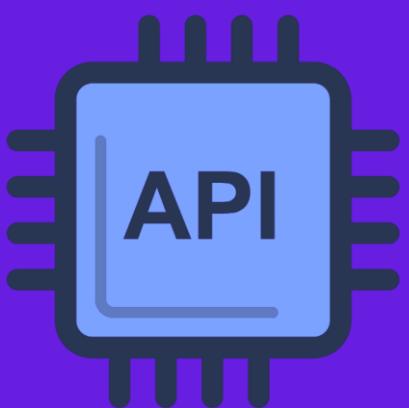
2

Pipeline



3

Tecnologie
scelte ed
utilizzate



4

Funzionamento
tecnico



**IL SOGNO
DELLA
VIRALITÀ'**



Cosa rende un
contenuto virale?





FAIL!

Carichi il tuo primo
video, ma non ottieni
visualizzazioni.

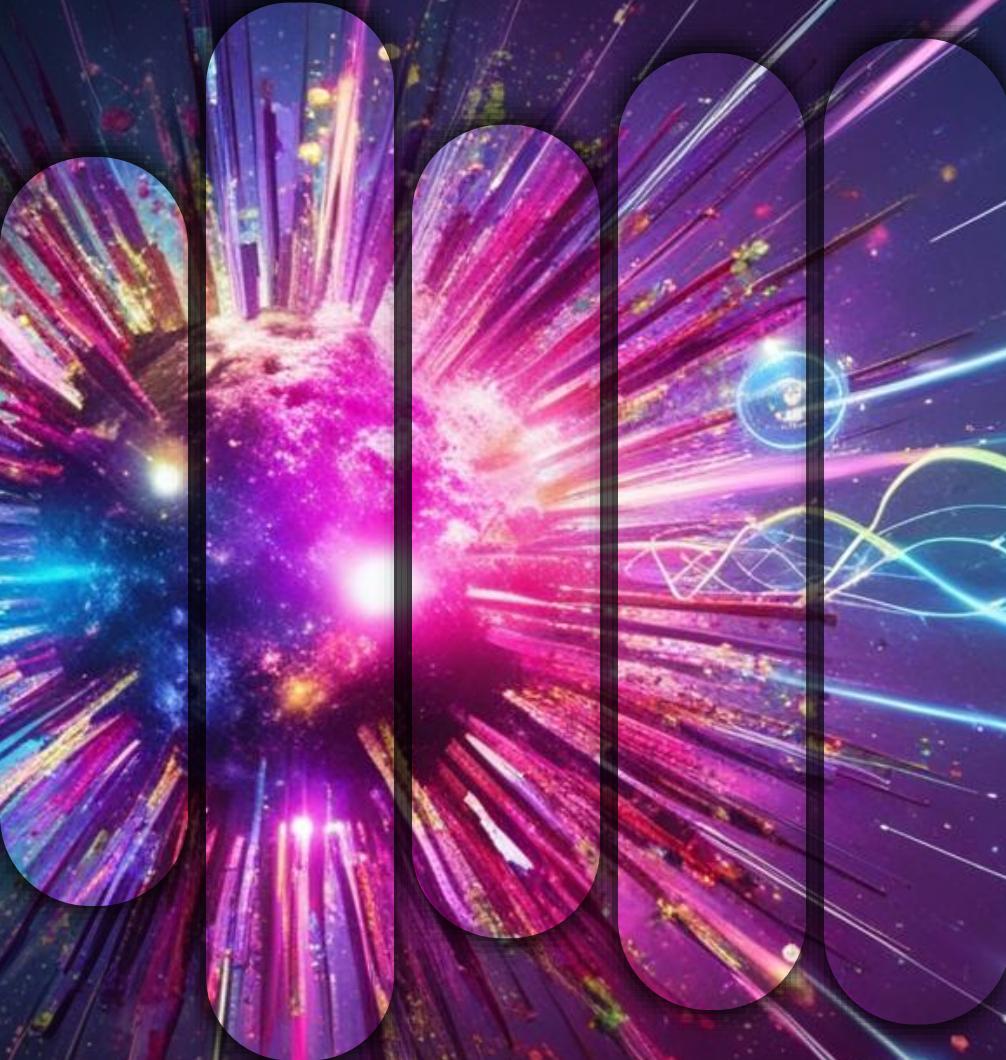
È frustrante, vero?





SUCCESS!

Poi vedi altri video che esplodono con milioni di visualizzazioni...



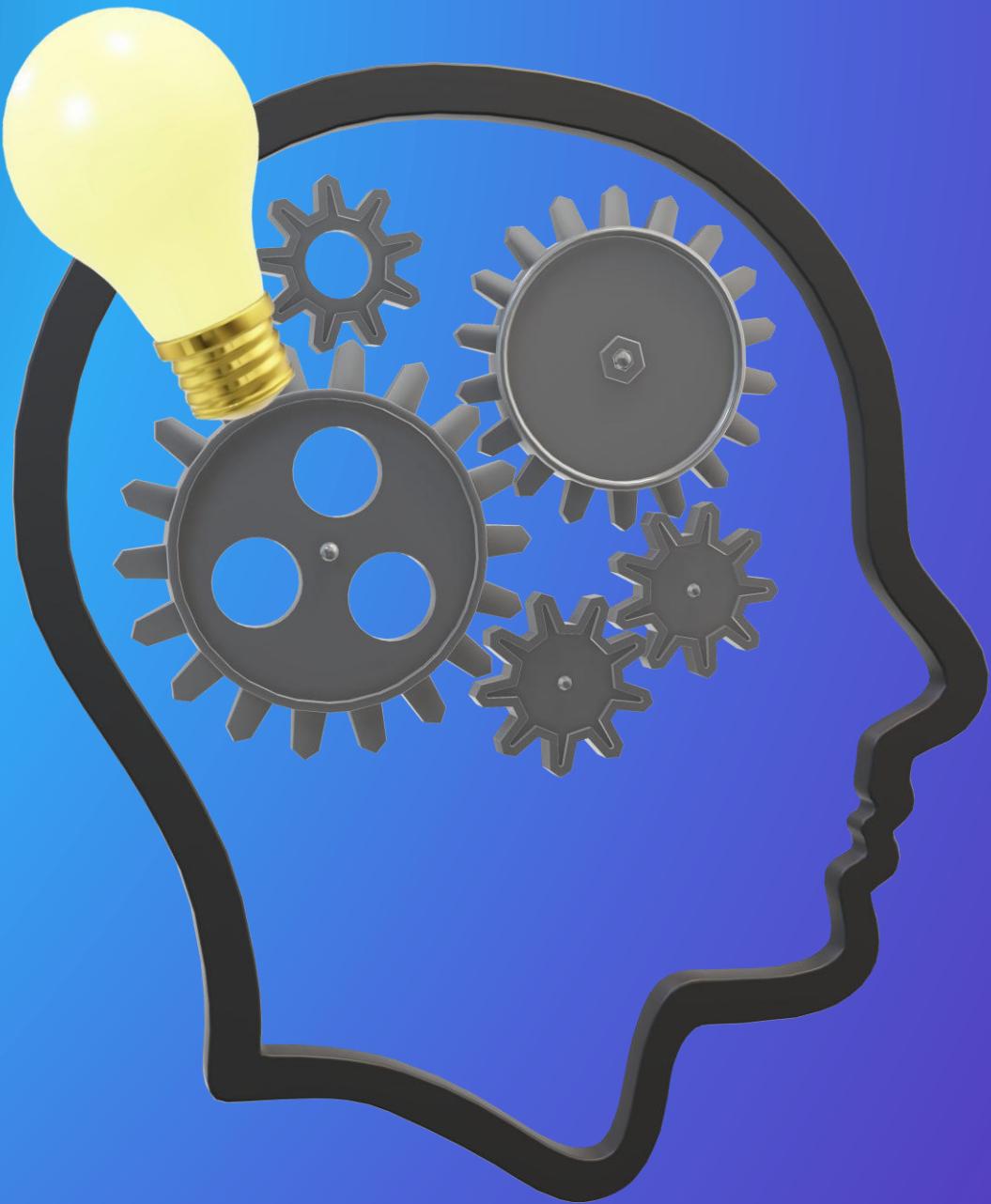


**LA VIRALITÀ
HA UN SUONO?**

NASCE



NASCE L'IDEA





HO DECISO DI ANALIZZARE
IN CHE MODO I SUONI
INFLUENZINO LA
POPOLARITÀ DI UN VIDEO.



L'OBIETTIVO?

IDENTIFICARE I FATTORI CHIAVE,
COME L'USO DI SUONI, CHE
TRASFORMANO UN SEMPLICE
VIDEO IN UN FENOMENO.



Nasce Vistorm:
dall'intuizione alla realtà



IL NOME

Vistorm, il nome del mio progetto, nasce dall'unione di «*visualization*» e «*storm*», per rappresentare un flusso inarrestabile di dati e informazioni.



LO SCOPO

Esaminare l'impatto delle canzoni virali sugli *YouTube Shorts*, analizzando i dati in tempo reale per prevedere la popolarità e la visibilità dei video in base a tendenze musicali e caratteristiche chiave.



PRODOTTO FINALE

Fornire uno strumento innovativo per monitorare il proprio canale, con funzionalità di predizione per pianificare i contenuti e strumenti utili per analizzare le performance. *Creators* e *marketers* possono così ottimizzare visibilità, coinvolgimento e successo, sfruttando anche l'impatto della musica virale.

PIPELINE

01
Data
Ingestion

02
Data
Streaming

03
Data
Processing

04
Machine
Learning

05
Indexing
Storage

06
Data
Visualization

Data Ingestion

Data Streaming

Data Processing

Machine Learning

Indexing/Storage

Data Visualization

Deployment

Logstash

Connette le API di
YouTube alla
pipeline.



Apache Kafka

Gestisce il flusso continuo

O

Data Ingestion

Data Streaming

Data Processing

Machine Learning

Indexing/Storage

Data Visualization

Deployment

Apache Kafka

Gestisce il flusso continuo
di dati.



Apache Spark

Consente analisi real-time

Data Ingestion

Data Streaming

Data Processing

Machine Learning

Indexing/Storage

Data Visualization

Deployment

Apache Spark
Consente analisi real-time.



Spark MLlib

Data Ingestion

Data Streaming

Data Processing

Machine Learning

Indexing/Storage

Data Visualization

Deployment

Spark MLlib

Libreria per creare modelli di apprendimento automatico.



ElasticSearch

MLLIB

Data Ingestion

Data Streaming

Data Processing

Machine Learning

Indexing/Storage

Data Visualization

Deployment

ElasticSearch

Archivia e recupera dati in modo rapido.



Kibana

Strumento di

Data Ingestion

Data Streaming

Data Processing

Machine Learning

Indexing/Storage

Data Visualization

Deployment

Kibana

Strumento di visualizzazione per creare dashboard interattive e aggiornate in tempo reale.



Deployment

Automation and orchestration



in tempo reale.

Data Ingestion

Data Streaming

Data Processing

Machine Learning

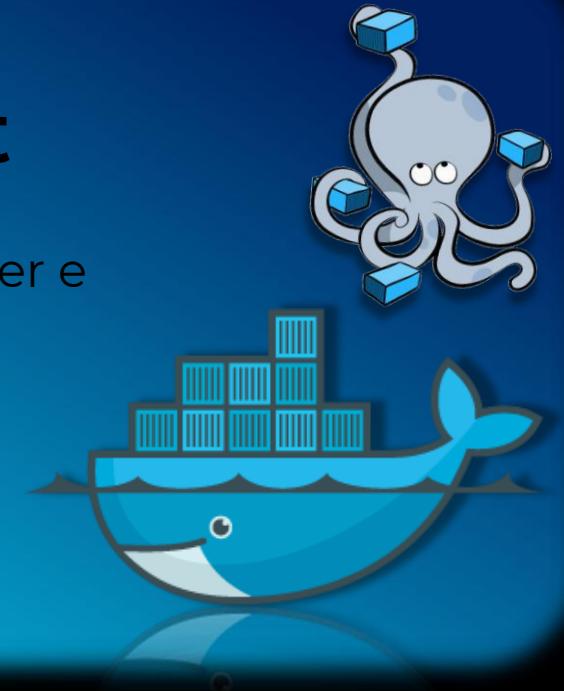
Indexing/Storage

Data Visualization

Deployment

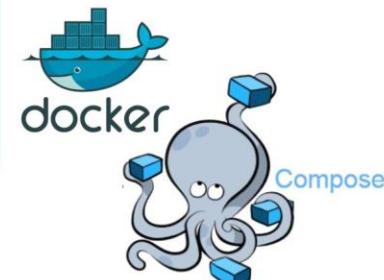
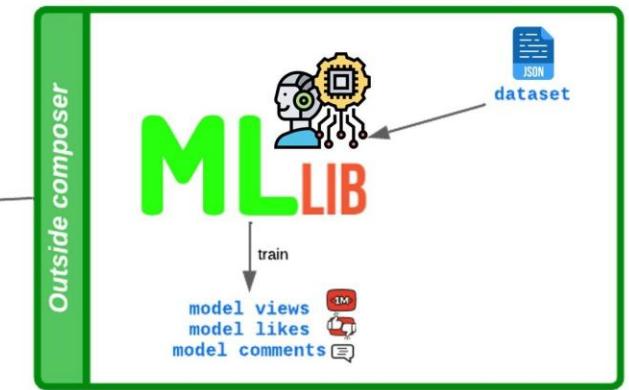
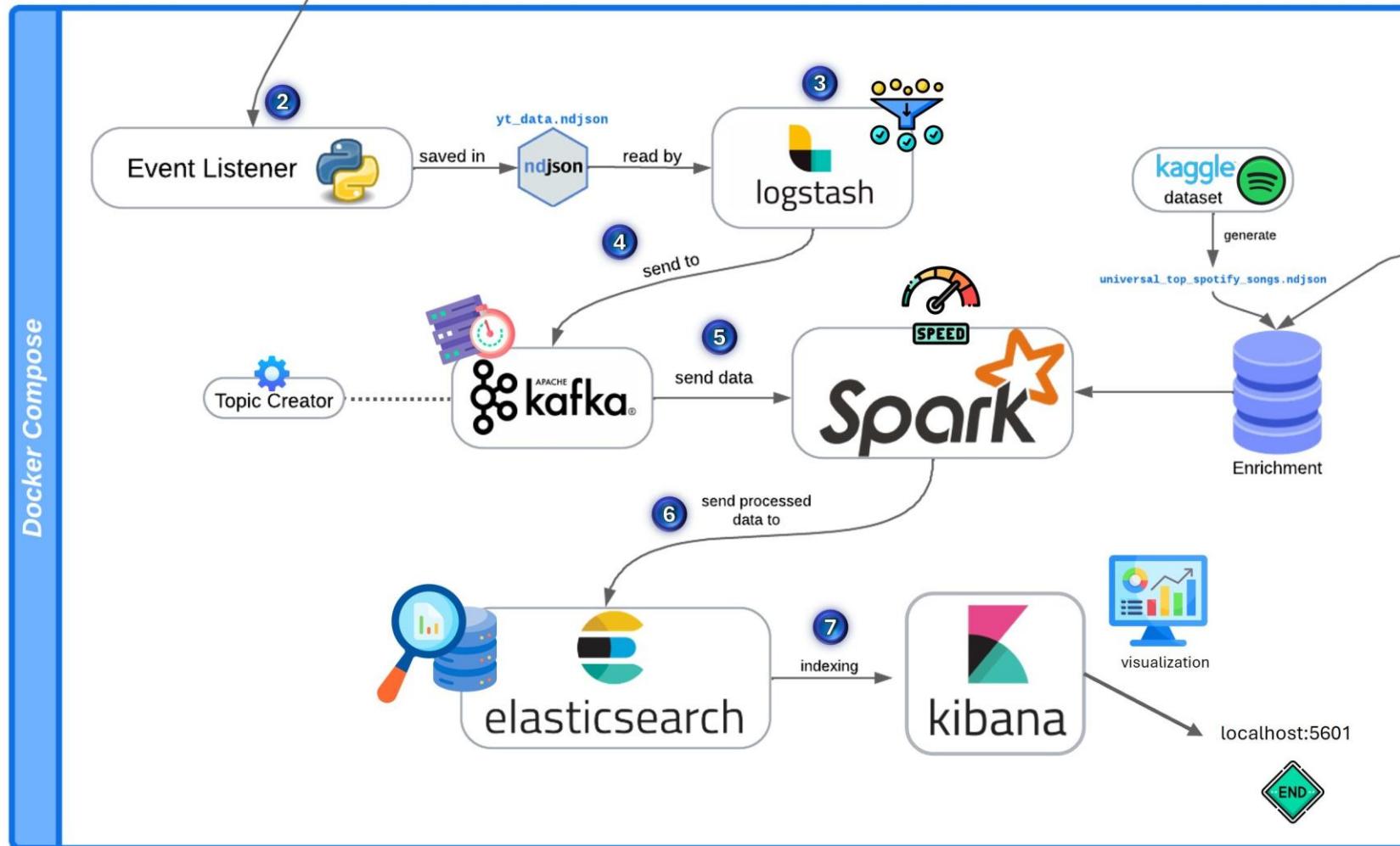
Deployment

Automatizzazione del deployment con Docker e Docker Compose





1 START



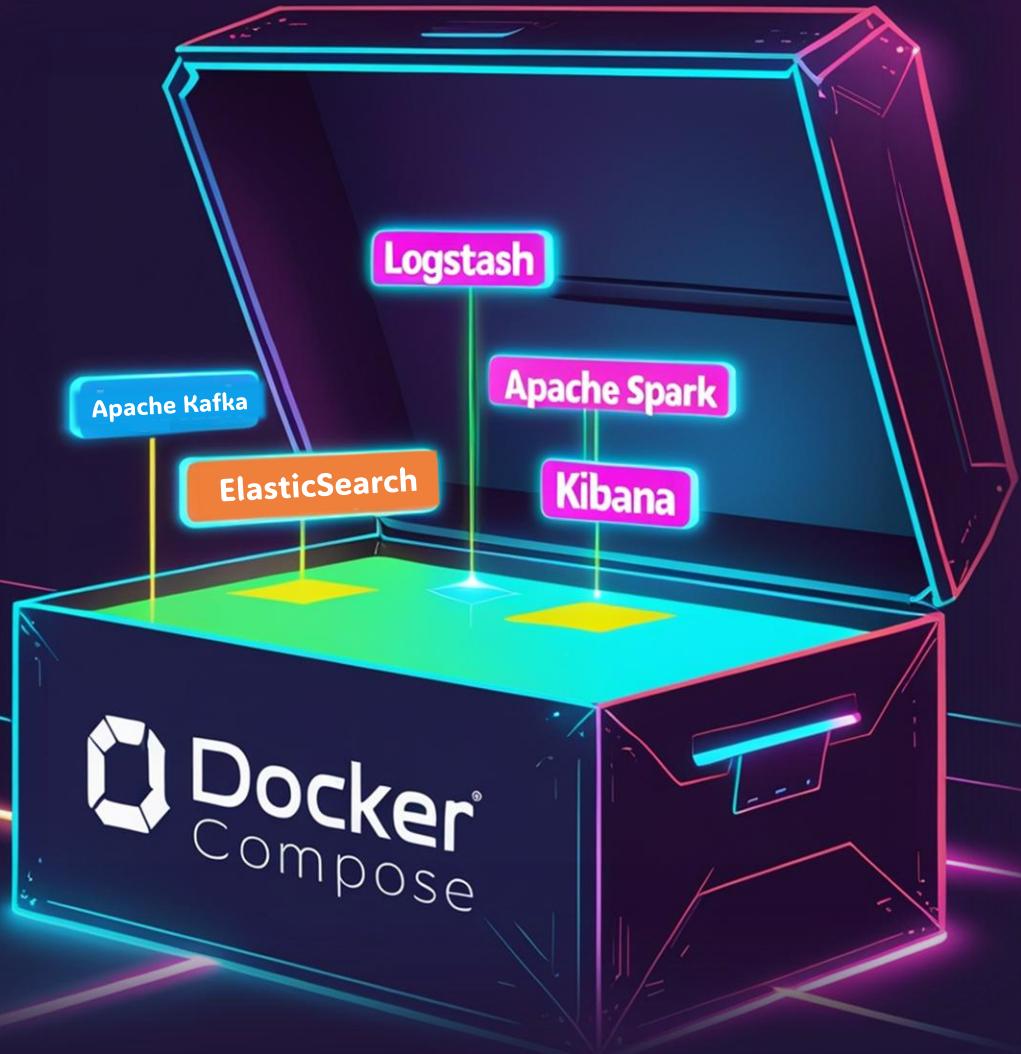
Data lineage



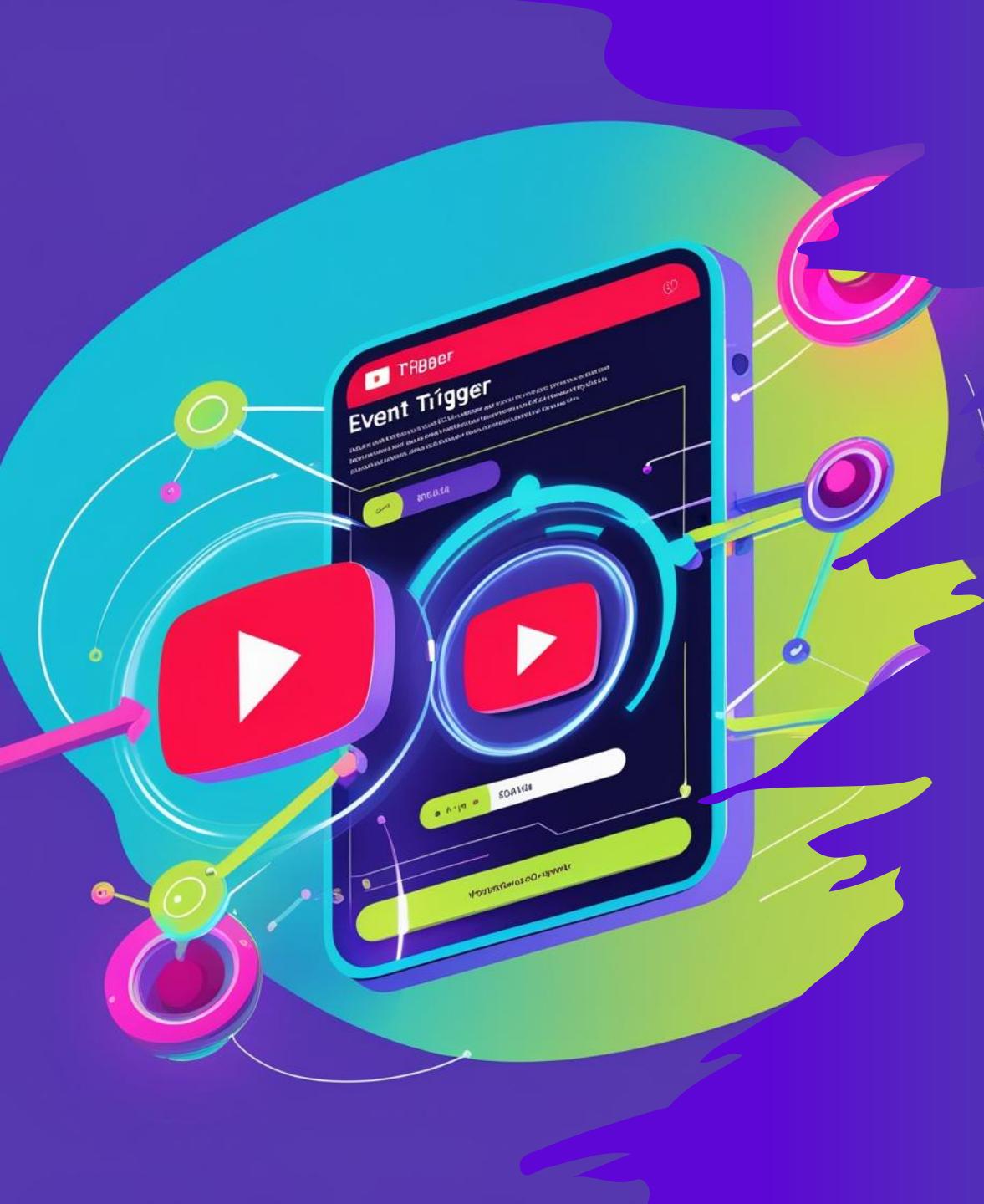
Il data lineage mappa l'intero ciclo di vita dei dati, tracciandone l'origine, le trasformazioni e le destinazioni, connettendo processi come ingestion, streaming, processing, indexing e visualization per garantire trasparenza e controllo.



Come le ho utilizzate?



FASE DI PRE-INGESTION



■ Caricamento e Monitoraggio

1. Carico uno YouTube Short su un canale creato appositamente per attivare un event trigger.
2. Uno script Python fa da listener monitora in tempo reale i nuovi caricamenti.

■ Controllo della Durata

- Lo script verifica se la durata del video è ≤ 60 secondi.
- I video che superano questa durata vengono ignorati perché non sono considerati Shorts.

■ Raccolta Dati e Analisi Audio

Lo script raccoglie i metadati del video.

■ Utilizza l'Audd API per analizzare la traccia audio:

1. Controlla la presenza di un suono.
2. Estrae e organizza informazioni come titolo, artista e altri dettagli musicali.

Genera un JSON strutturato, validato e trasformato, pronto per essere passato a Logstash per la data ingestion.



DATA INGESTION

- Logstash prende in input il JSON prodotto durante la pre-ingestion.
- Applica filtri per rimuovere dati inutili, mantenendo solo le informazioni rilevanti).
- Invia i dati filtrati e trasformati a un topic di Kafka, dove saranno pronti per essere elaborati dai componenti successivi della pipeline.

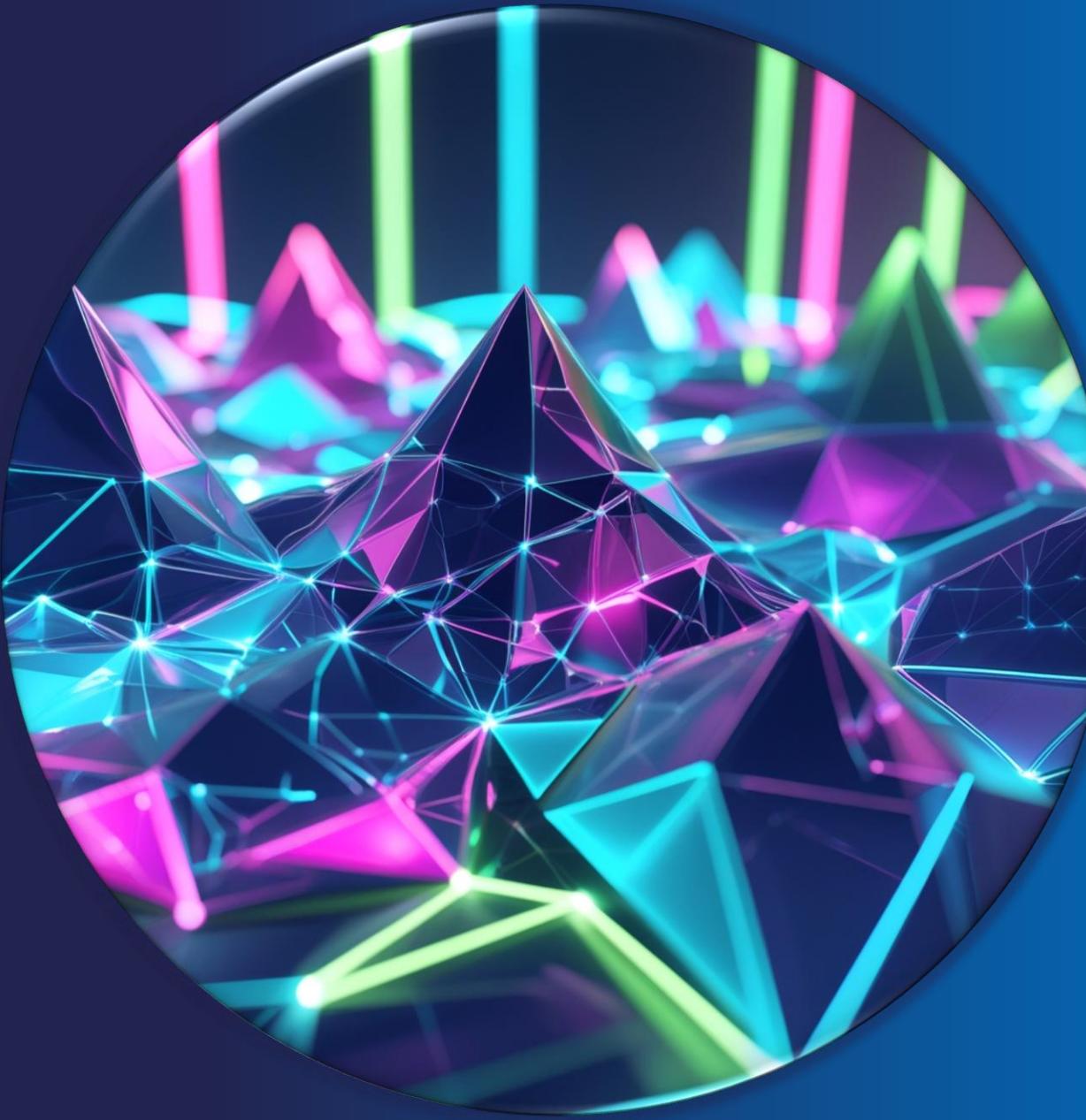


DATA STREAMING

Utilizzando Apache Kafka per la gestione del flusso continuo dei dati raccolti:

- Tutte le informazioni (metriche video, tracce musicali, dati virali) sono distribuite attraverso un *topic* dedicato.
- Questo approccio assicura un trasferimento dei dati in tempo reale, eliminando i *bottleneck*.





DATA PROCESSING

Utilizzando Apache Spark per processare i dati, con un modello di MLlib che verrà spiegato tra poco, i dati dei video sono stati integrati con i dati musicali, verificandone la corrispondenza con le tracce virali di Spotify.

Come risultato, ci sarà un set di dati pronti ad essere indicizzati.

MACHINE LEARNING

Il modello predittivo con Spark MLlib utilizza un **GBT Regressor** (*Gradient-Boosted Trees*).

Il modello analizza tre scenari:

1. **Senza musica**: stima l'impatto del video in assenza di contenuti musicali.
2. **Con musica non virale**: valuta l'influenza di una canzone non virale.
3. **Con musica virale**: valuta l'influenza di una canzone riconosciuta.



INDEXING/STORAGE

Elasticsearch indicizza i dati elaborati, garantendo:

- **Scalabilità** nella gestione delle informazioni.
- **Ricerche rapide** per analizzare i risultati predittivi.

I dati sono sempre disponibili per ulteriori analisi.



DATA VISUALIZATION

La dashboard interattiva creata con Kibana, presenta:

- Metriche video dettagliate.
- Confronti tra gli scenari predittivi.
- Analisi dell'impatto delle canzoni virali.

Questo rende immediata l'analisi di trend e insight.



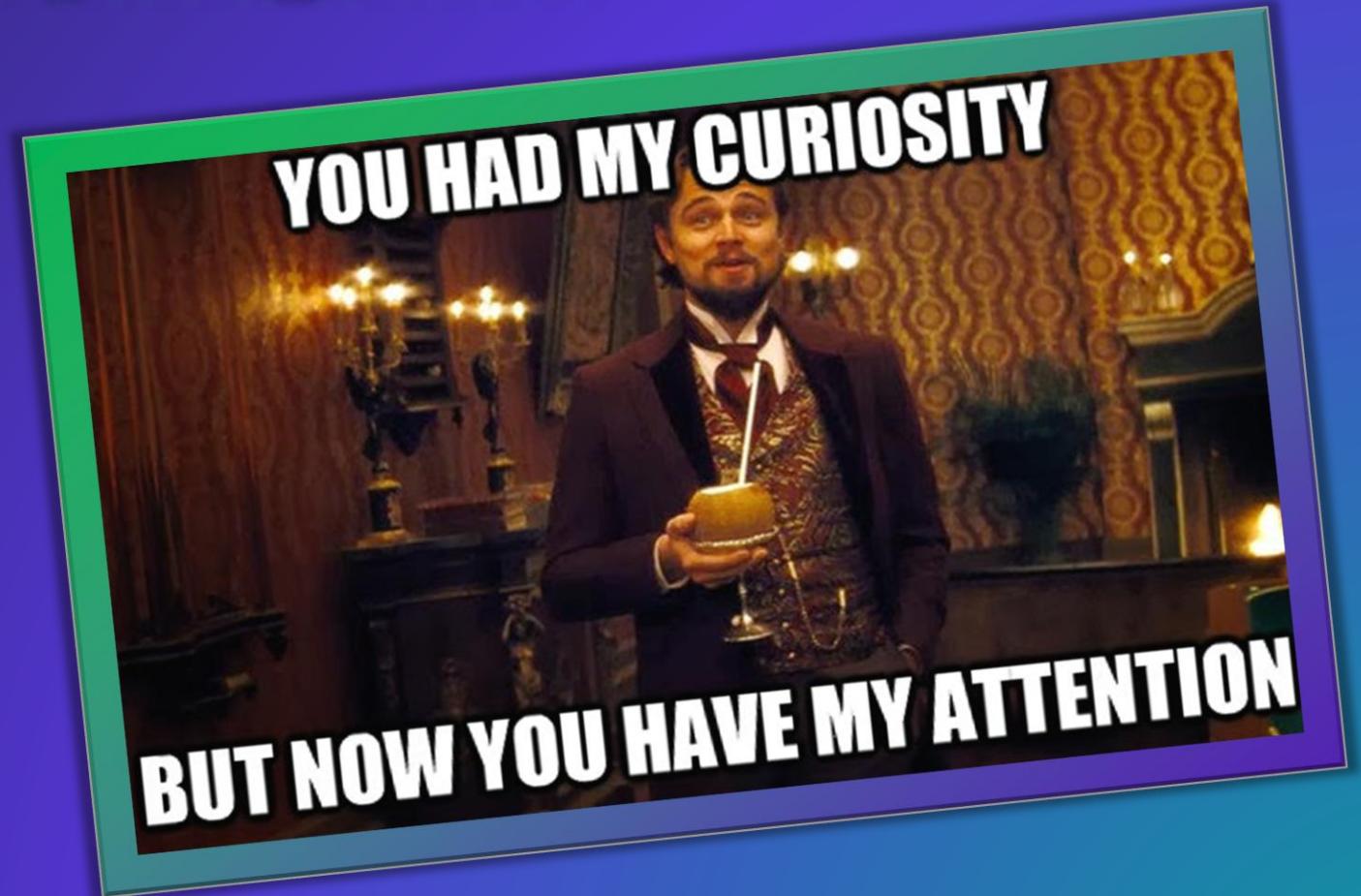


MA NON È STATO
così SEMPLICE...

Backstage del progetto

Ogni progetto ha il suo lato serio, ma anche momenti in cui tutto sembra andare storto (e poi ci ridiamo sopra).

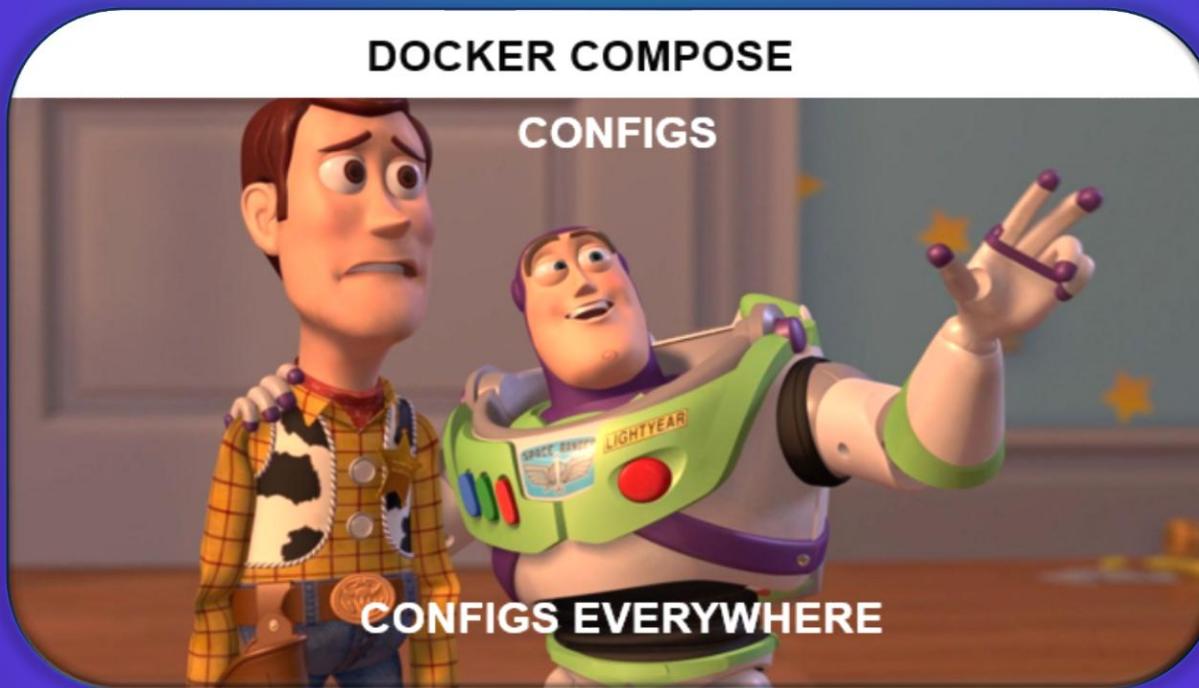
Ecco la mia raccolta di meme per raccontarvi cosa è successo davvero dietro le quinte...



Moments that almost broke me (but Didn't)



Moments that almost broke me (but Didn't)



Moments that almost broke me (but Didn't)



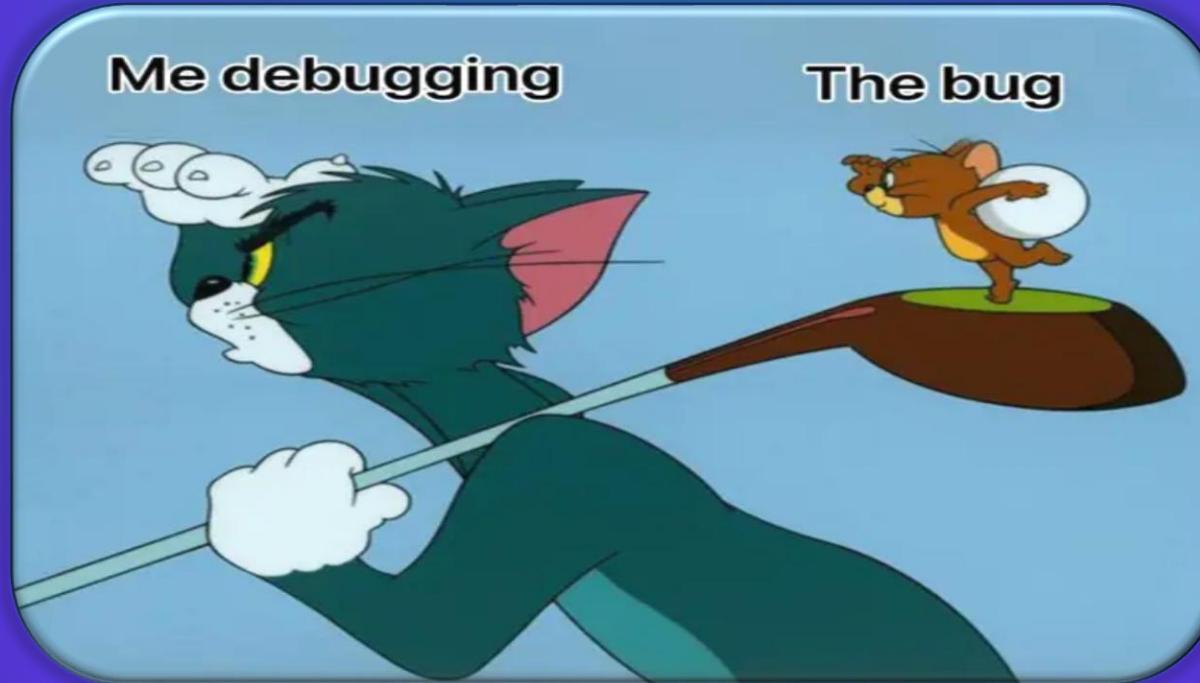
Moments that almost broke me (but Didn't)



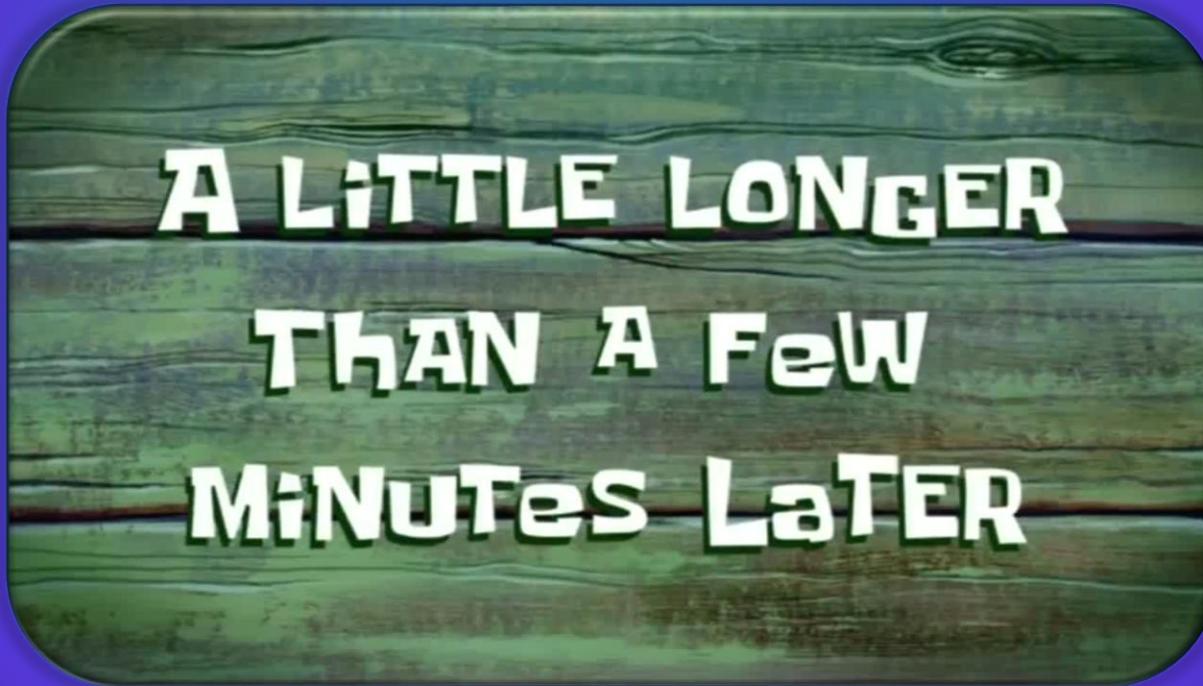
Moments that almost broke me (but Didn't)



Moments that almost broke me (but Didn't)



Moments that almost broke me (but Didn't)

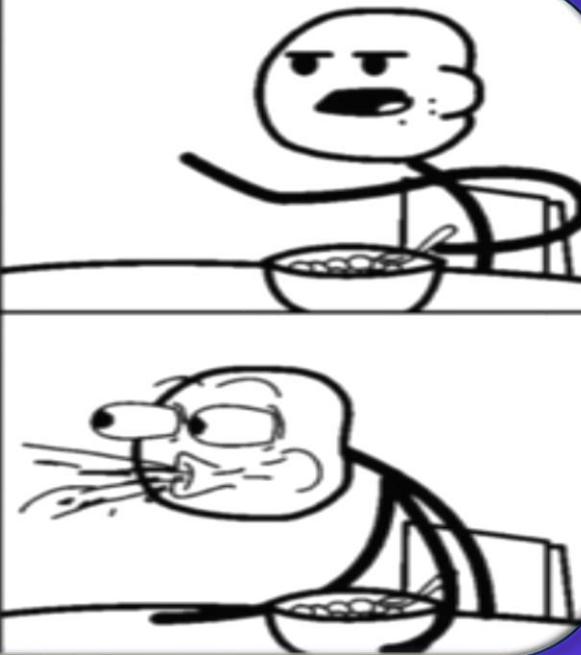


A LITTLE LONGER
THAN A FEW
MINUTES LATER



Moments that almost broke me (but Didn't)

**Quando ti aspetti
l'ennesimo errore**



**Ma improvvisamente
funziona**



**gnolissim
atnemssivvotdmi em**



Moments that almost broke me (but Didn't)



QUANDO VISTORM TI FA CAPIRE CHE IL SUONO
CHE HAI UTILIZZATO NON CONTRIBUISCE ALLA
TUA POPOLARITÀ



Github Profile



queenofh3ll20



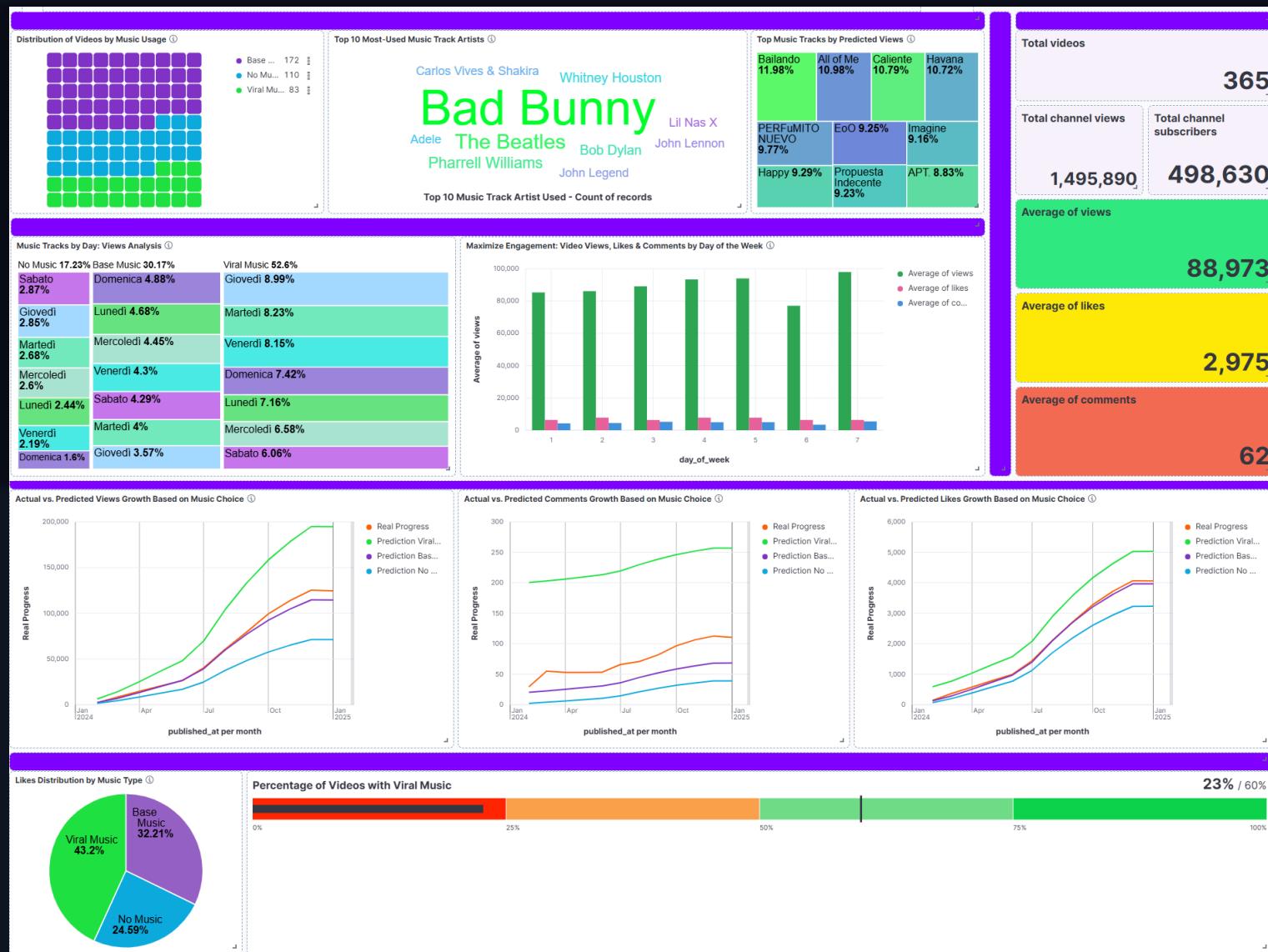
Project Repository



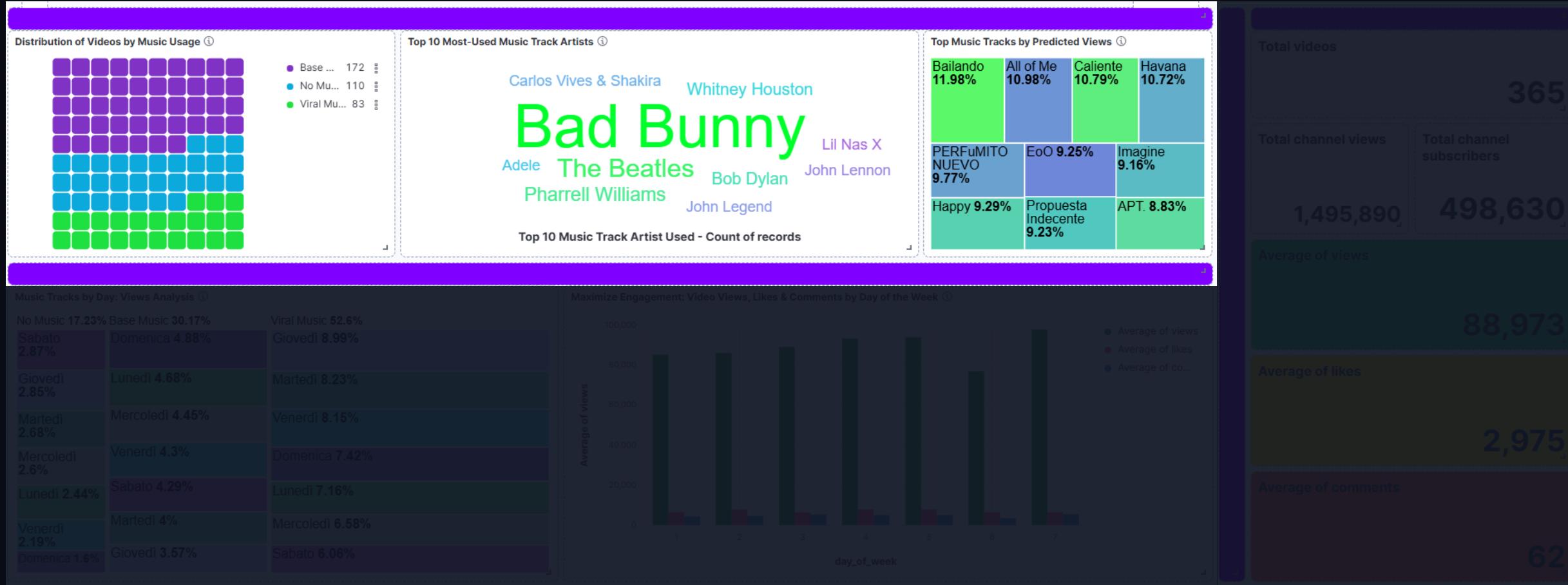
Vistorm



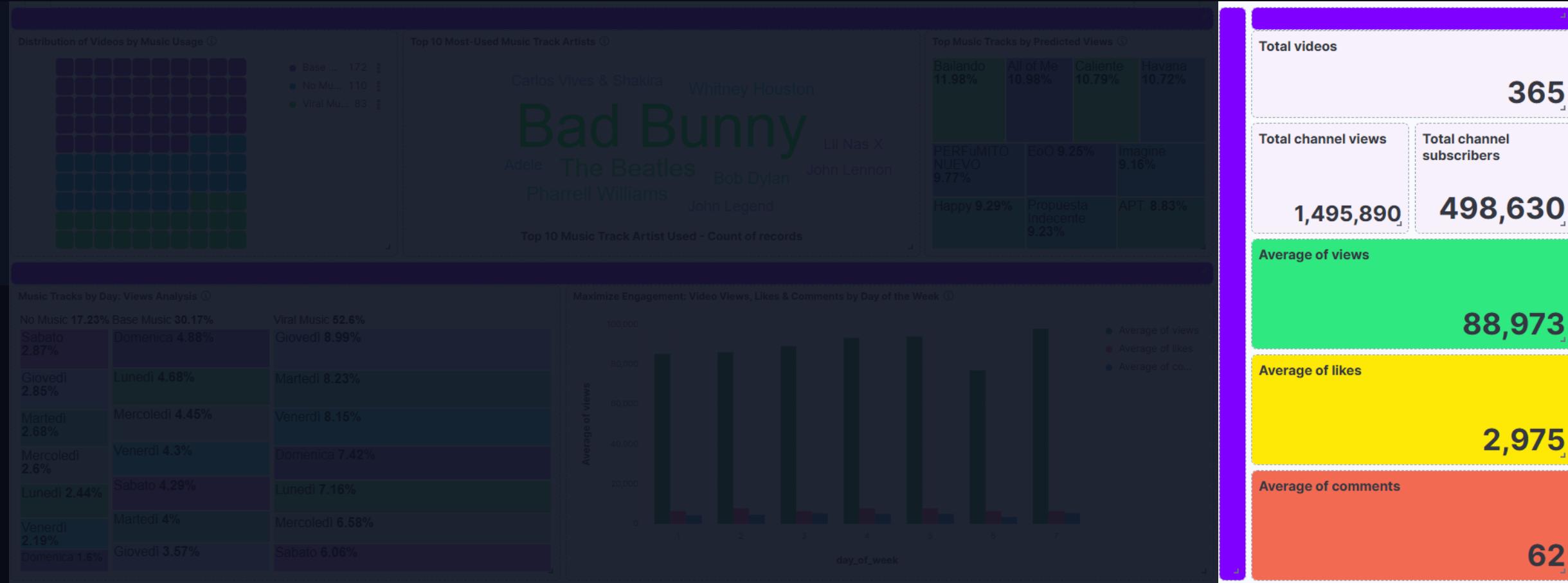
Dati a colpo d'occhio: la Dashboard che comunica



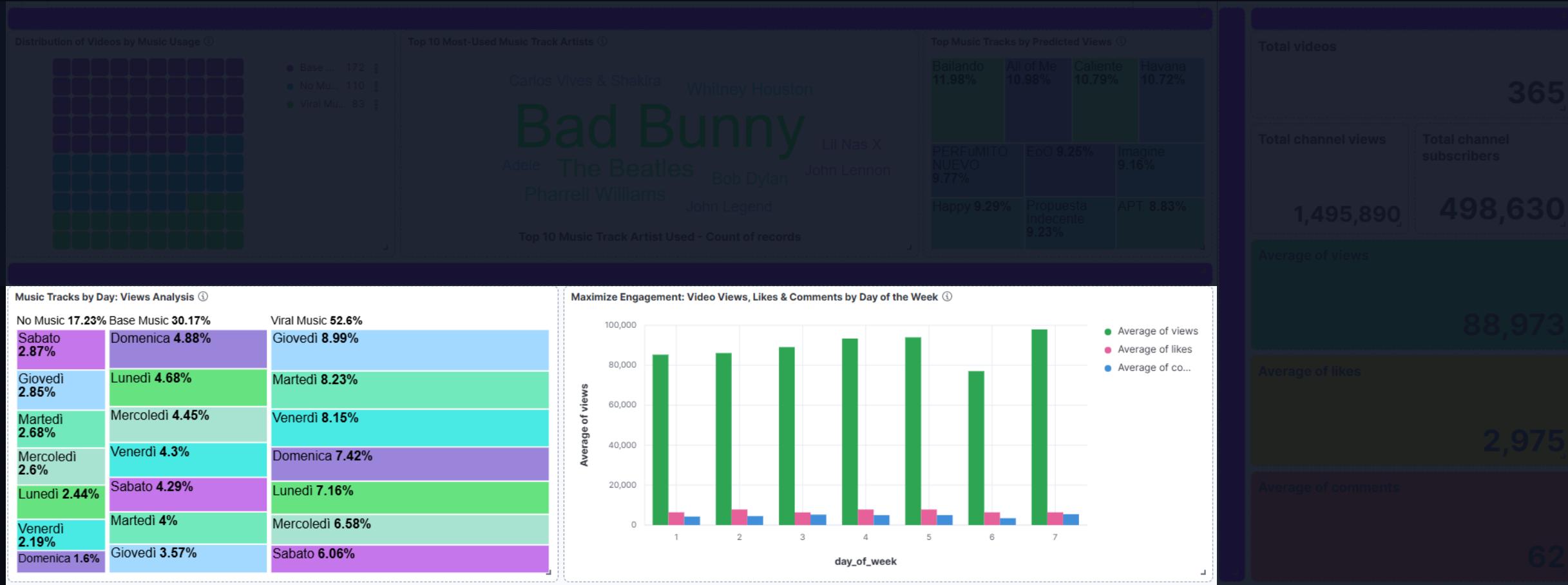
Section 1: Music Impact Insights



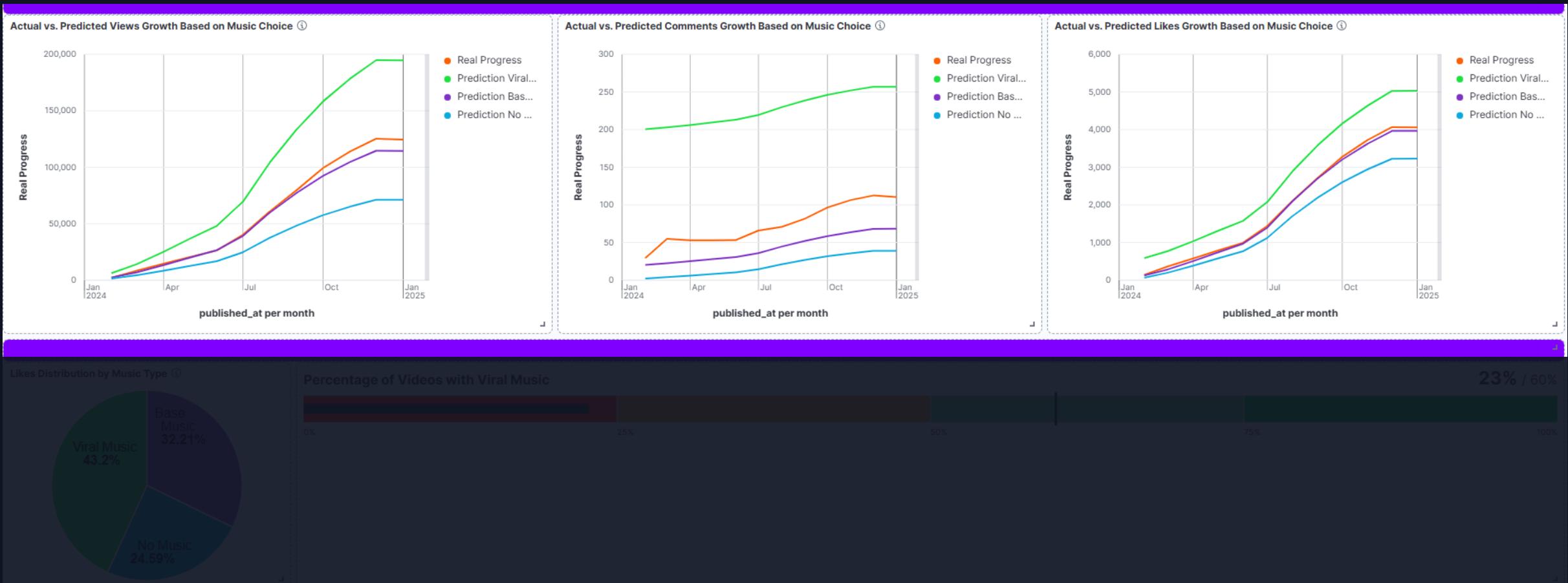
Section 2: Channel Metrics



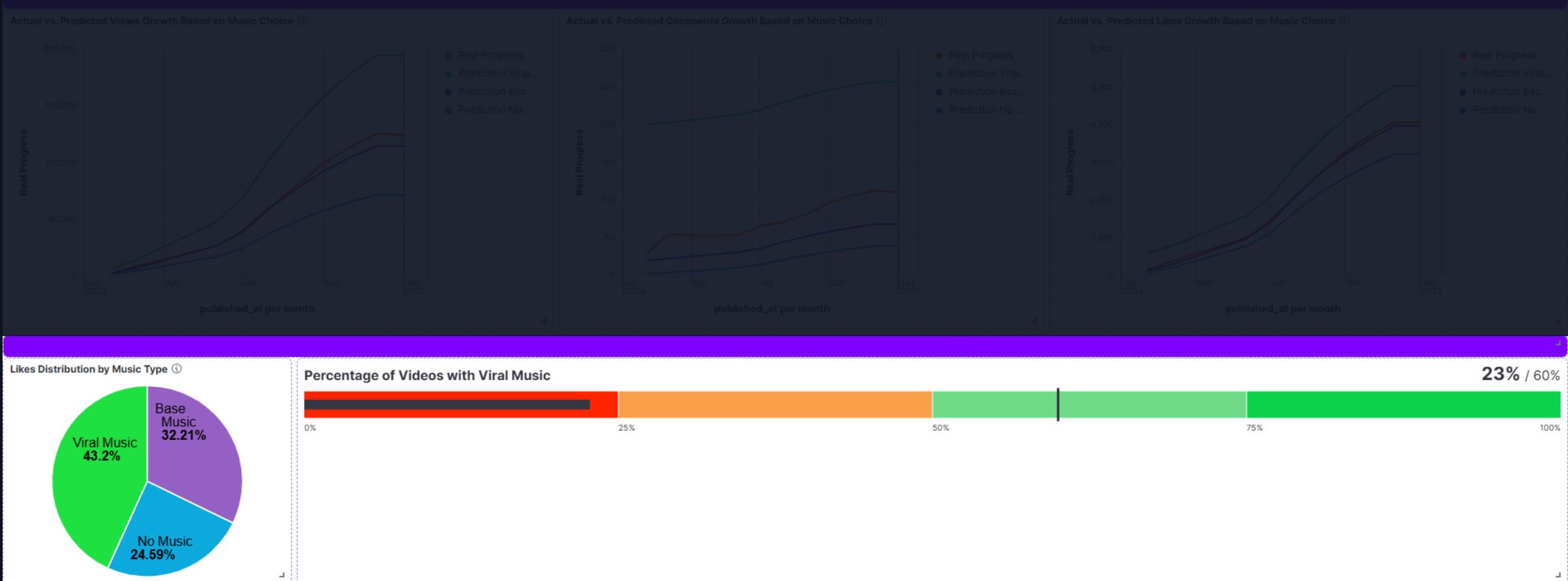
Section 3: Engagement Explorer



Section 4: Performance Predictions



Section 5: Viral Vibes





Vistorm unisce creatività e tecnologia, fornendo ai creator strumenti avanzati per comprendere e dominare il mondo della viralità, trasformando ogni contenuto in un'opportunità di successo.

A cartoon illustration of Bugs Bunny wearing a blue suit jacket over a purple shirt. He is looking towards the right with a wide-eyed, slightly worried expression. The background consists of a dense pattern of colorful, radiating comic book-style speed lines in shades of blue, yellow, pink, and white, creating a sense of motion and energy.

HAT'S
FOLKS



THANK YOU
FOR YOUR
ATTENTION

Giulia Pulvirenti