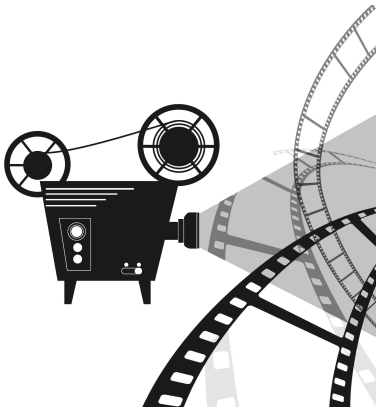
C-Rating™ Project

Movies that do rock!

Contents

Today's movie?



Clarity

We are all into movies
Choice is a pain
We want a market-unique, augmented rating

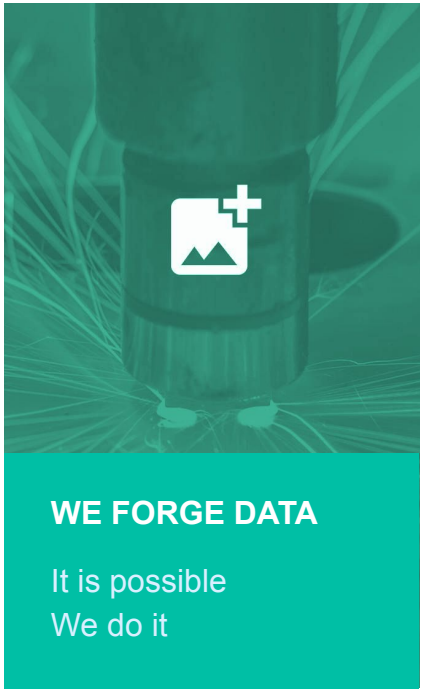
Goal

Reconcile the data out there
Augment our knowledge
Share it with the world

Movies data sources

from partner platforms

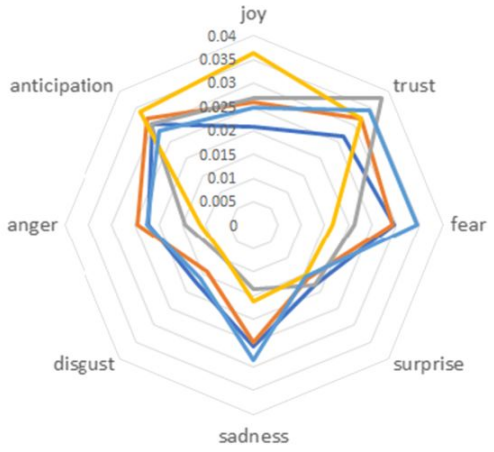
- Web scrapers and crawlers
  - IMD
  - Rotten Tomatoes
  - TMDB
  - ...
- FTP files received from 3rd parties
- XML / JSON Web Services

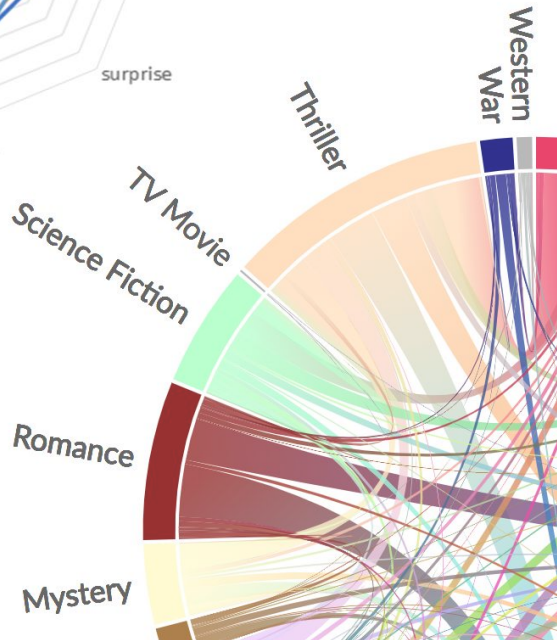


Augmented AI processes

C-Movie added value

- Title elucidation
- Genre and topic classifiers
- Enhanced ratings
- Targeted, personalized and context-aware recommendations





IDEAS THAT WILL BLOW YOU AWAY

It is possible - We do it

- Twitter Sentiment Analysis
- News contextual analysis
- Community and network analysis

Augmented AI processes

C-Movie added value

Retos

- Heterogeneous genres / no standard
- Inconsistent categories / if any
- Ratings & information are boring

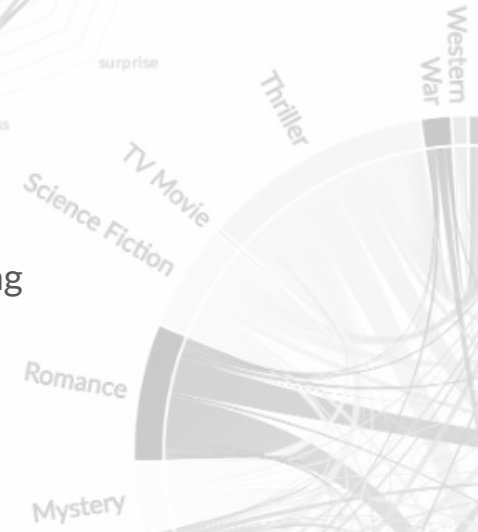
Soluciones

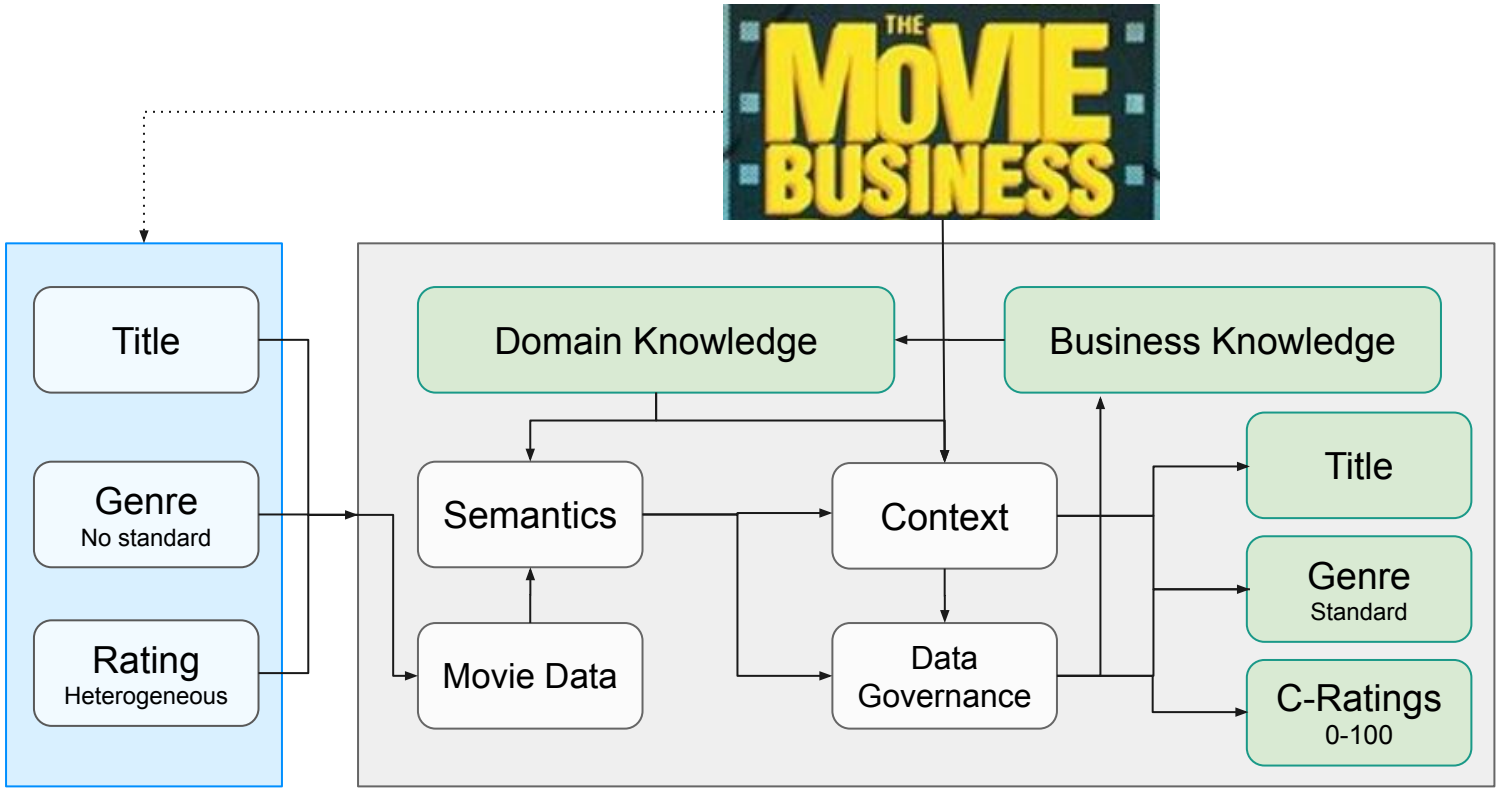
- Twitter Sentiment Analysis
- News contextual analysis

Data is augmented with real-world information

- C-Movie's Augmented A.I. & Governance
- Genre classifier with Natural Language Understanding
- Feedback Loop : Learning every day

The system learns in real time and is designed to govern business knowledge





External sources : Evolving ecosystem

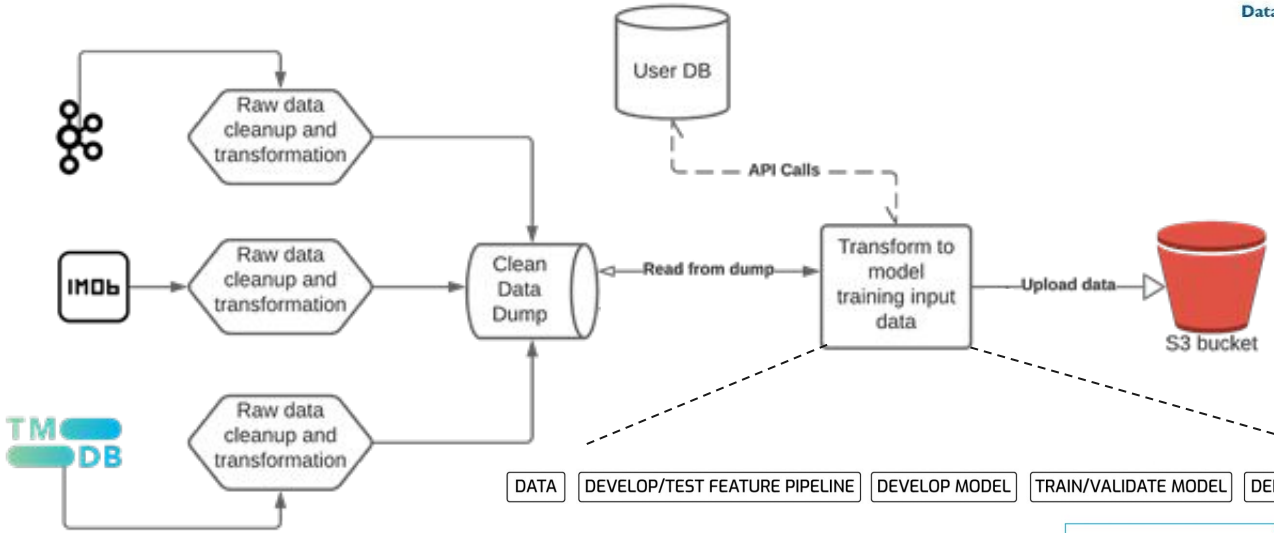
Data providers

C-Ratings Augmented Data Processes

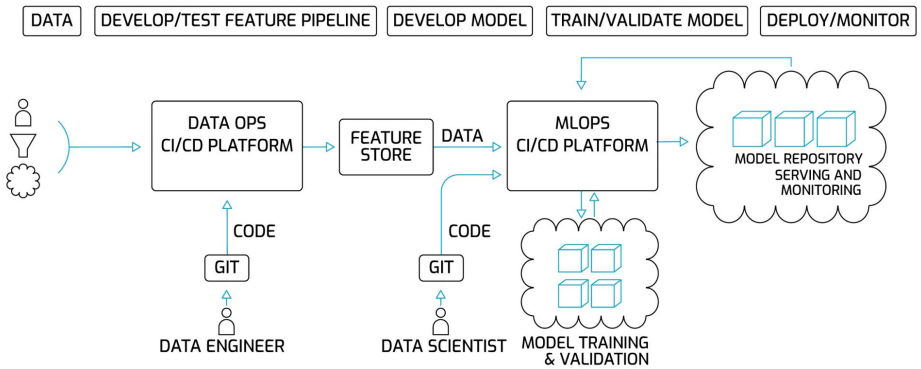
Use Case Solution + Augmented Domain/Business Knowledge

FEEDBACK LOOP

Simple Architecture Inspiration



Data Marketplace

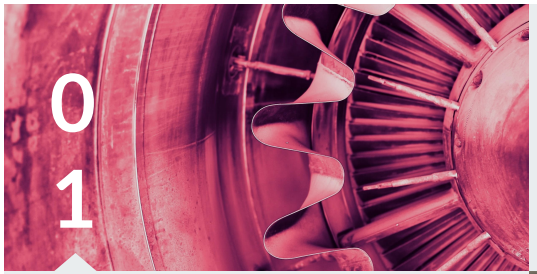


Storyline

1. **Data acquisition** using C-Movies data connectors
2. **Technical Solution**
   a. Aggregate & consolidate heterogeneous data (applying data augmentation when possible)
      i. Movie Title Aggregation (Profile / Pattern)
      ii. Genre Aggregation (Profile / Pattern)
      iii. Rating normalization, aggregation, weighted average
   b. Apply C-Ratings Proprietary Model
      i. Elucidates the true Title
      ii. Proposes the most suitable genre
      iii. Calculate C-Rating
3. **Data augmentation** using NLU agents learn from what is happening in the real world
   a. What do viewers feel?
      i. Rating enhancement
   b. Contextual analysis
      i. C-Category assignment
      ii. Performance, screenplay, soundtrack
4. **Business Solution**
   a. C-Ratings Proprietary Model Lifecycle
      i. Train & Retrain
      ii. Consumption
   b. Business Governance
      i. Business Process Modeling (BPM)
      ii. Manual Process curation (Domain Experts)
      iii. Business Glossary
      iv. Business Concepts
   c. Data Marketplace
      i. Governed Business Data Views
      ii. Internal use : IAM and BPM
      iii. External sharing: Data Contracts
      iv. Governed APIs

Key players

Containerized pipelines with Apache Airflow
Business Knowledge augmentation applying semantics / ontologies
Business Governance : BPM, IAM and a Business Data Marketplace designed to provision, trade & comply



Data processing

Apache Spark
Python & Dusk
Scikit - Learn
pytorch

Parquet files
S3 buckets
Redshift
MongoDB

Data acquisition

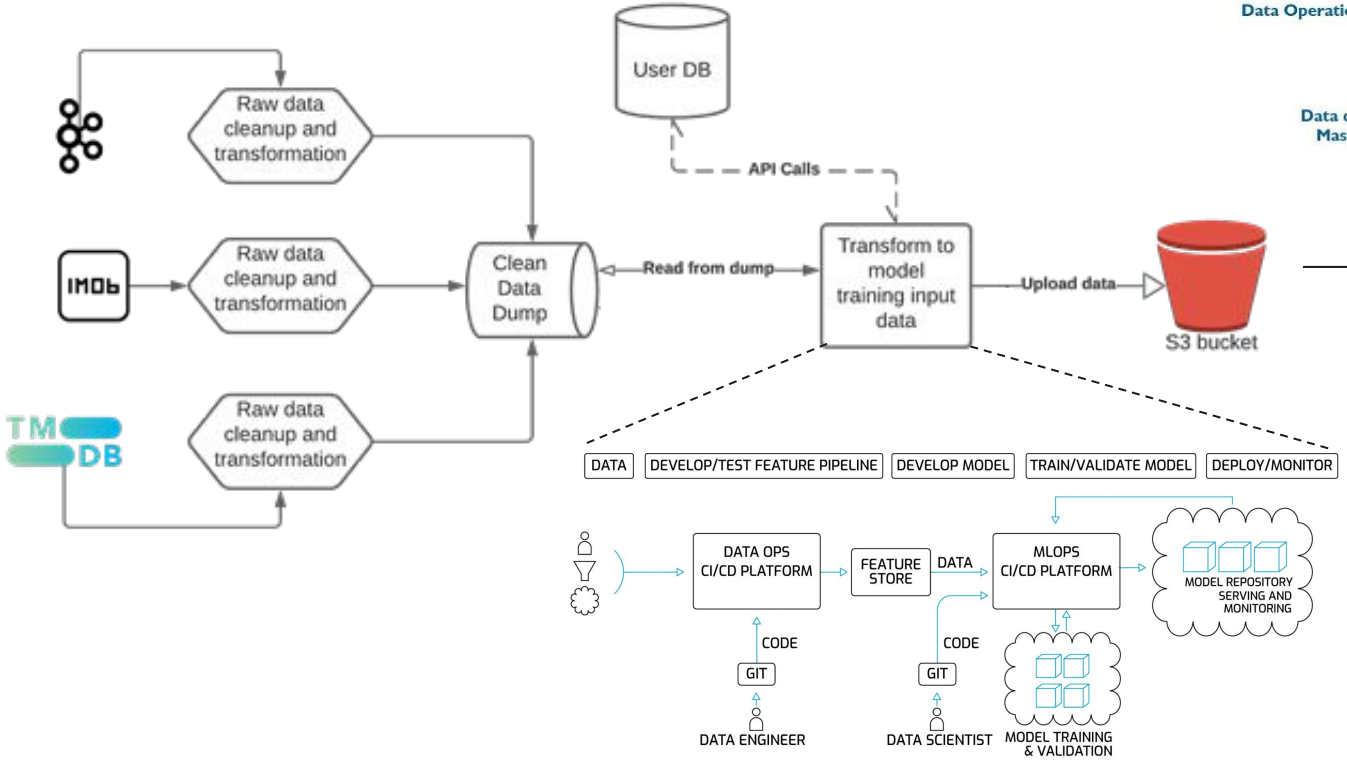
Confluent Kafka
Python
Apache Spark

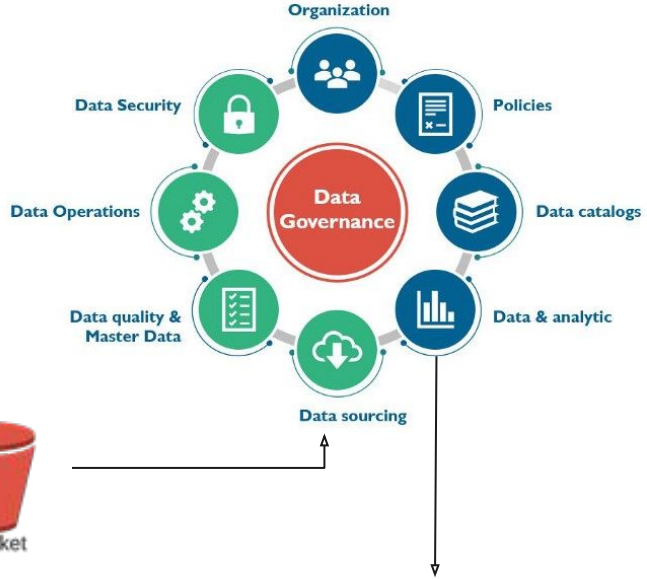
Productivization

Seldon
Jenkins
Hooks
MLflow

Python
fastai
fastapi
Flask

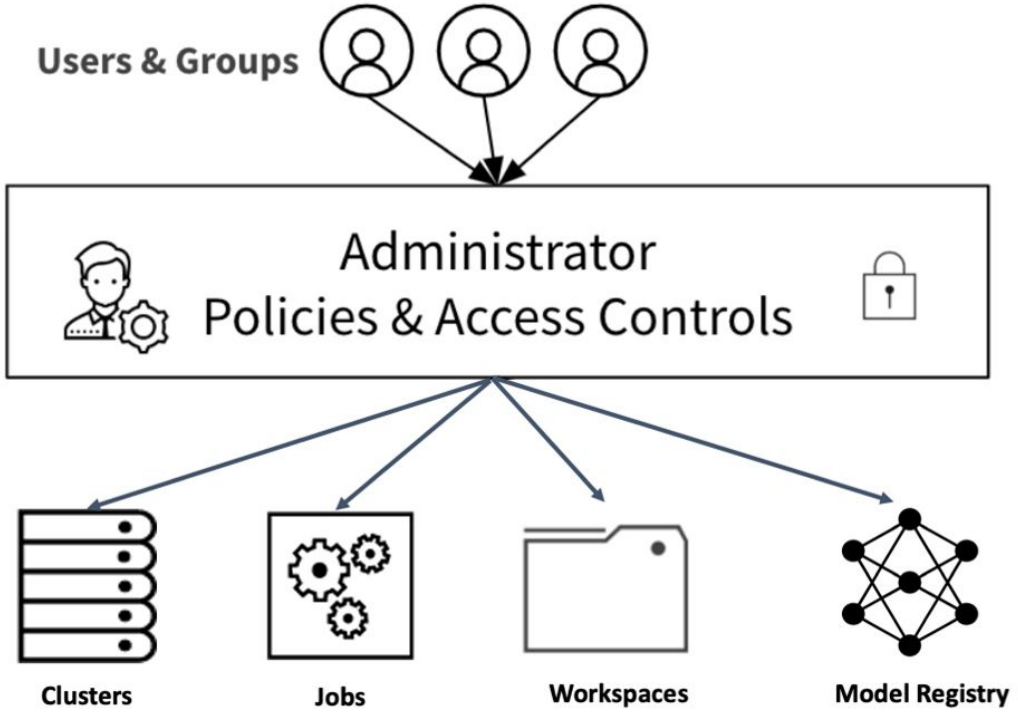
Simple Architecture Inspiration



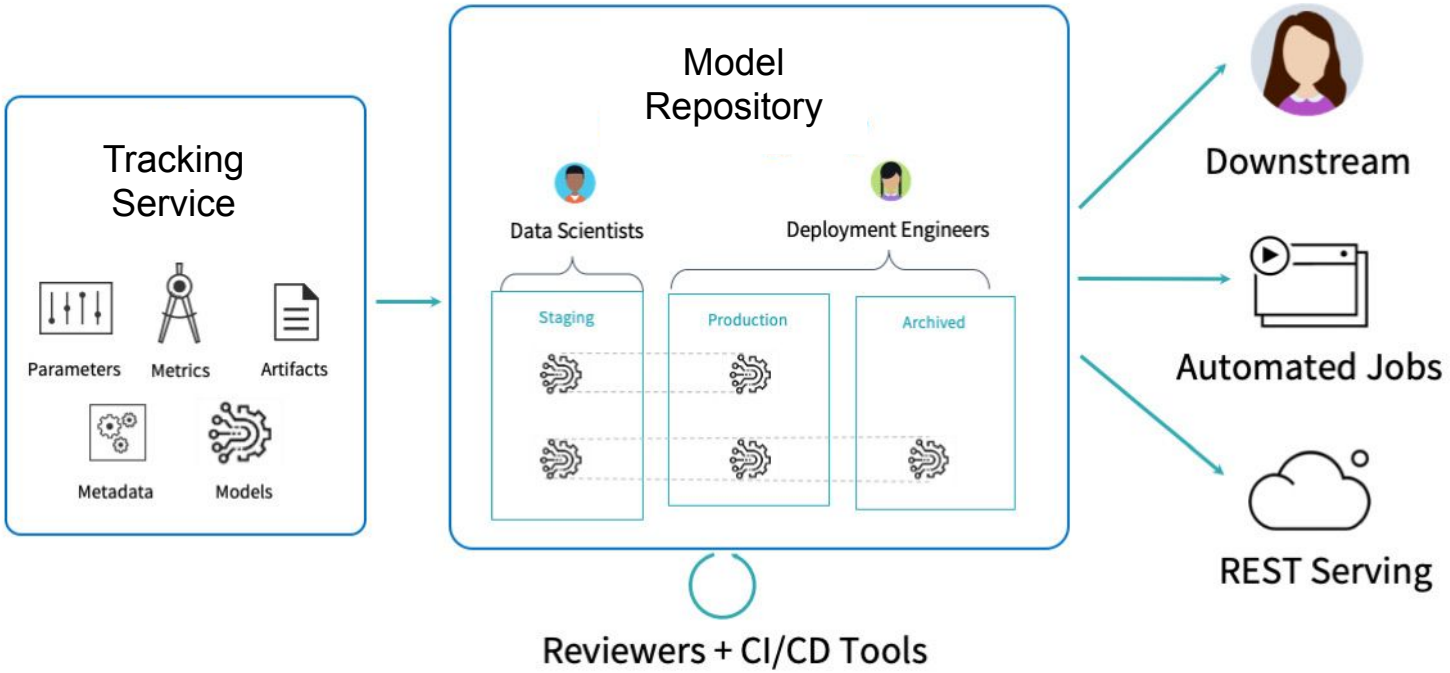


Data Marketplace

Governing Data & Business Processes



MLOps : C-Ratings Proprietary Model Lifecyle



PRIORITIZED BACKLOG

1. Required data connectors
2. Title aggregator
3. Genre aggregator
4. Rating aggregator
5. Title consolidator (most represented title)
6. Genre aggregator (expert genre mapping)
7. C-Ratings Proprietary Algorithm Consumption process
8. Results simple ungoverned load process
9. Simple API : Returns Title, Genre and C-Ratings
10. Simple Front (C-Movie UX)
11. Twitter data connector
12. News data connectors
13. Semantics engine: Enhance C-Ratings Proprietary Algorithm with Artificial Intelligence
    a. Enhanced multilingual title and data processing
    b. Enhanced genre classification
    c. Enhanced ratings using real-world information
    d. Examples / Inspirational:
        i. LDA Topic detection, clustering and genre mapping (simple)
        ii. NLU and NLG: Understand and generate Texts (complex)

Many Thanks.

quijanocsm@gmail.com