

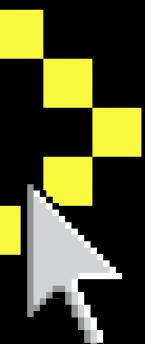


••DCODE  
HACKATHON

# TEAM- THIRST E

**Problem Statement #6**  
**Discoverability Engine for OSS Projects**

**WEBSITE NAME- RANKmeHARDER**

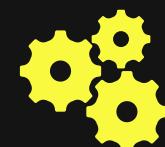


# PROBLEM STATEMENT

## Context:

Developers face a major challenge in discovering high-quality open-source projects on platforms like GitHub.

Despite millions of repositories, current discovery methods rely mainly on popularity metrics (stars, forks) which often overlook true quality or innovation.



Maintenance quality.



Documentation.

## As a result:

Technically strong but lesser-known projects remain hidden.

Developers waste time exploring poorly maintained repositories.

Contributors struggle to find beginner-friendly or innovative projects.

## Impact:

High-value projects lack visibility and support.

OSS collaboration and innovation slow down.

Ecosystem becomes skewed toward already popular projects.



## Key Insight

A repository's true value is multi-dimensional, depending on:



Community engagement.



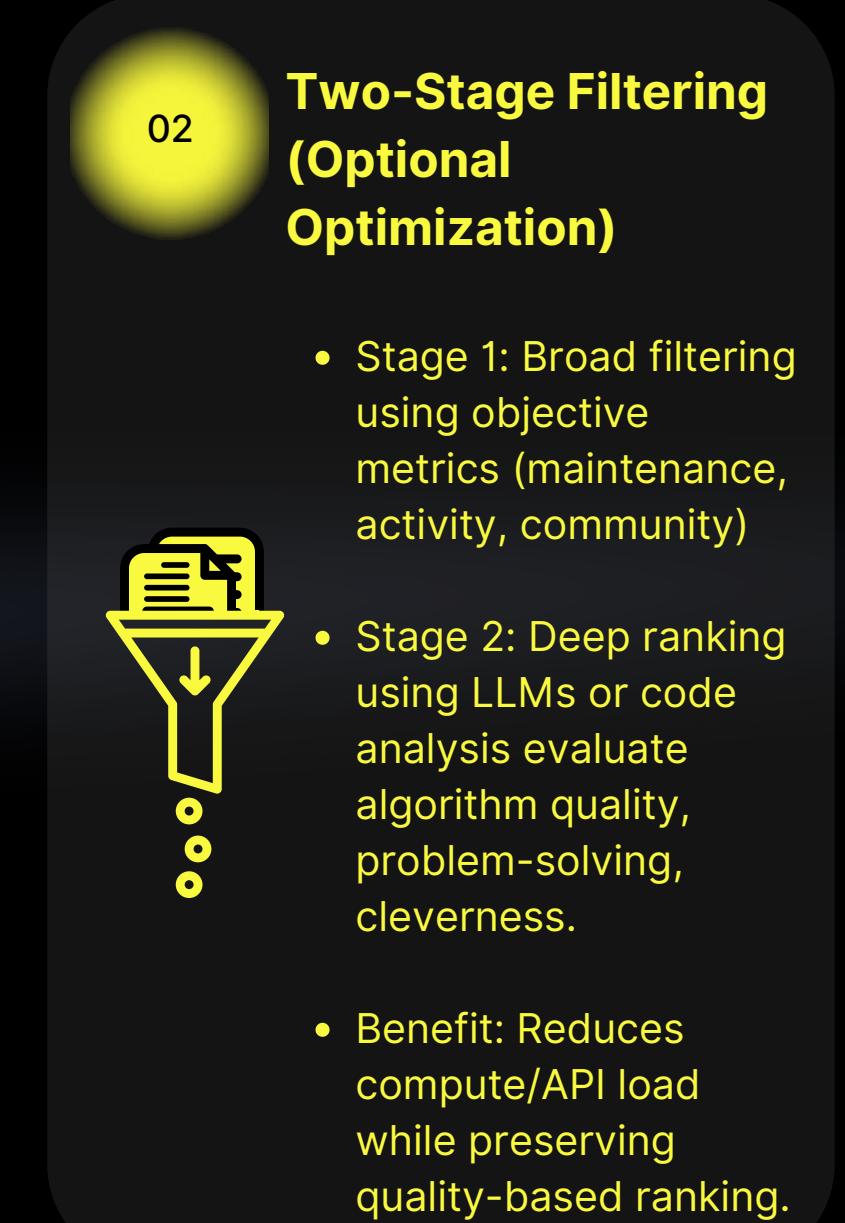
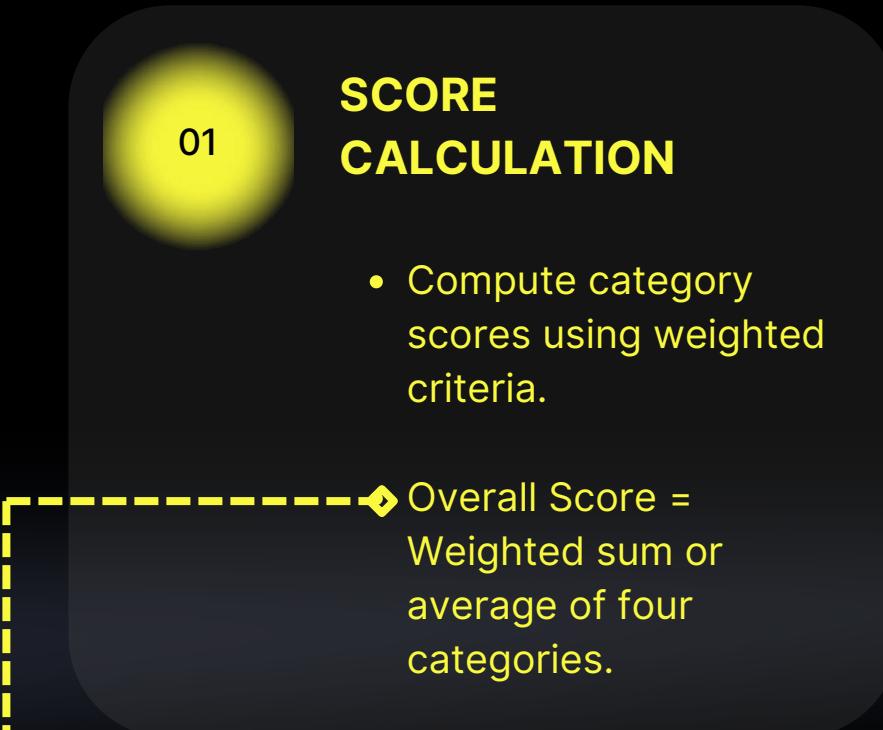
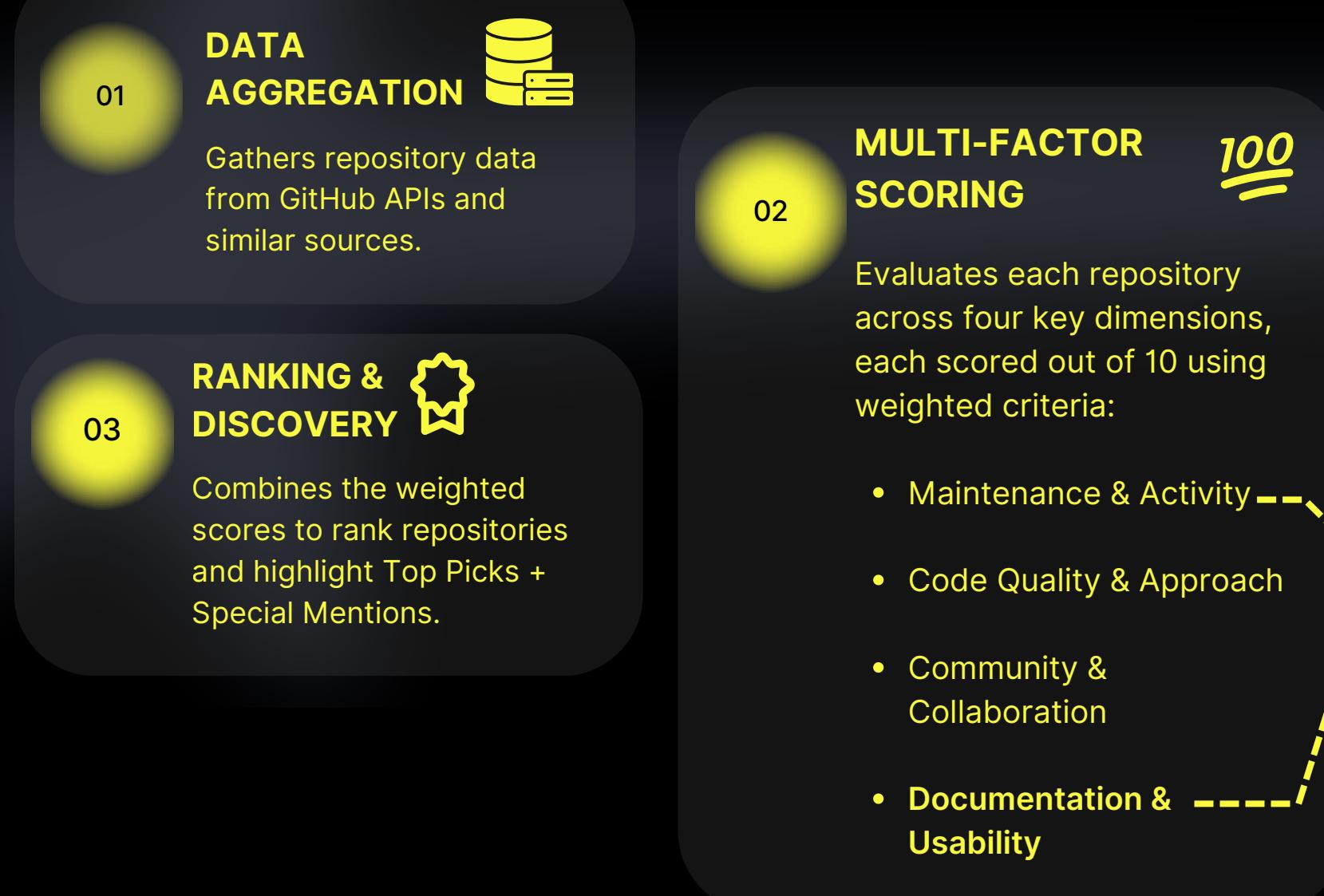
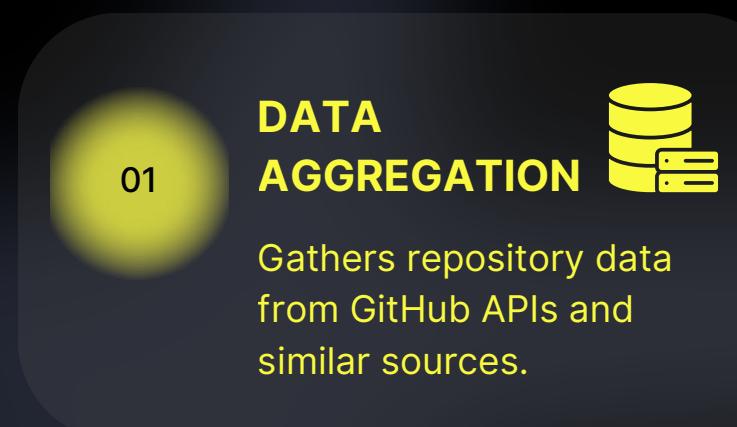
Innovation.

# HIGH-LEVEL SOLUTION OVERVIEW

## RANKING & DISPLAY LOGIC

### Our Approach

We propose a Scored Discoverability Engine that evaluates open-source repositories holistically, not just by popularity.

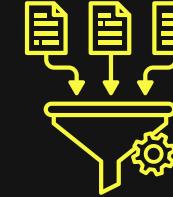


# TECHNICAL APPROACH

## SYSTEM ARCHITECTURE

### 01 Data Ingestion:

- Fetch repo stats via GitHub GraphQL + REST
- Cache & queue jobs (Redis + Celery)



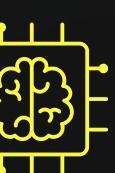
### 02 Scoring Engine:

- Normalize metrics (log + percentile scaling)
- Compute 4 category scores → weighted overall



### 03 AI/LLM Module:

- Analyze README + code for clarity & novelty
- Generate qualitative scores (0–10)



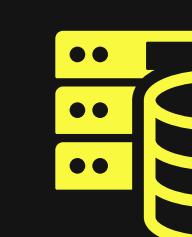
### 04 API & Frontend:

- FastAPI backend serves scores & filters
- React dashboard visualizes rankings



### 05 Database:

- PostgreSQL stores repos, metrics, scores, history



## # Key Technologies

Python

FastAPI

Redis

Celery

PostgreSQL

Github API

GPT-5

React

Docker

GitHub Actions

## # Core Principles



- Transparency: Explainable, reproducible scores



- Fairness: Normalized for project size & age



- Scalability: Async workers + cached API



- Extendability: Plug-in metrics or LLM modules



- Cost-Efficiency: AI analysis only for top repos

## # Outcome

An explainable, ranked list of open-source projects by real quality signals enabling smarter discovery for developers, investors, and communities.

# FEASIBILITY, SCALABILITY & DEVELOPMENT PLAN

## Built on Proven Foundations

- Leverages mature GitHub APIs and stable open-source tools (React, FastAPI/Node.js, PostgreSQL).
- Modular microservice architecture ensures each component (data, scoring, AI, frontend) scales independently.



## Scalable & Performant Design

Caching (Redis) + two-tier filtering pipeline for high responsiveness even with millions of repos.

- Async task queues (Celery/Worker nodes) handle heavy compute workloads efficiently.
- Cloud-ready deployment using Docker + CI/CD (GitHub Actions) for seamless updates.



## Transparent & Extensible Scoring

- Weighted scoring system is explainable, reproducible, and easy to extend with new metrics.
- Plug-in architecture for adding future AI/LLM modules without major refactors.



## Future Enhancements

- Real-time community sentiment tracking
- Integration with package registries (npm, PyPI, crates.io)
- Contributor & maintainer reputation metrics
- Fun additions during search loading like Monkeytype, Dino Game etc.

# IMPACTS AND BENEFITS

TRANSFORMING THE OPEN-SOURCE LANDSCAPE

Our Repository Ranker will empower developers worldwide to:

Discover quality, impactful, and well-maintained projects previously hidden by popularity bias.



Support newcomers by spotlighting projects with excellent documentation and welcoming communities.



Encourage contributions to diverse, sustainable projects, strengthening the open-source ecosystem.



Empower users, organizations, and maintainers with transparent, actionable metrics for better decisions and incentives.



# **THANK YOU**

**BY TEAM MEMBERS:**

**SRIHARSHA MALLAVARPU**

**PARUSHYA DATA**

**YUVRAJ SINGH**

**AAMIR HAMZA**

**VATSALYA VIKRAM**