# Recommender Metrics Framework

## Measuring the success of a Recommender System

Nikolaos Triantafyllis (GRNET)

Kostas Kagkelidis, Nikolaos Triantafyllis, Themis Zamani, Kostas Koumantaros
{kaggis, ntriantafyl, themis, kkoum}@admin.grnet.gr
National Infrastructures for Research and Technology (GRNET)

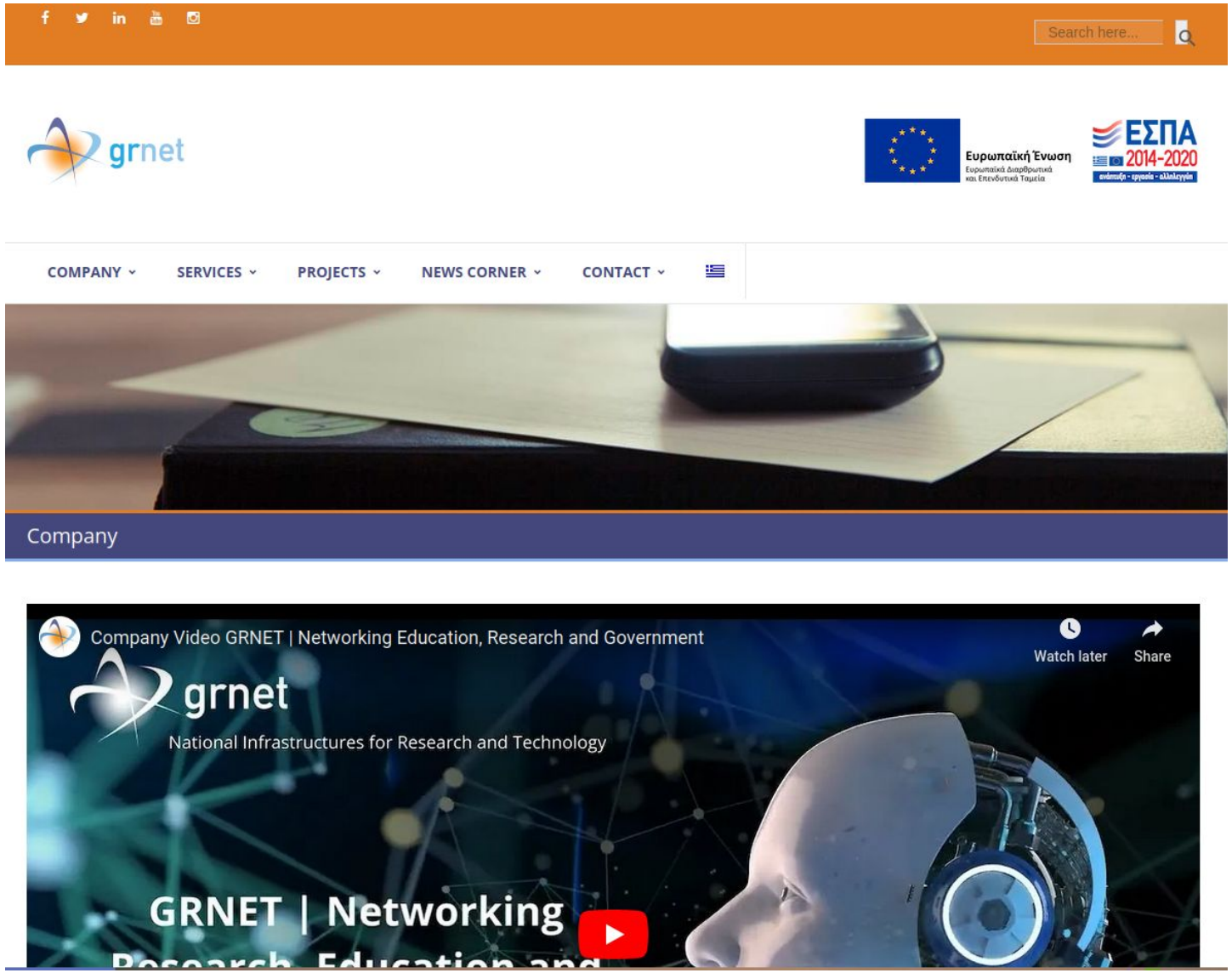
FOSS COMM 23
FREE & OPEN SOURCE SOFTWARE COMMUNITIES MEETING

The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement 101017536

CC BY ND

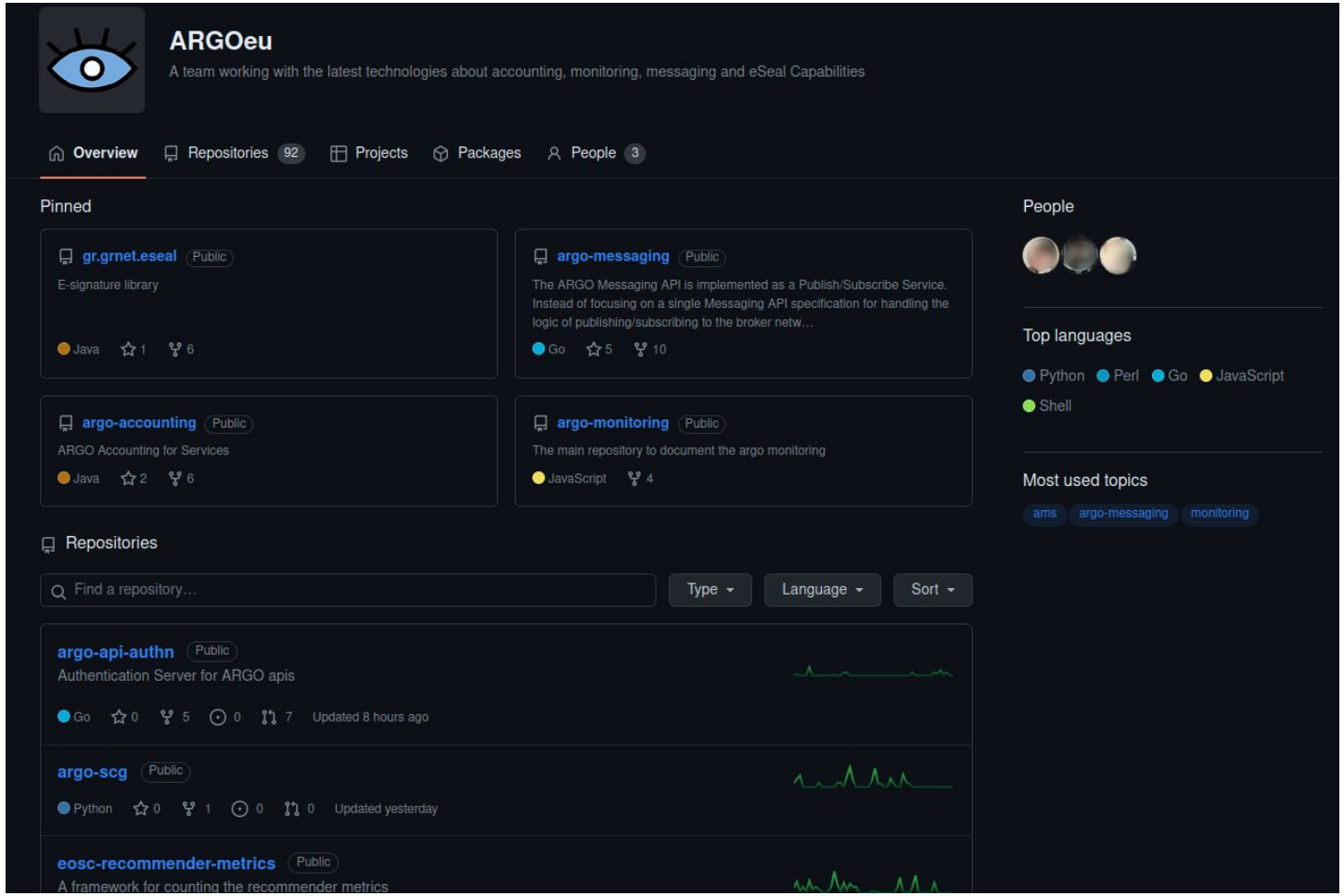
# Who we are?

- GRNET S.A. – National Infrastructures for Research and Technology, is one of the largest public sector technology companies in Greece.
- It provides networking, cloud computing, HPC, data management services and e-Infrastructures and services to academic and research institutions, to educational bodies at all levels, and to all agencies of the public sector.
- https://grnet.gr



# Who we are?

- The European Infrastructures and Projects Directorate.
- Among others, it develops software solutions in various research projects across EU and Greece.
- One of the subteams of the Directorate that lies in GitHub Organizations: https://github.com/ARGOeu
- Recommender Metrics Framework (RMF) is developed by GRNET and used as an open-source solution in the EOSC-Future project.



# What are we going to talk about?

A Recommender System's Metrics Framework that can produce measurable results for a Recommender System's evaluation.

# Where is it used?

Will explain in a few minutes

To monitor and report diagnostic metrics for the EOSC Marketplace Recommender Service.

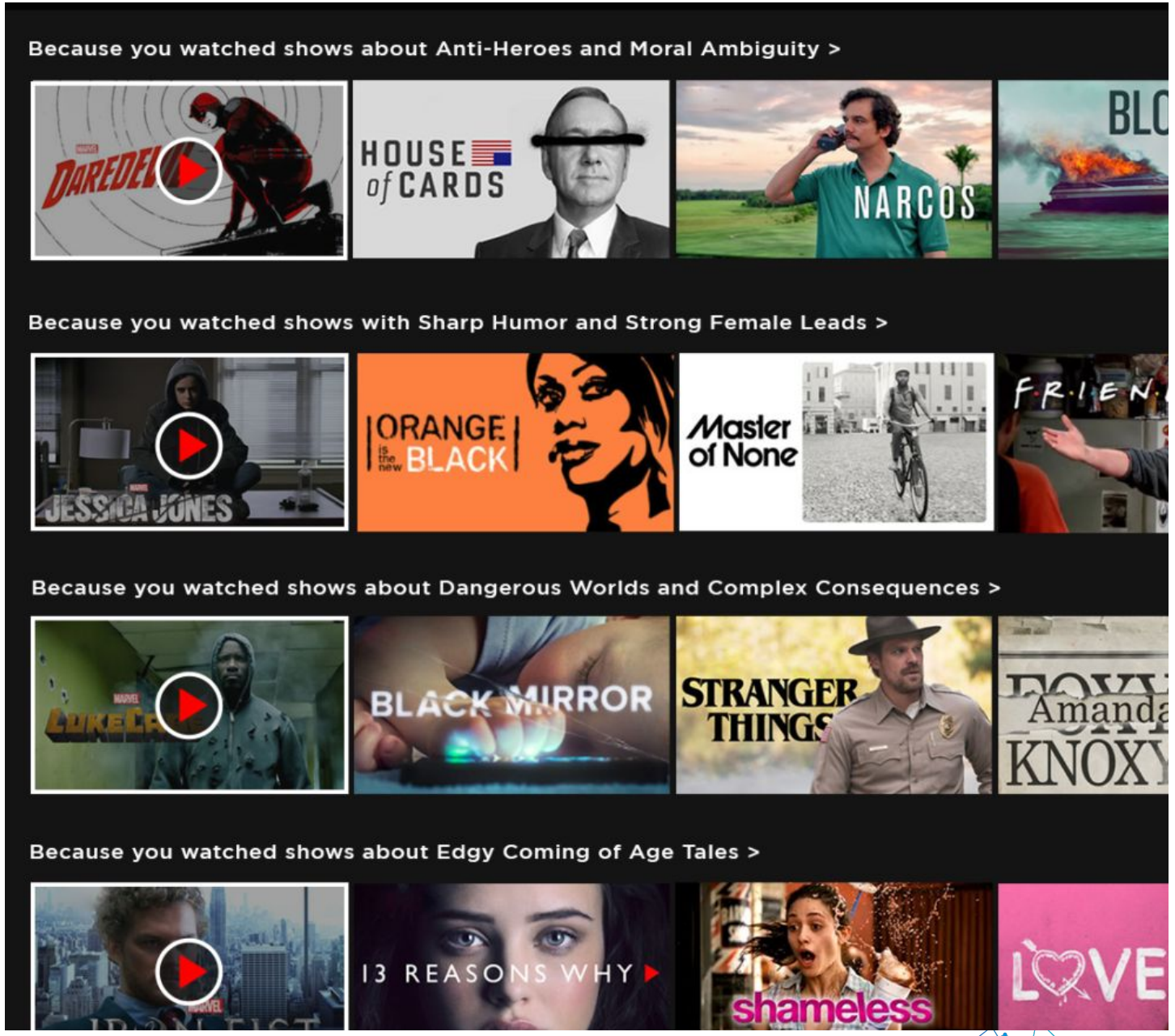
**Recommender Service**

- **Owners**
- **Developers**
- **Engineers**

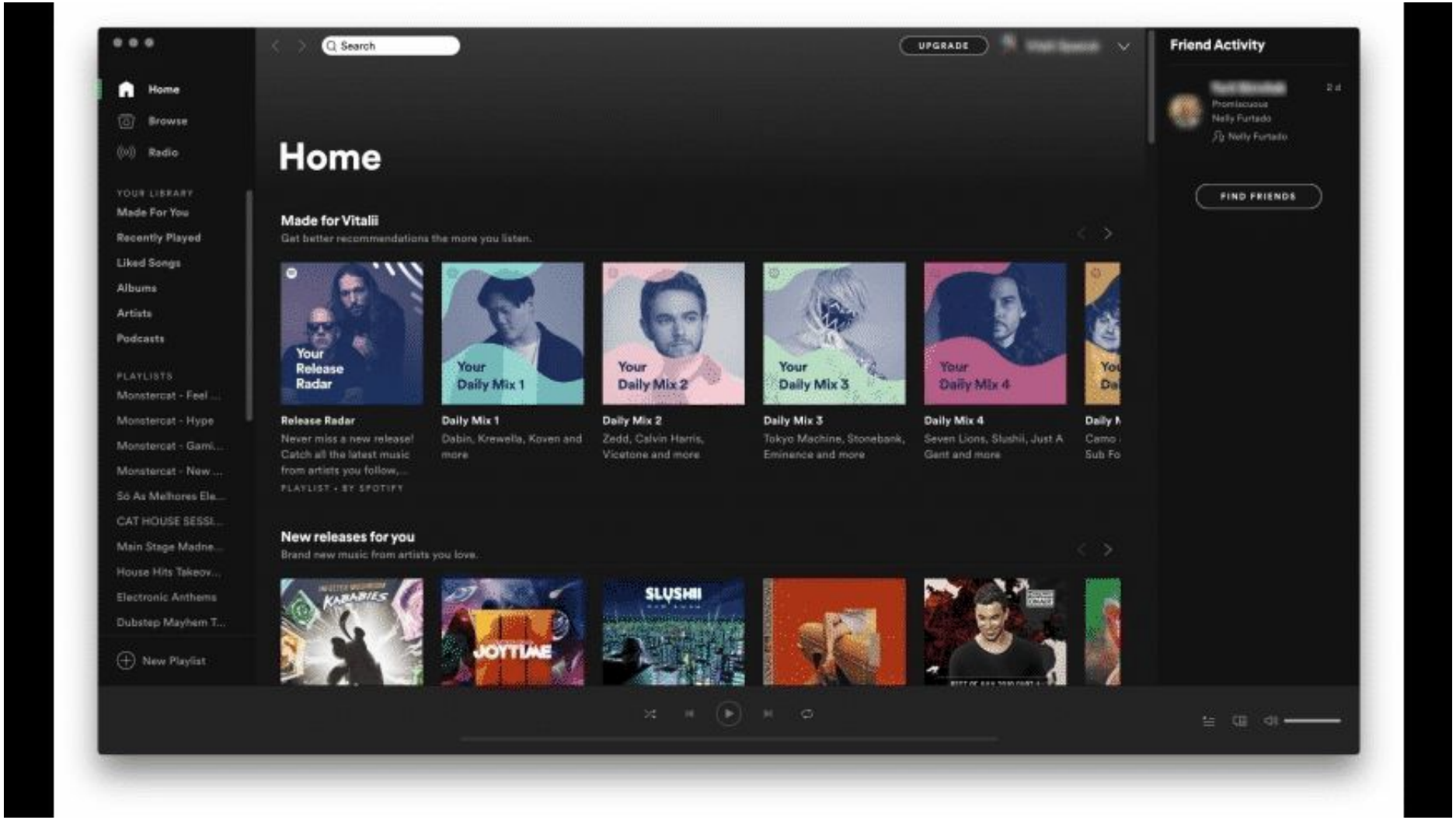
**Service using the RS Marketing Team**

# What is a Recommender System (RS)?

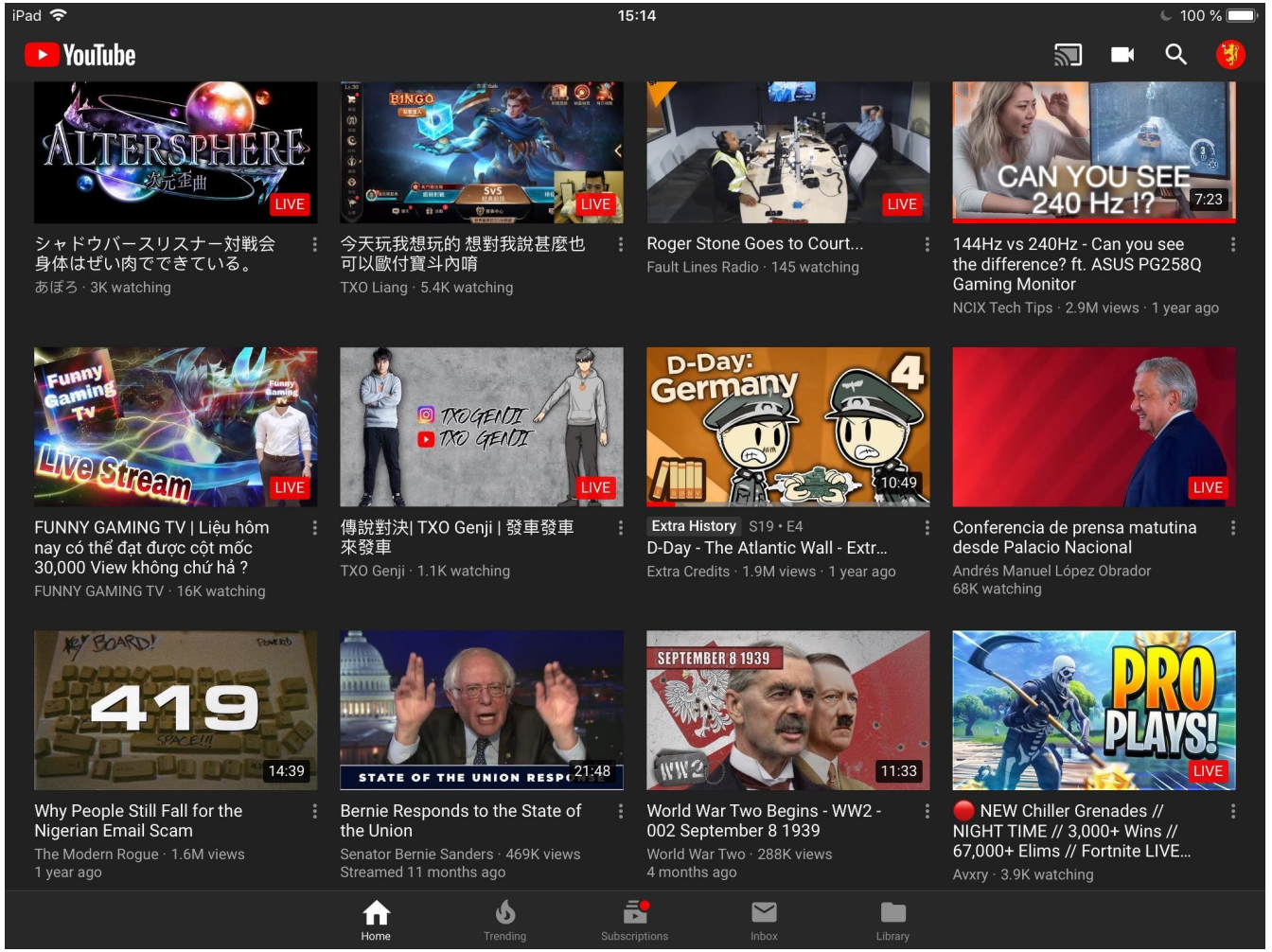
- Offers personalized suggestions to users.
- Recommendations are based on user preferences, behaviors, patterns.
- Recommendations can include products, content, services, or connections.
- Aims to enhance user experience.
- Applications include e-commerce, content streaming, and social media.
- Addresses the information overload problem in the digital age.



# More...


Search
UPGRADE
Home
Browse
Radio
YOUR LIBRARY
Made For You
Recently Played
Liked Songs
Albums
Artists
Podcasts
PLAYLISTS
Monstercat - Feel ...
Monstercat - Hype
Monstercat - Gami...
Monstercat - New ...
Só As Melhores Ele...
CAT HOUSE SESSI...
Main Stage Madne...
House Hits Takeov...
Electronic Anthems
Dubstep Mayhem T...
New Playlist
Home
Made for Vitalii
Get better recommendations the more you listen.
Your Release Radar
Release Radar
Never miss a new release! Catch all the latest music from artists you follow,...
PLAYLIST • BY SPOTIFY
Your Daily Mix 1
Daily Mix 1
Dabin, Krewella, Koven and more
Your Daily Mix 2
Daily Mix 2
Zedd, Calvin Harris, Vicetone and more
Your Daily Mix 3
Daily Mix 3
Tokyo Machine, Stonebank, Eminence and more
Your Daily Mix 4
Daily Mix 4
Seven Lions, Slushii, Just A Gent and more
New releases for you
Brand new music from artists you love.
KABABIES
JOYTIME
SLUSHII
Friend Activity
FIND FRIENDS

# More...


iPad
15:14
100 %
YouTube
ALTERSPHERE
次元歪曲
LIVE
BINGO
5VS
LIVE
LIVE
CAN YOU SEE
240 Hz !?
7:23
シャドウバースリスナー対戦会
身体はぜい肉でできている。
あぽろ · 3K watching
今天玩我想玩的 想對我說甚麼也
可以歐付寶斗內啃
TXO Liang · 5.4K watching
Roger Stone Goes to Court...
Fault Lines Radio · 145 watching
144Hz vs 240Hz - Can you see
the difference? ft. ASUS PG258Q
Gaming Monitor
NCIX Tech Tips · 2.9M views · 1 year ago
Funny
Gaming
Tv
Live Stream
LIVE
TXOGENJI
TXO GENJI
LIVE
D-Day:
Germany
4
10:49
LIVE
FUNNY GAMING TV | Liệu hôm
nay có thể đạt được cột mốc
30,000 View không chứ hả ?
FUNNY GAMING TV · 16K watching
傳說對決| TXO Genji | 發車發車
來發車
TXO Genji · 1.1K watching
Extra History S19 • E4
D-Day - The Atlantic Wall - Extr...
Extra Credits · 1.9M views · 1 year ago
Conferencia de prensa matutina
desde Palacio Nacional
Andrés Manuel López Obrador
68K watching
419
14:39
STATE OF THE UNION
21:48
SEPTEMBER 8 1939
WW2
11:33
PRO
PLAYS!
LIVE
Why People Still Fall for the
Nigerian Email Scam
The Modern Rogue · 1.6M views
1 year ago
Bernie Responds to the State of
the Union
Senator Bernie Sanders · 469K views
Streamed 11 months ago
World War Two Begins - WW2 -
002 September 8 1939
World War Two · 288K views
4 months ago
NEW Chiller Grenades //
NIGHT TIME // 3,000+ Wins //
67,000+ Elims // Fortnite LIVE...
Avxry · 3.9K watching
Home
Trending
Subscriptions
Inbox
Library

# More...

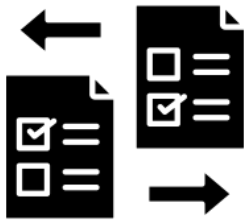
# But…

- How “good” recommendations are?

- What is “good”?

# Why do we need RS metrics?

**Evaluation:** They assess RS performance and effectiveness.



**Comparison:** They help compare different systems or variations.

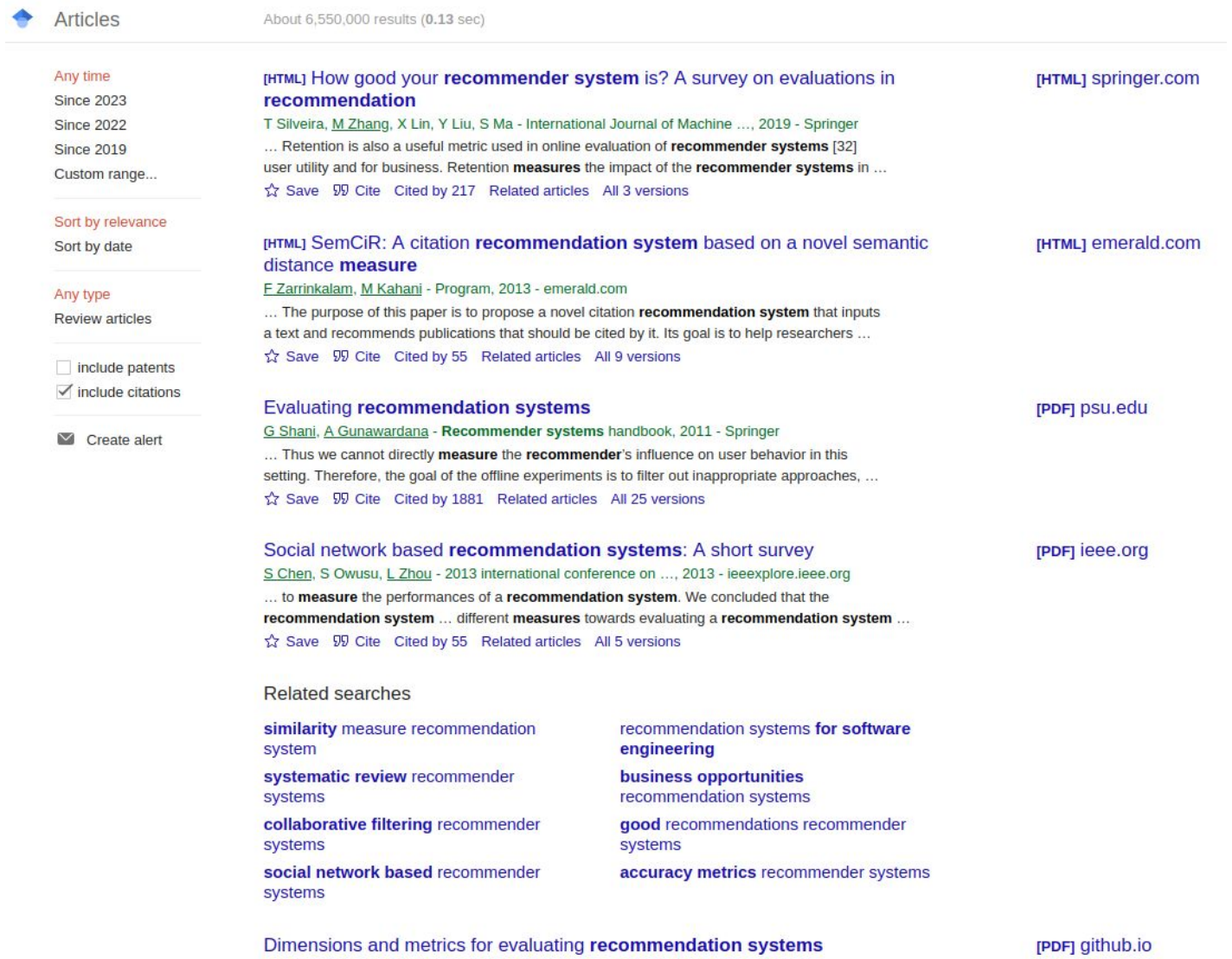
**User Satisfaction:** They measure user engagement and they can help in building user trust.

**Fairness**: They can be used to mitigate biases in recommendations, promoting fairness and inclusivity.

**Optimization**: They guide improvements and fine-tuning in algorithms and parameters, whereas they support continuous system improvement.

**Business Impact:** They can be tied to financial outcomes (KPIs).

# Science in the background...



Articles

About 6,550,000 results (**0.13** sec)

Any time
Since 2023
Since 2022
Since 2019
Custom range...

Sort by relevance
Sort by date

Any type
Review articles

include patents
include citations

Create alert

[HTML] How good your **recommender system** is? A survey on evaluations in **recommendation** — [HTML] springer.com
T Silveira, M Zhang, X Lin, Y Liu, S Ma - International Journal of Machine ..., 2019 - Springer
... Retention is also a useful metric used in online evaluation of **recommender systems** [32] user utility and for business. Retention **measures** the impact of the **recommender systems** in ...
Save Cite Cited by 217 Related articles All 3 versions

[HTML] SemCiR: A citation **recommendation system** based on a novel semantic distance **measure** — [HTML] emerald.com
F Zarrinkalam, M Kahani - Program, 2013 - emerald.com
... The purpose of this paper is to propose a novel citation **recommendation system** that inputs a text and recommends publications that should be cited by it. Its goal is to help researchers ...
Save Cite Cited by 55 Related articles All 9 versions

Evaluating **recommendation systems** — [PDF] psu.edu
G Shani, A Gunawardana - **Recommender systems** handbook, 2011 - Springer
... Thus we cannot directly **measure** the **recommender**'s influence on user behavior in this setting. Therefore, the goal of the offline experiments is to filter out inappropriate approaches, ...
Save Cite Cited by 1881 Related articles All 25 versions

Social network based **recommendation systems**: A short survey — [PDF] ieee.org
S Chen, S Owusu, L Zhou - 2013 international conference on ..., 2013 - ieeexplore.ieee.org
... to **measure** the performances of a **recommendation system**. We concluded that the **recommendation system** ... different **measures** towards evaluating a **recommendation system** ...
Save Cite Cited by 55 Related articles All 5 versions

Related searches

**similarity** measure recommendation system
**systematic review** recommender systems
**collaborative filtering** recommender systems
**social network based** recommender systems
recommendation systems **for software engineering**
**business opportunities** recommendation systems
**good** recommendations recommender systems
**accuracy metrics** recommender systems

Dimensions and metrics for evaluating **recommendation systems** — [PDF] github.io

# How to measure success?

**Statistics:** quantifies the occurrences of various data entities, such as user interactions, item popularity, or recommendation relevance scores.

**Metrics**: goes beyond simple **counts** and offer sophisticated characterizations of Recommender System's performance.

**KPIs**: Key Performance Indicators (KPIs) focused on business-oriented metrics, which are aligned with the overarching goals of the organization.

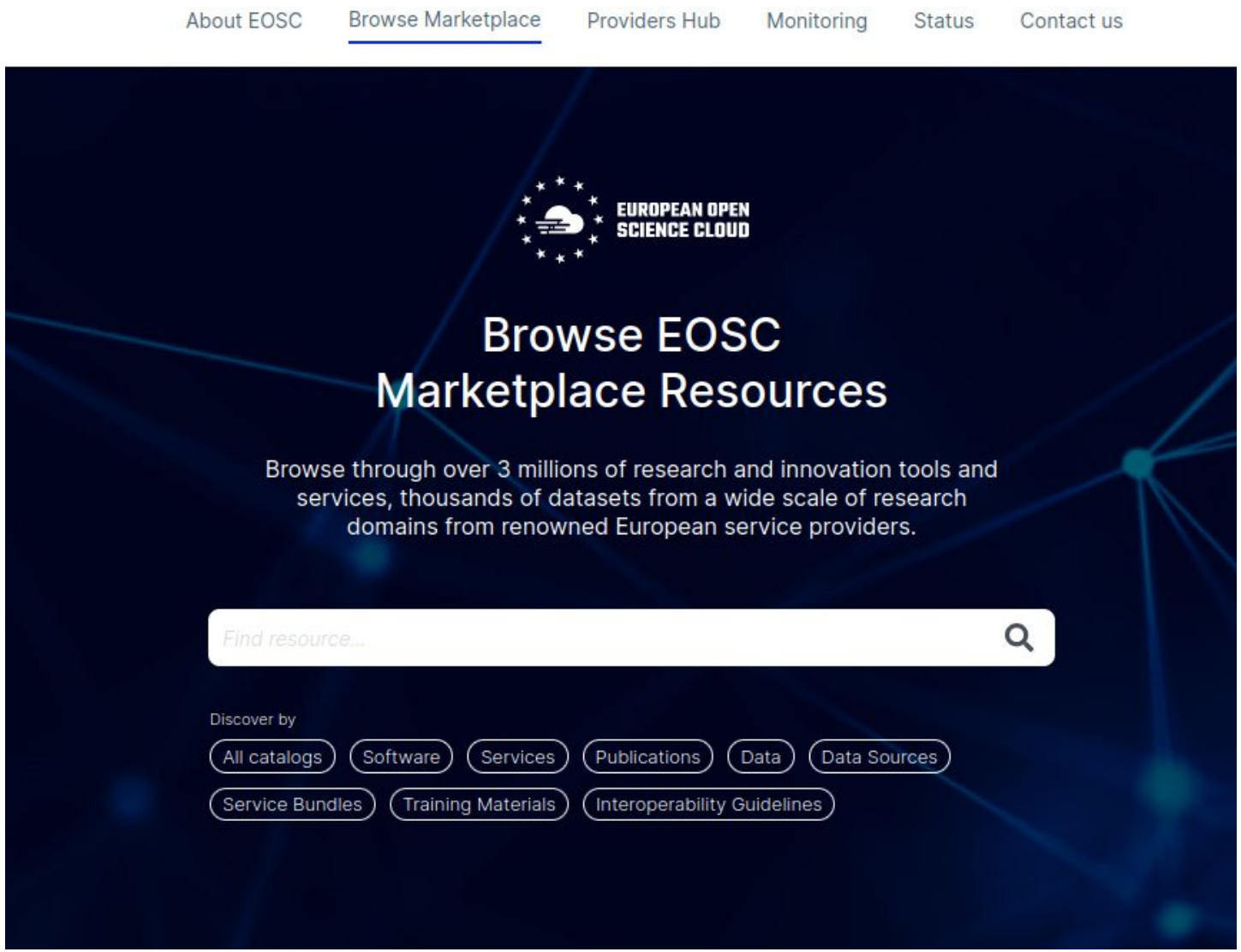
**Graphs**: visualizations of statistics/metrics across time helping in tracking trends, and identifying seasonality.

# What is EOSC?

- EOSC stands for the "European Open Science Cloud".
- It is a European initiative aimed at creating a unified, open environment for researchers and scientists.
- Its objectives include enabling access, sharing, and reuse of data, as well as providing access to research services and resources.
- EOSC promotes open science principles, fostering transparency and accessibility in research.
- Collaboration among researchers from various disciplines and institutions is a central goal of EOSC.

# What is EOSC Marketplace?

- The EOSC Marketplace is an online platform and ecosystem within EOSC: https://marketplace.eosc-portal.eu/
- It is designed to facilitate access to a wide range of digital resources, services, and data.
- It serves as a centralized hub for researchers and scientists in the EU research community.
- Researchers can use the marketplace to discover, access, and use resources that support their research activities.
- Resources available through the marketplace include data repositories, software applications, computing and storage facilities, trainings, and various research services.



# What is Recommender Metrics Framework (RMF)?

- A Python open-source software that monitors, analyzes, and evaluates recommendation mechanisms.
- Measures the effectiveness of the EOSC Marketplace RS to enhance the user experience and improve AI algorithms.
- Incorporates diagnostic statistics, metrics and visualizations for deeper insights into model performance.
- Presents reports as a web service and visualizes statistics, metrics, and Key Performance Indicators (KPIs) through a RESTful API and UI dashboard.
- Quantitative evaluation is taking into account EOSC Marketplace resources, user actions, and recommendations.
- Supports real-time and offline data ingestion, multiple resource types, and various recommendation engines as sources.
- Evolves over time, adding features and utilities to promote the development of more reliable and high-quality RS designs.

# Use Cases

IS CURRENTLY USED

## Monitoring the EOSC Marketplace RS

- Monitors and reports diagnostic metrics for the EOSC Marketplace RS.
- Analyses user actions and recommendations.
- Provides Statistics, Metrics, KPIs, Graphs in a REST API and dashboard UI.
- Delivers comprehensive documentation.

CAN BE USED

## Evaluate a third-party RS

- An analysis tool of the recommendation engine.
- Data preparation with the necessary input information.
- Tasks involve retrieving data from multiple sources, removing irrelevant data, correlating information, and generating **statistical insights**.

# What it offers?

## Statistics

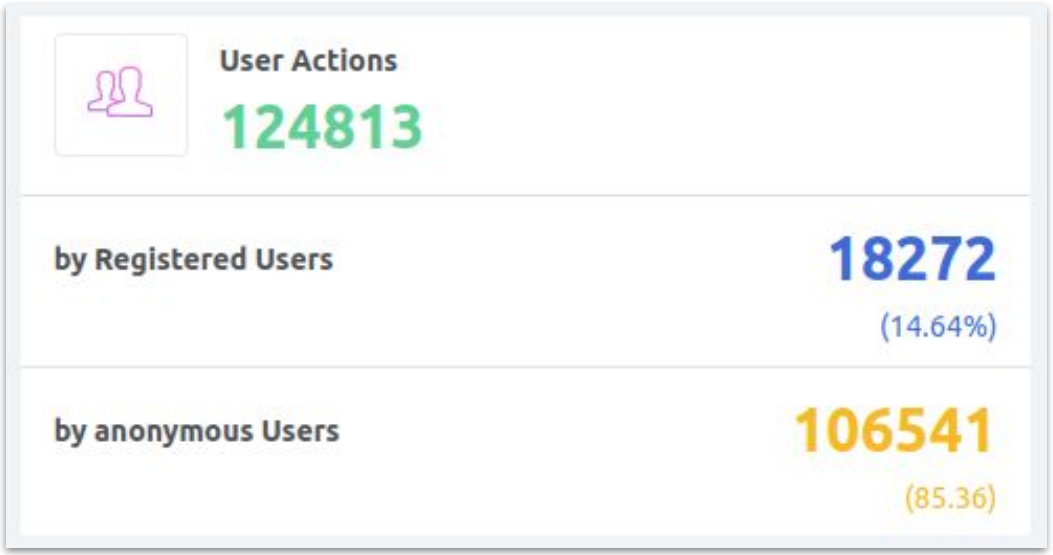
- Number of Users
- Number of Resources
- Number of Recommended items
- Number of User Actions by
  - Registered or
  - Anonymous users
- Total Views

# What it offers?

User Actions

## Statistics

- Number of Users
- Number of Resources
- Number of Recommended items
- Number of User Actions by
  - Registered or
  - Anonymous users
- Total Views



# What it offers?

## Metrics

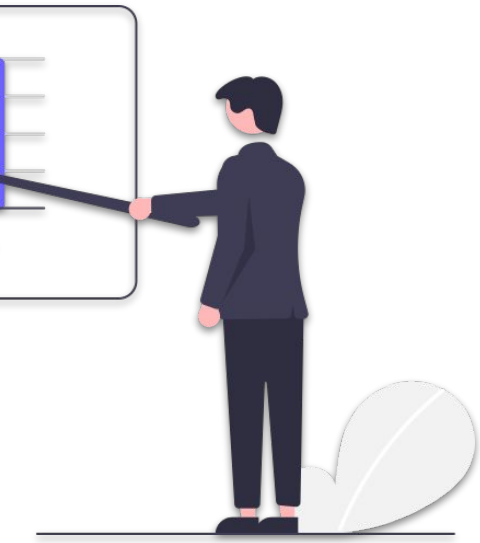
- Accuracy
- Catalog Coverage
- Diversity Gini Index
- Diversity Shannon Entropy
- Novelty
- User Coverage

# What it offers?

## Metrics

- **Accuracy**
- Catalog Coverage
- Diversity Gini Index
- Diversity Shannon Entropy
- Novelty
- User Coverage

Measures Recommendations' accuracy based on users' access to the services. A value of 1, indicates that the RS model got all the predictions right, and a value of 0 indicates that the RS model did not make a single correct prediction.

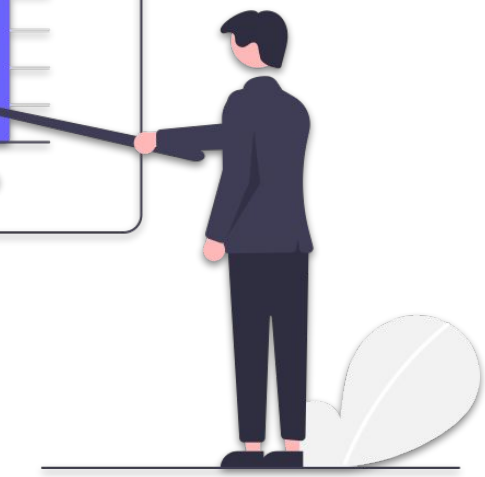


# What it offers?

## Metrics

- Accuracy
- **Catalog Coverage**
- Diversity Gini Index
- Diversity Shannon Entropy
- Novelty
- User Coverage

The percentage (%) of the division of the unique services found in recommendations to the total number of published services.

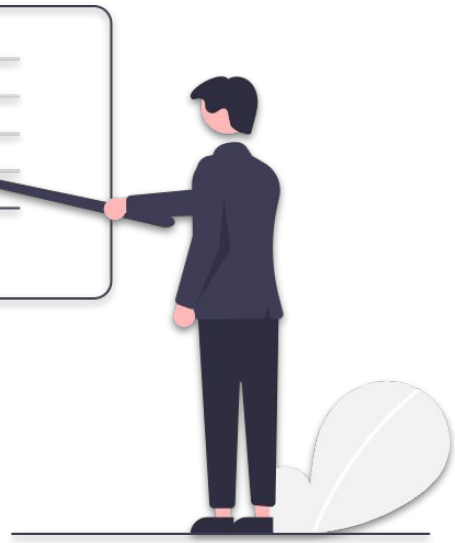


# What it offers?

## Metrics

- Accuracy
- Catalog Coverage
- **Diversity Gini Index**
- Diversity Shannon Entropy
- Novelty
- User Coverage

Measures Recommendations' diversity. The index is 0 when all items are chosen equally often, and 1 when a single item is always chosen.

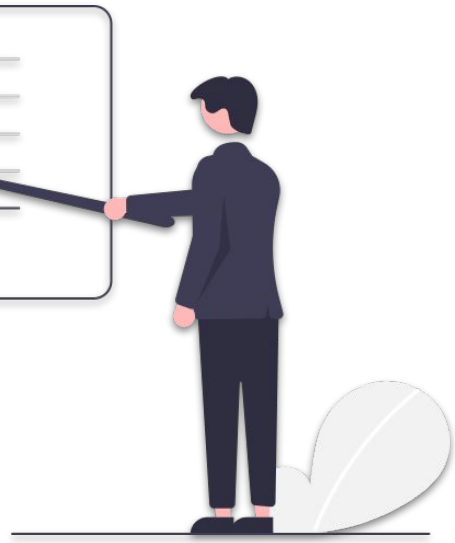


# What it offers?

## Metrics

- Accuracy
- Catalog Coverage
- Diversity Gini Index
- **Diversity Shannon Entropy**
- Novelty
- User Coverage

Measures Recommendations' diversity. The entropy is o when a single item is always chosen or recommended, and log n when n items are chosen or recommended equally often.

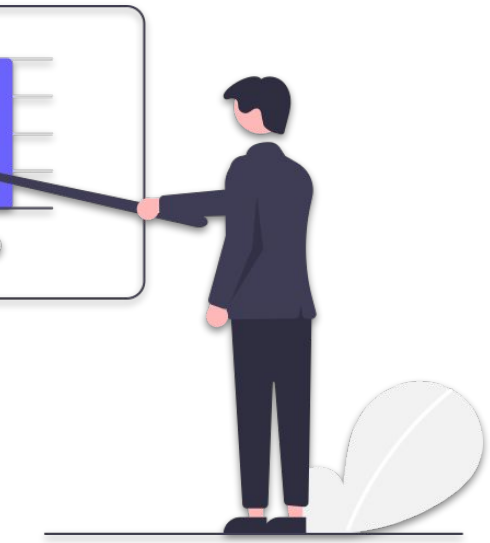


# What it offers?

## Metrics

- Accuracy
- Catalog Coverage
- Diversity Gini Index
- Diversity Shannon Entropy
- **Novelty**
- User Coverage

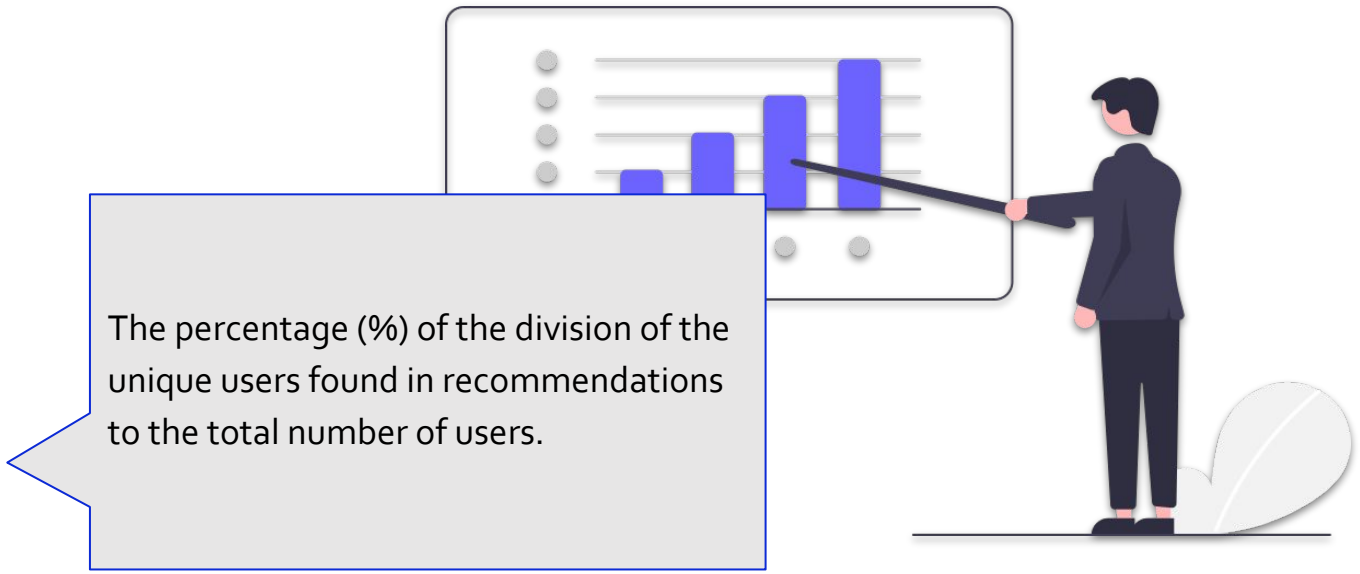
The novelty metric expresses the ability of the system to recommend items not generally seen before by the population of users.



# What it offers?

## Metrics

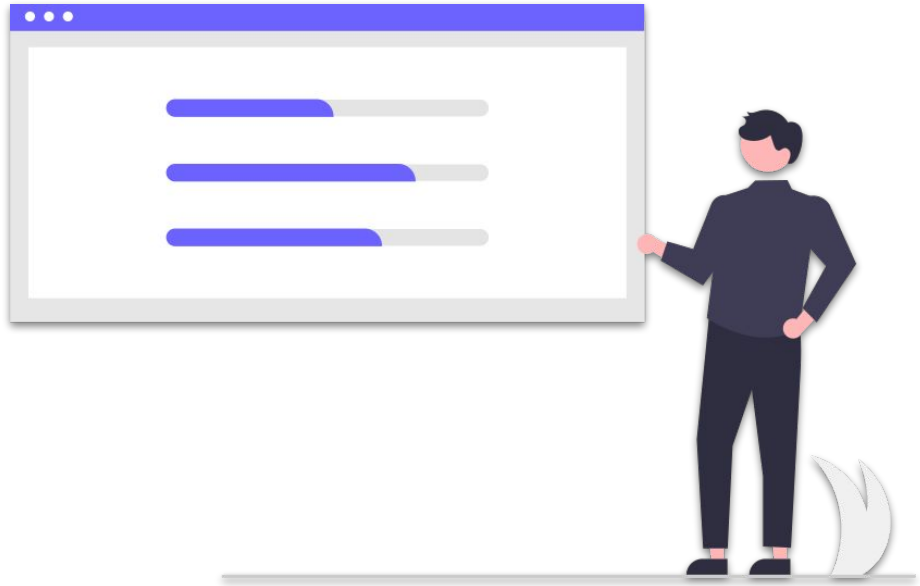
- Accuracy
- Catalog Coverage
- Diversity Gini Index
- Diversity Shannon Entropy
- Novelty
- **User Coverage**



# What it offers?

## KPIs

- Click-Through Rate
- Hit-Rate
- Top 5 viewed Items
- Top 5 recommended Items
- Top 5 viewed categories
- Top 5 recommended categories
- Top 5 viewed scientific domains
- Top 5 recommended scientific domains

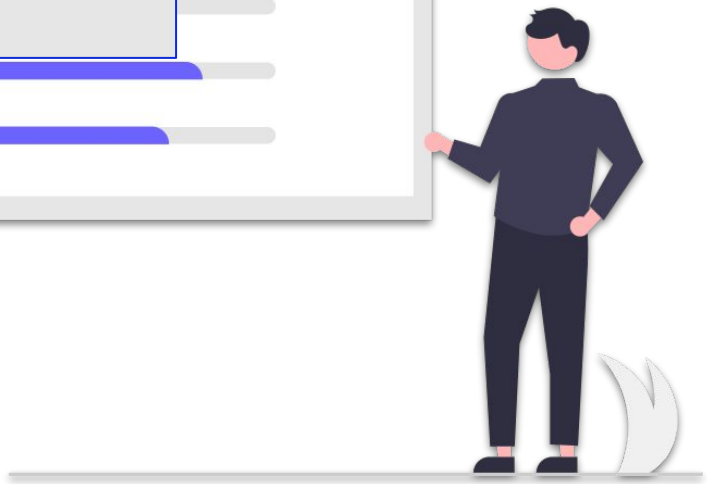


# What it offers?

## KPIs

- **Click-Through Rate**
- Hit-Rate
- Top 5 viewed Items
- Top 5 recommended Items
- Top 5 viewed categories
- Top 5 recommended categories
- Top 5 viewed scientific domains
- Top 5 recommended scientific domains

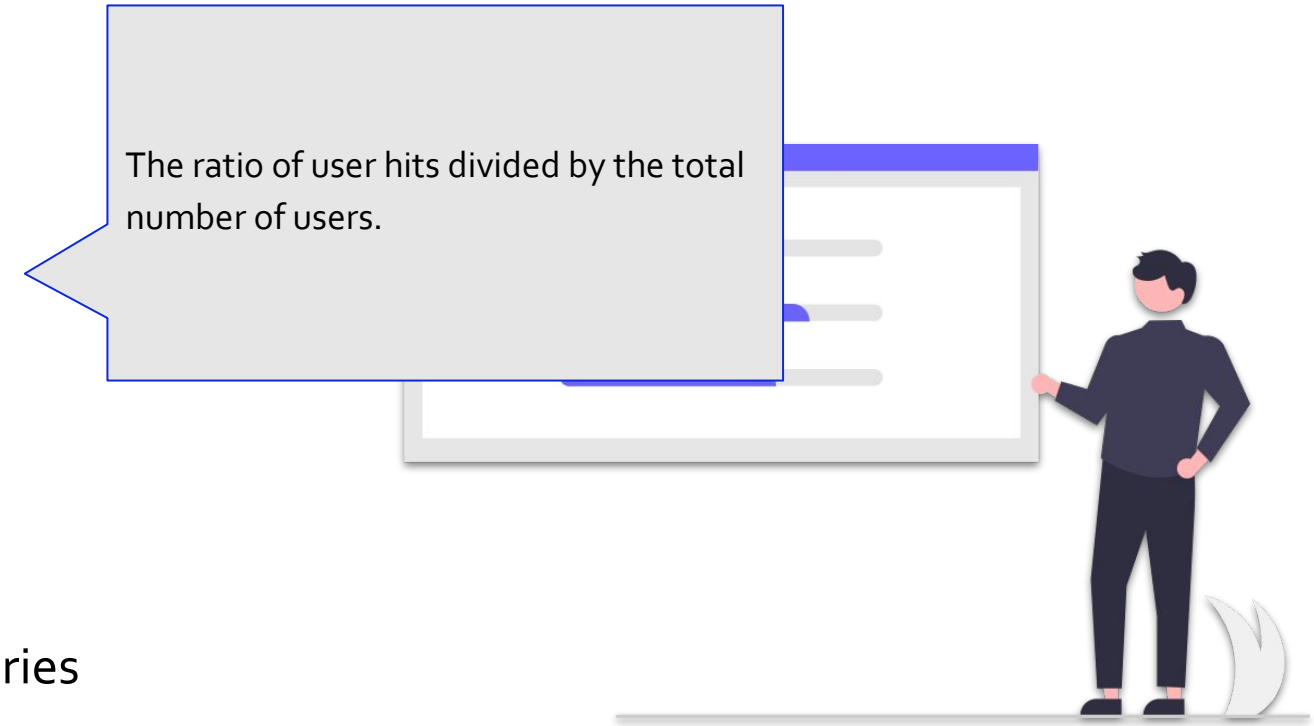
The number of user clicks through recommendations panels divided by the total times recommendation panels were presented to users.



# What it offers?

## KPIs

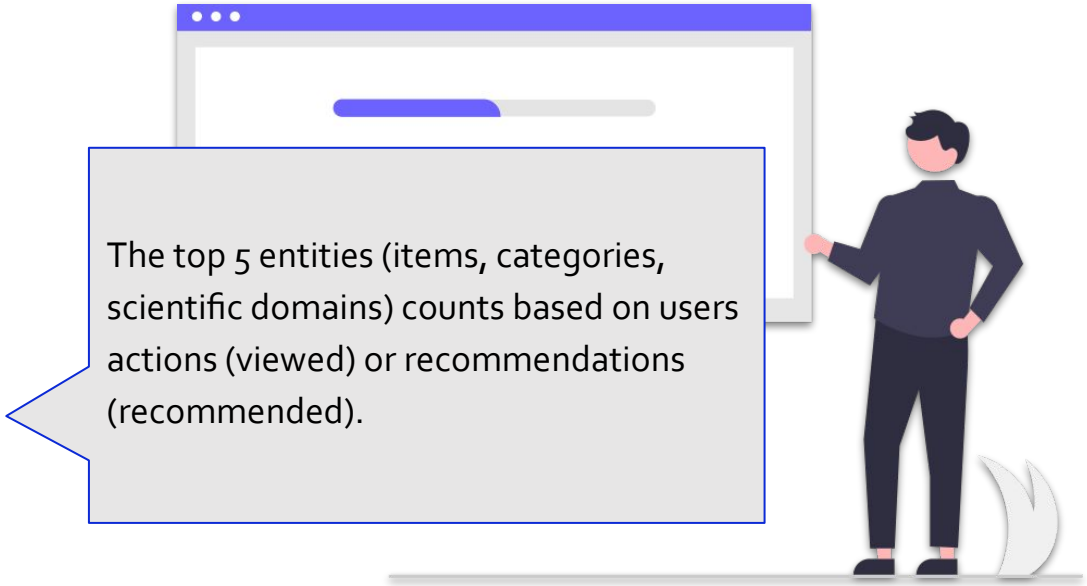
- Click-Through Rate
- **Hit-Rate**
- Top 5 viewed Items
- Top 5 recommended Items
- Top 5 viewed categories
- Top 5 recommended categories
- Top 5 viewed scientific domains
- Top 5 recommended scientific domains



# What it offers?

## KPIs

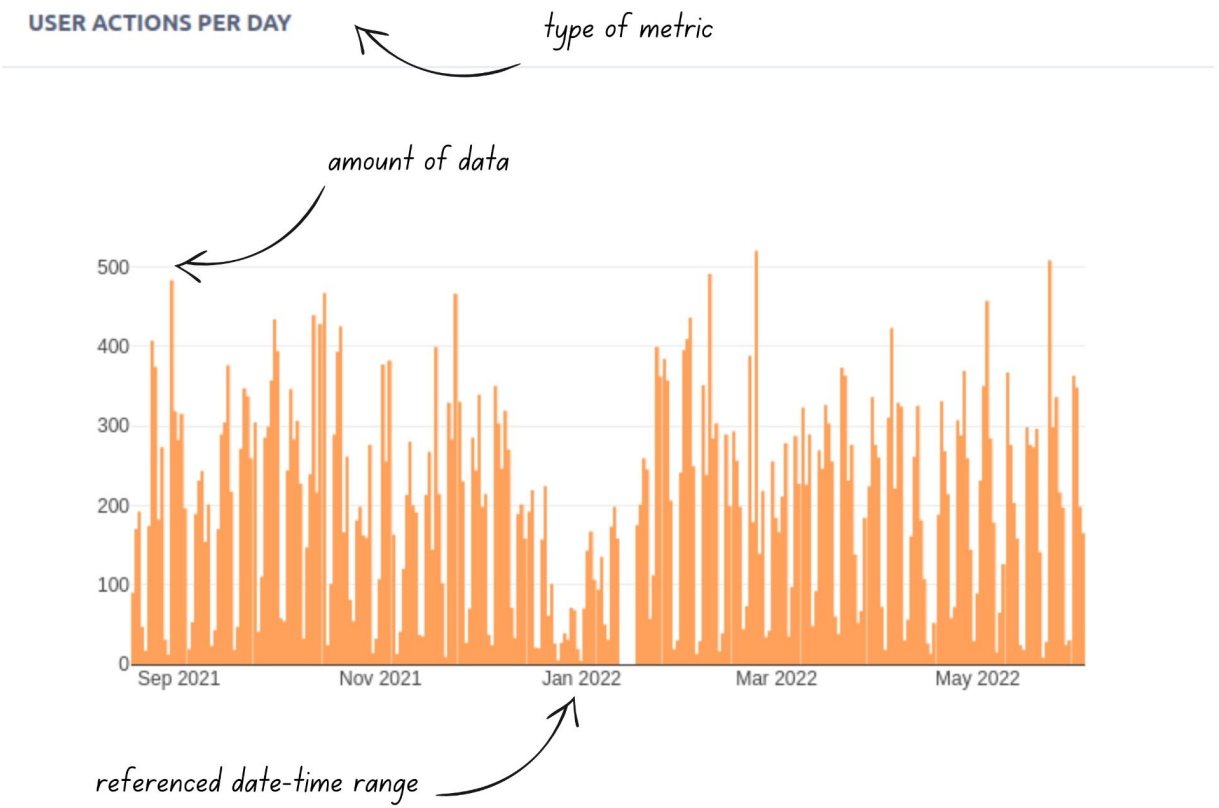
- Click-Through Rate
- Hit-Rate
- **Top 5 viewed Items**
- **Top 5 recommended Items**
- **Top 5 viewed categories**
- **Top 5 recommended categories**
- **Top 5 viewed scientific domains**
- **Top 5 recommended scientific domains**



# What it offers?

## Graphs

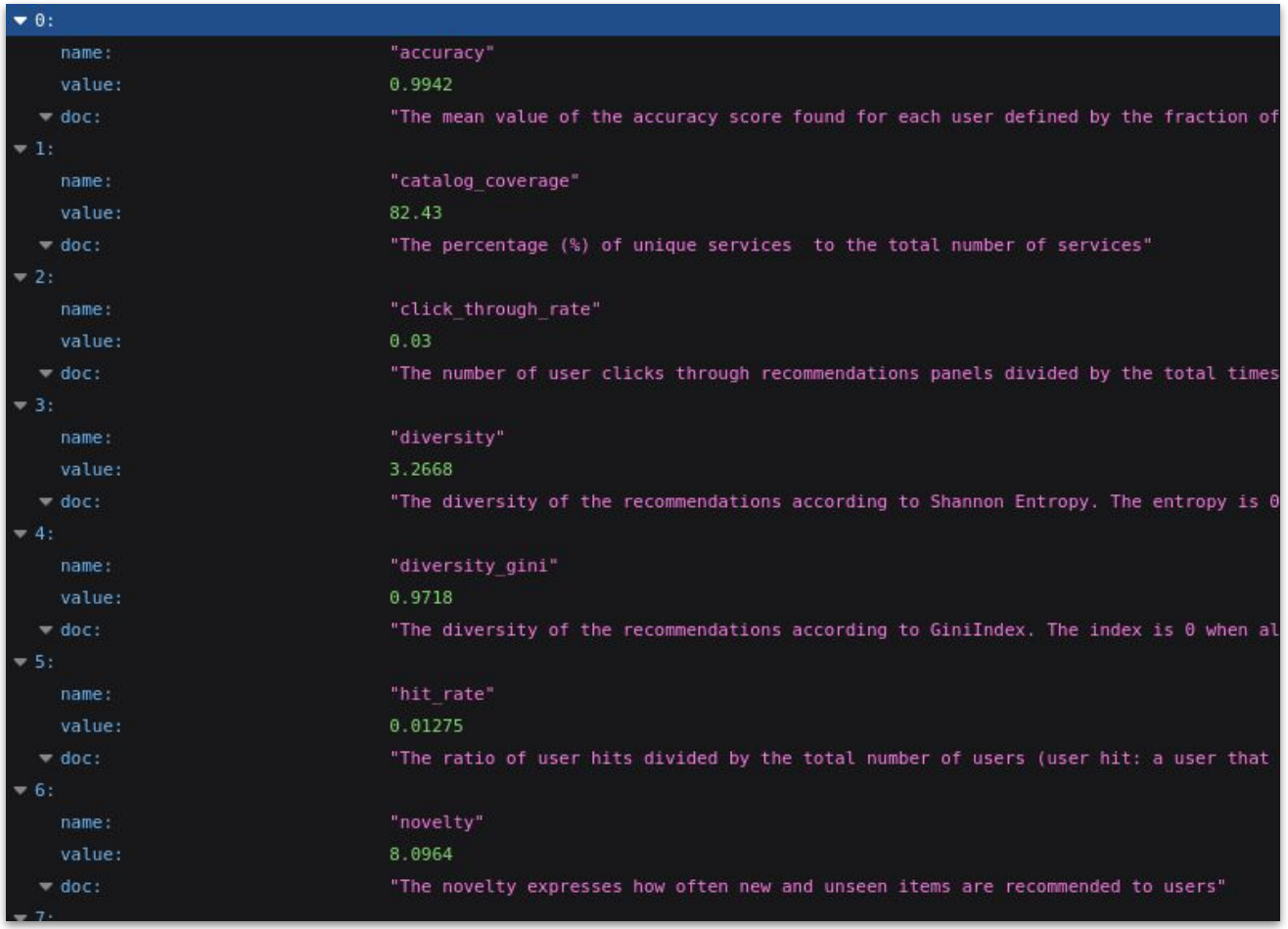
- User Actions per day
- Recommended Items per day
- User Actions per month
- Recommended Items per month



# What it offers?

## Rest API

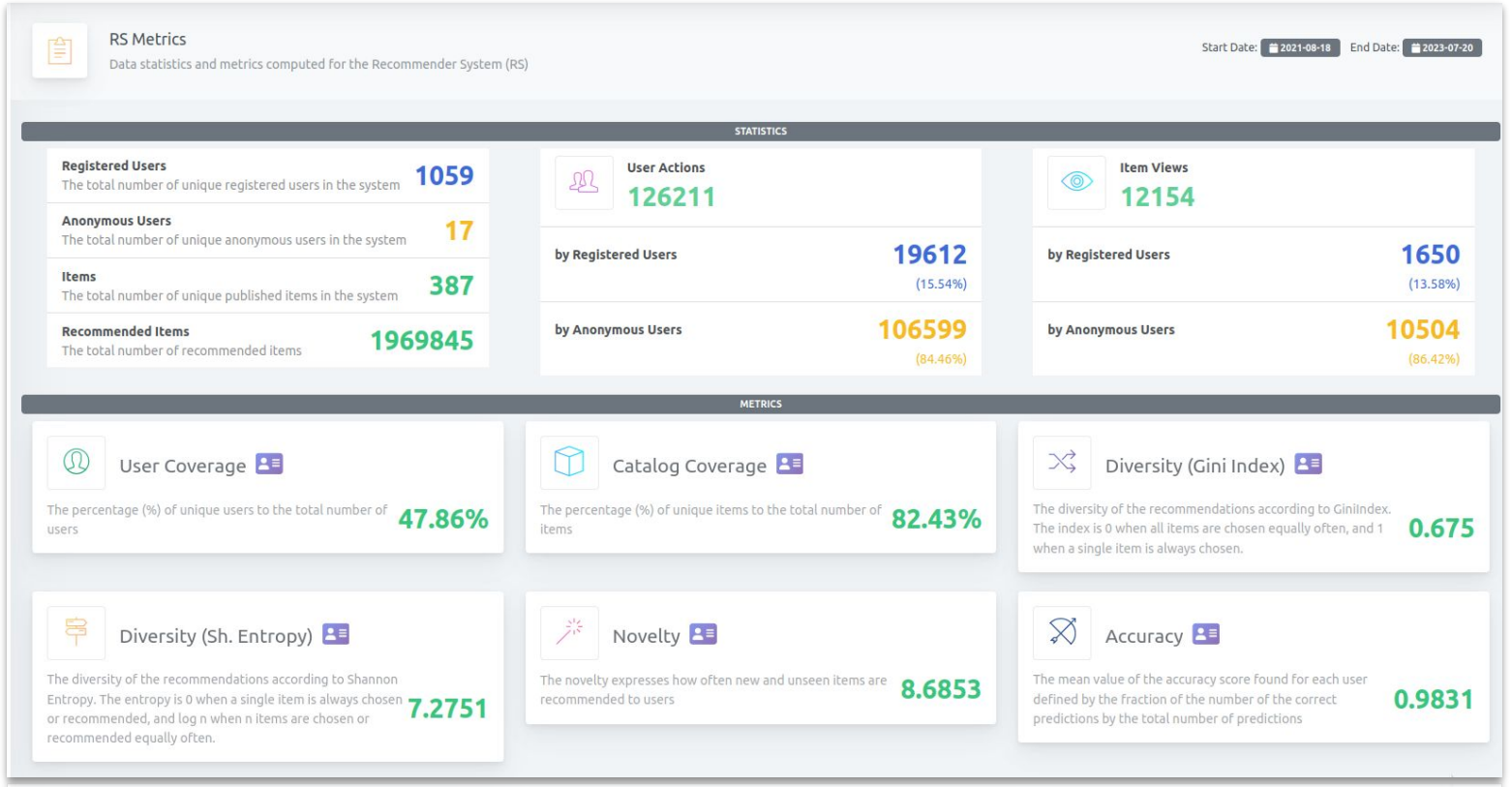
- Statistics
- Metrics
- KPIs
- Graphs' Data



# What it offers?

## UI Dashboard

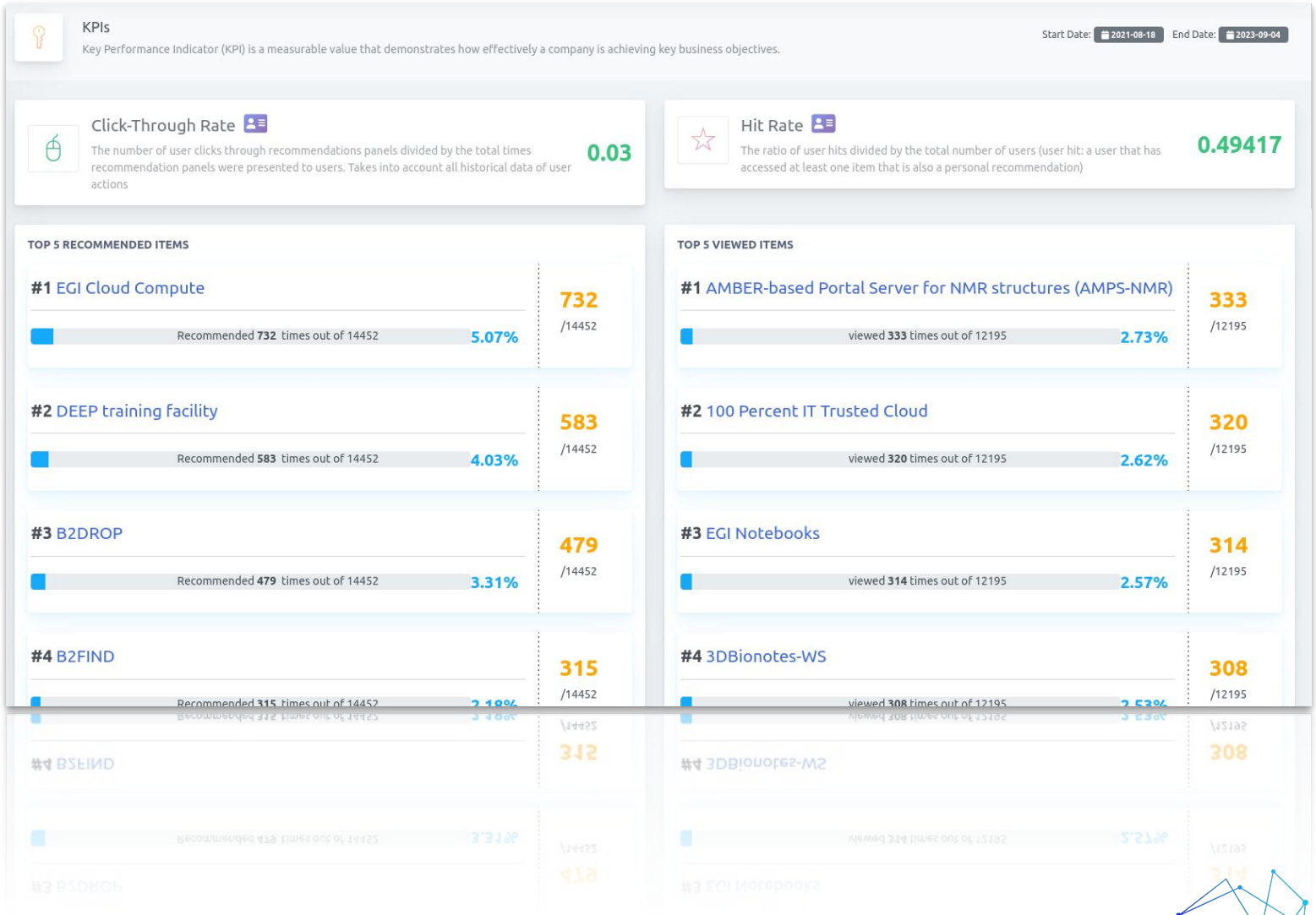
- Statistics
- Metrics
- KPIs
- Graphs



# What it offers?

## UI Dashboard

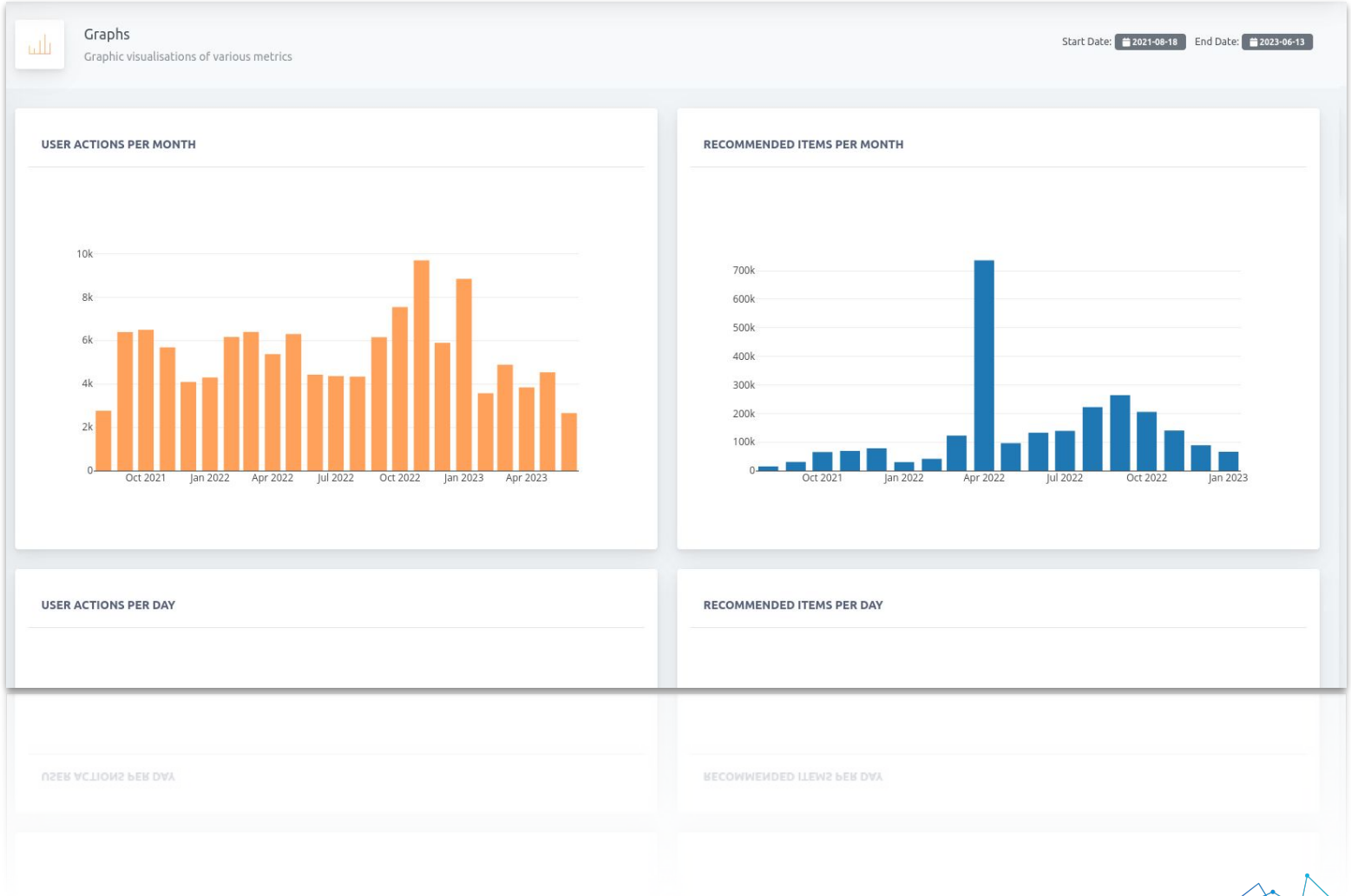
- Statistics
- Metrics
- KPIs
- Graphs



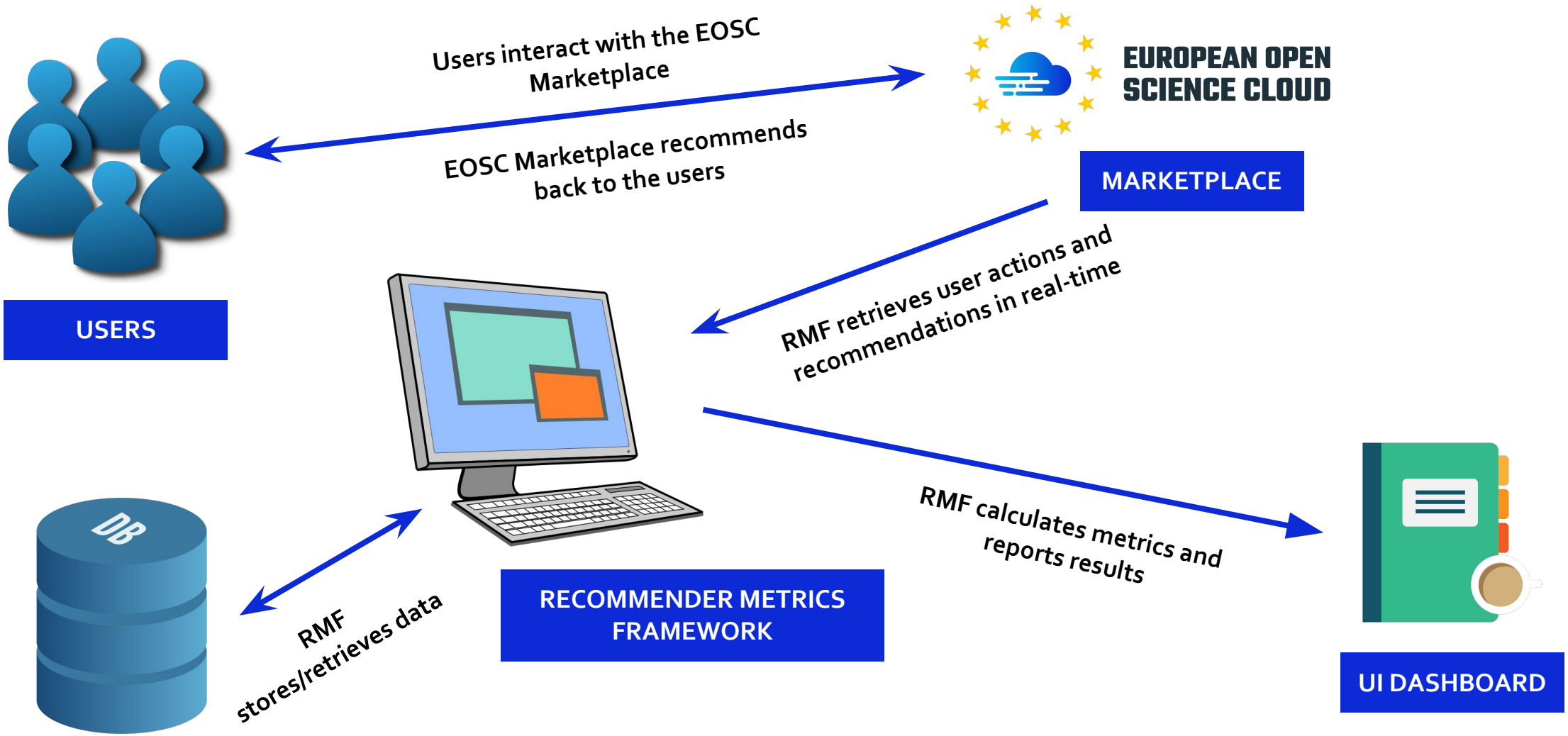
# What it offers?

## UI Dashboard

- Statistics
- Metrics
- KPIs
- Graphs



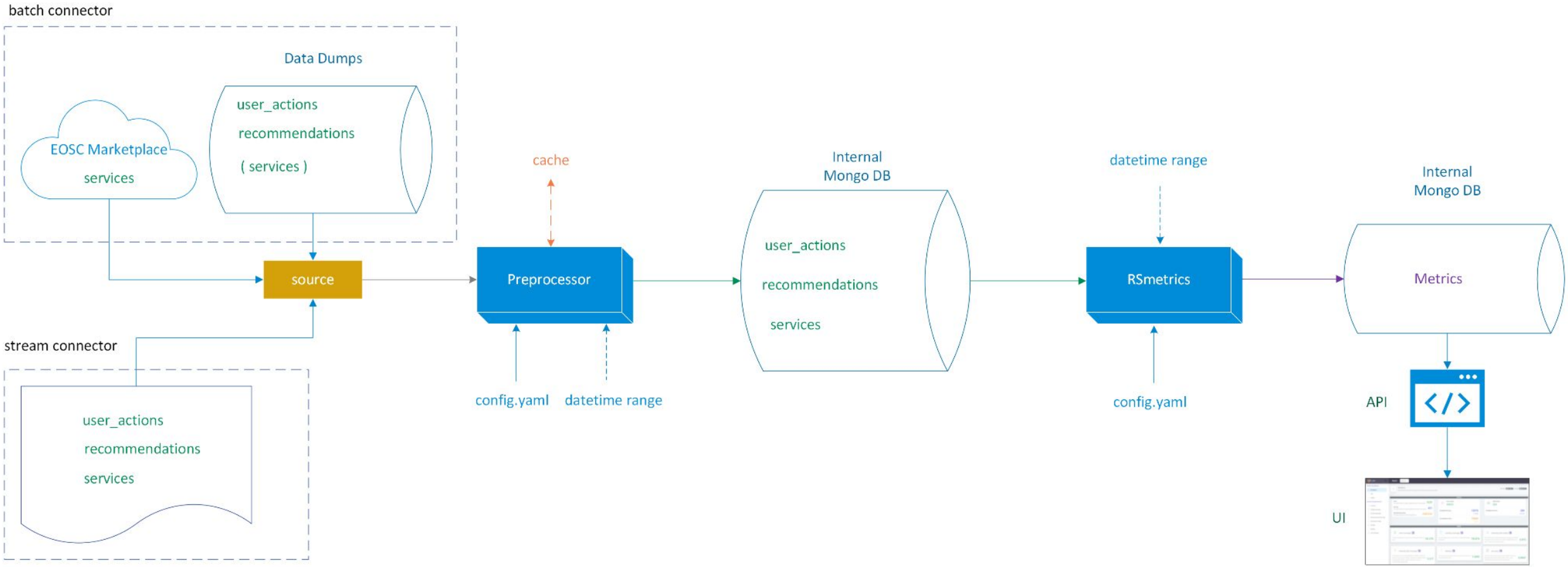
# Process Flow



# RMF Components

All Units

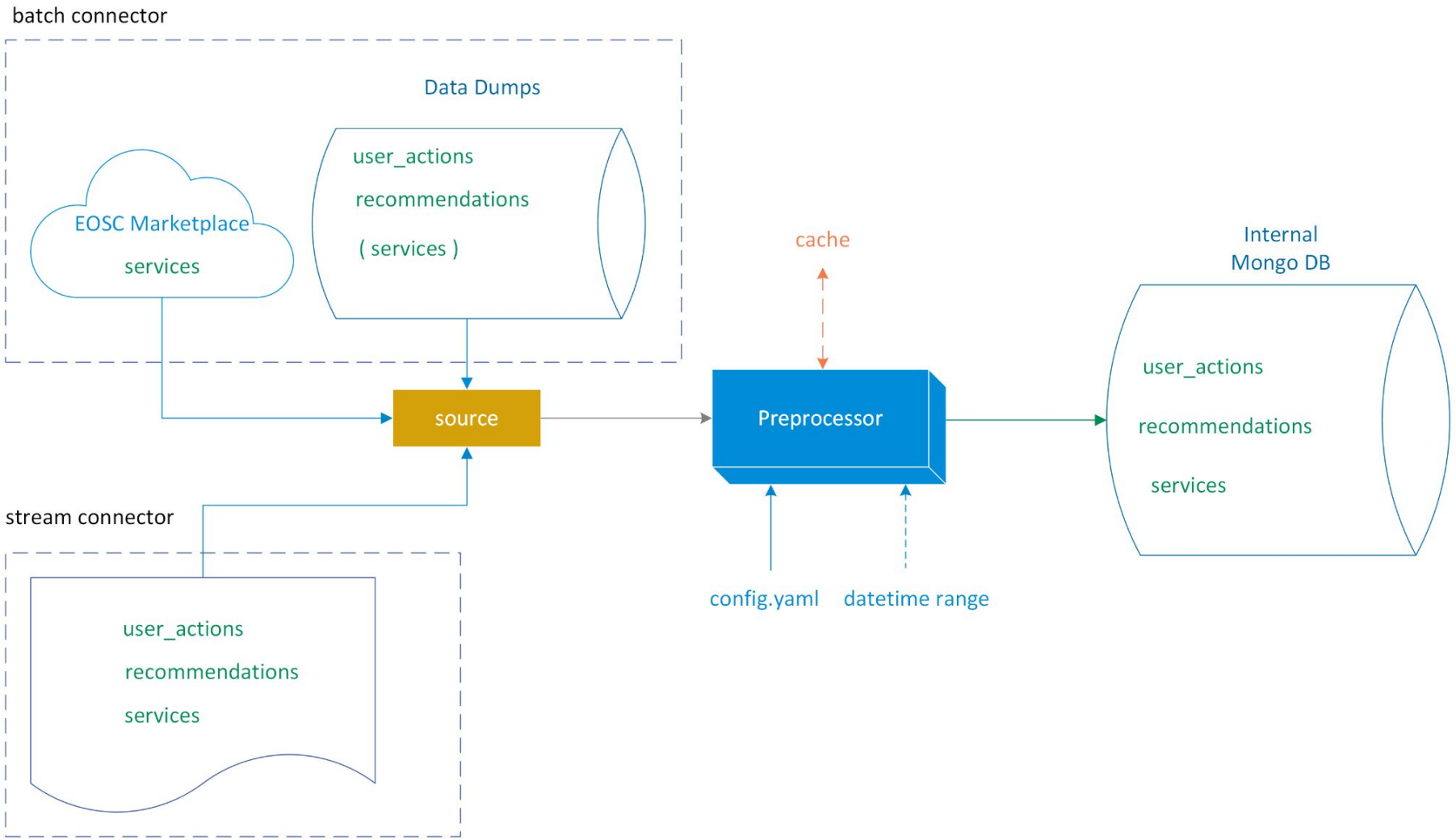
# Framework's components



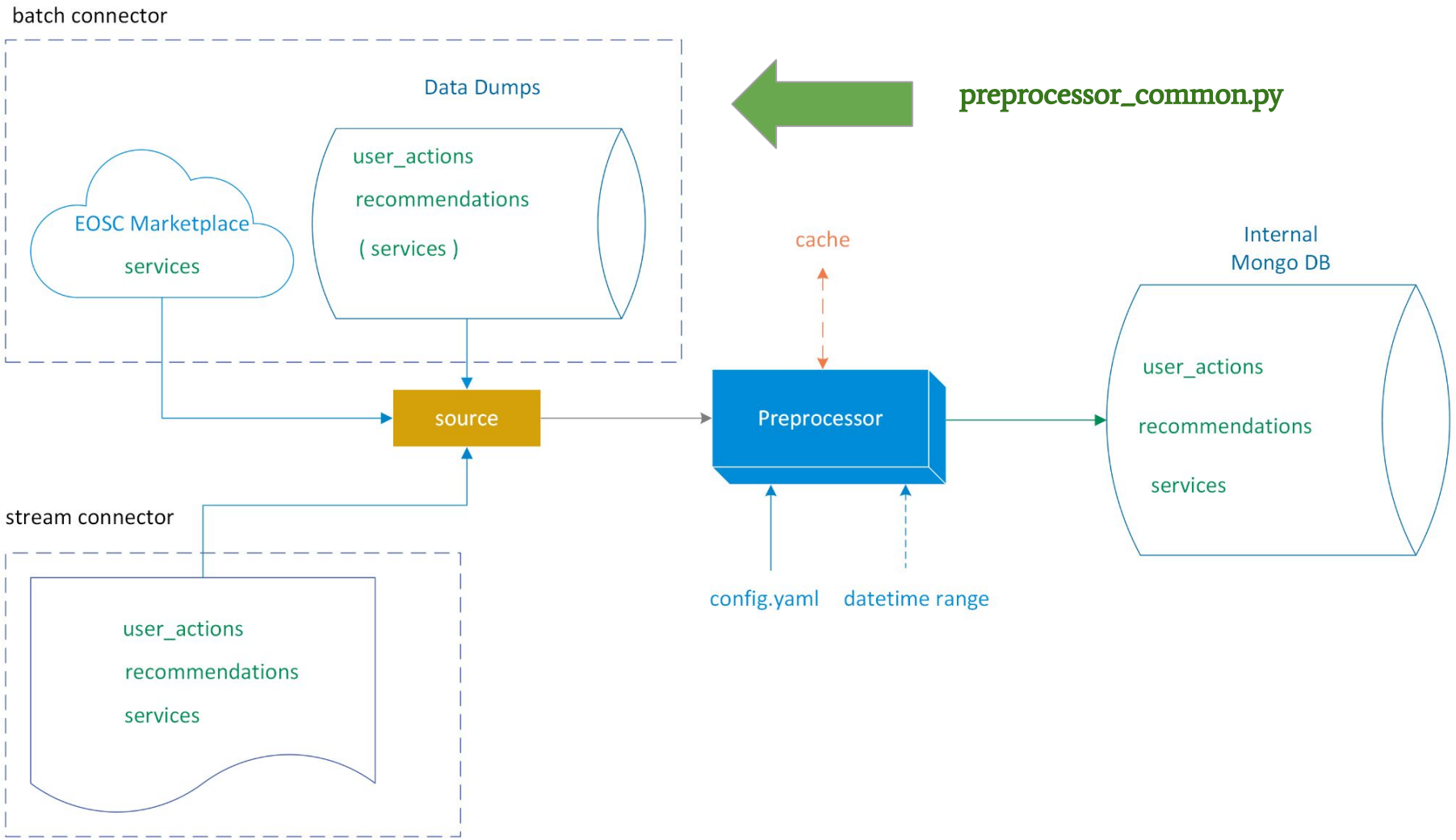
# RMF Components

Preprocessor Unit

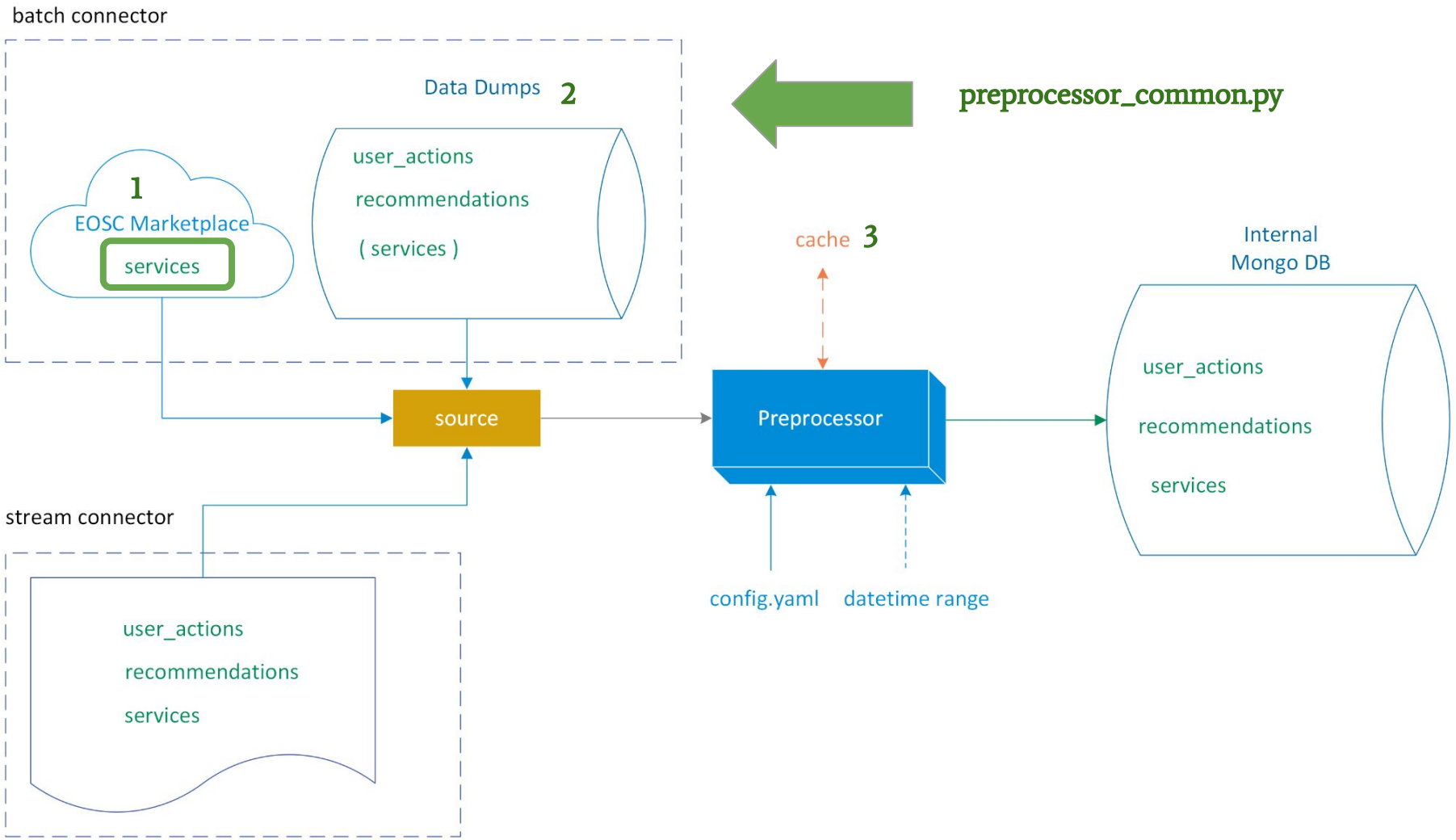
# Preprocessor Unit



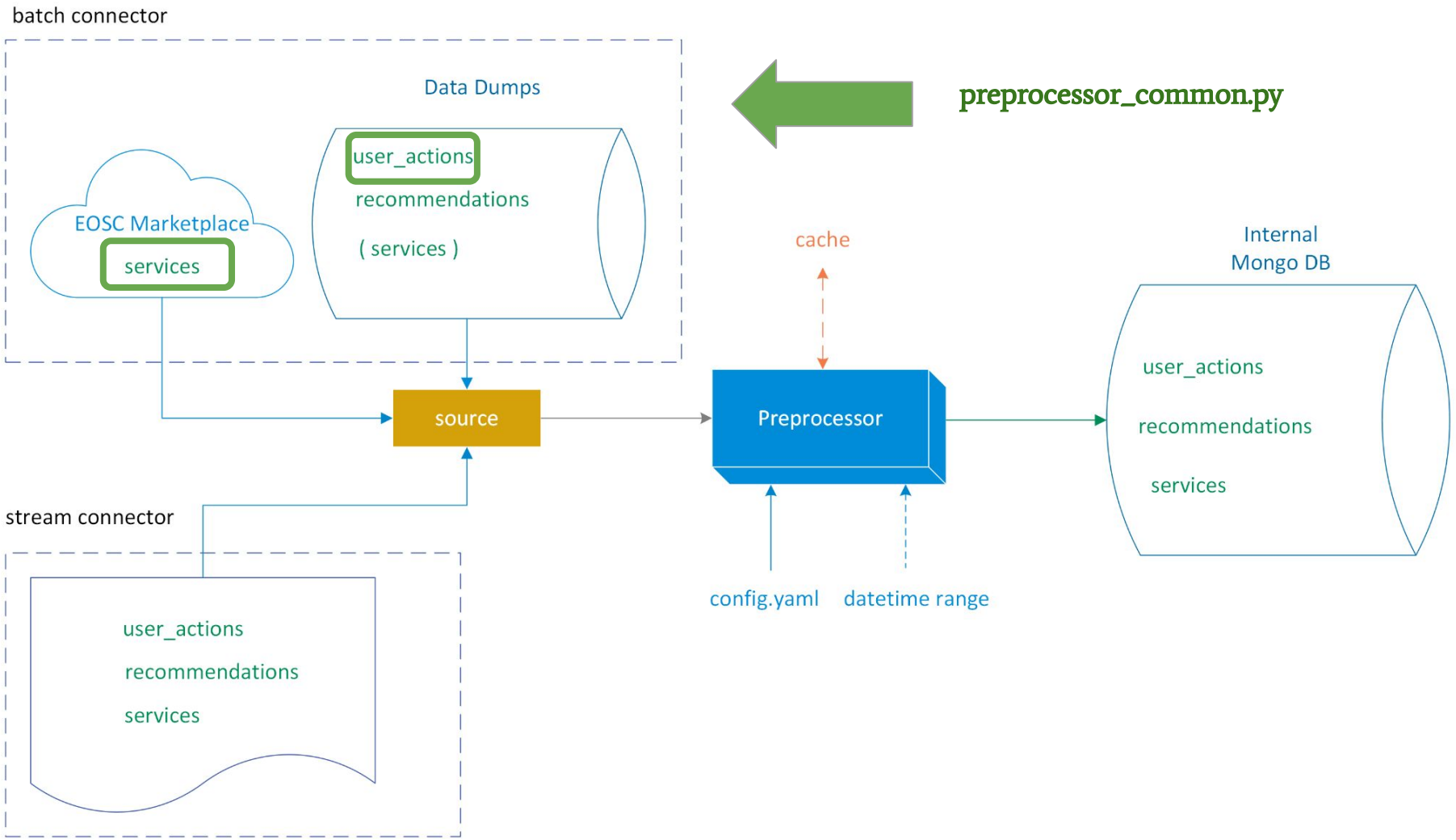
# Preprocessor Unit



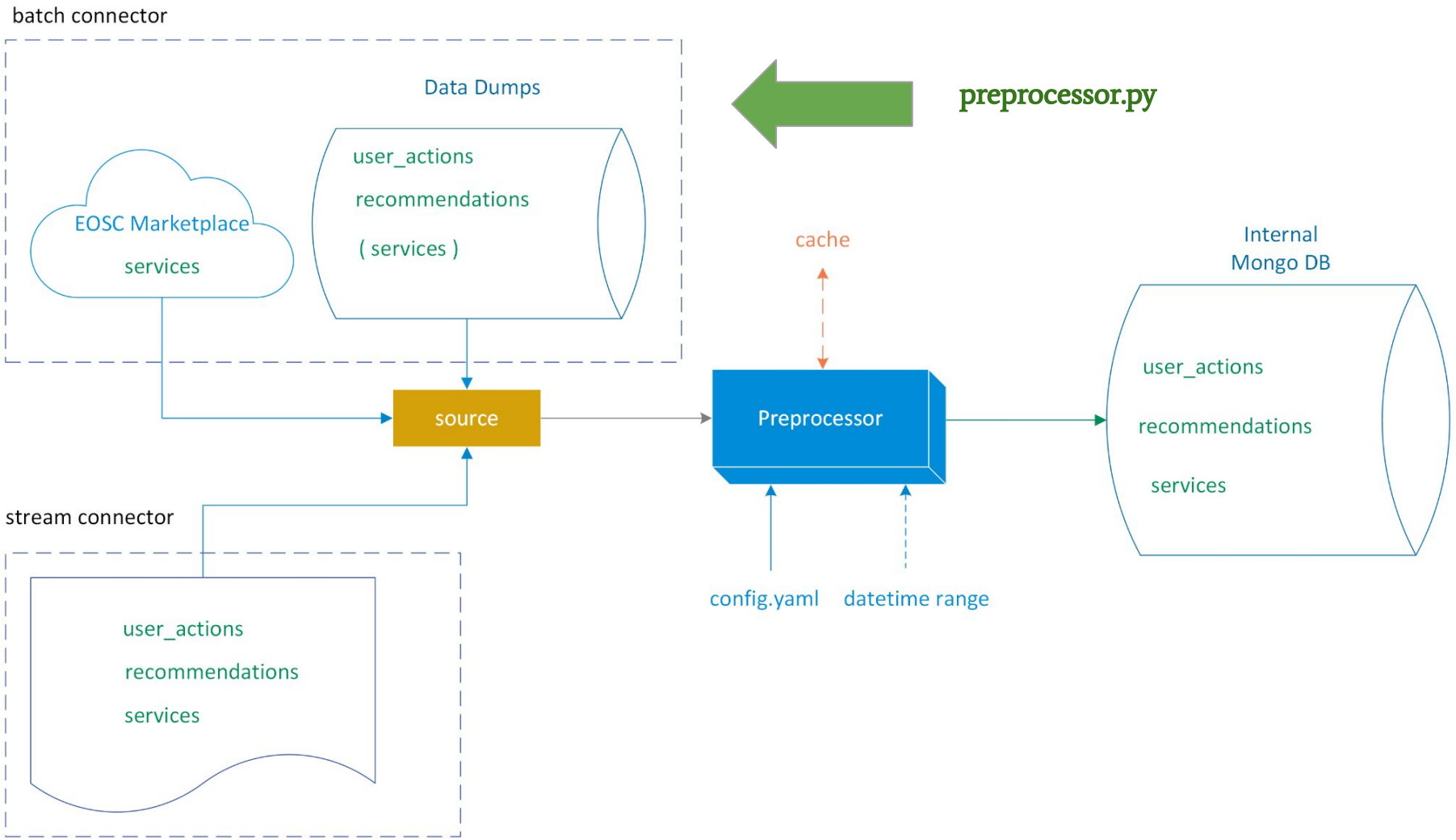
# Preprocessor Unit



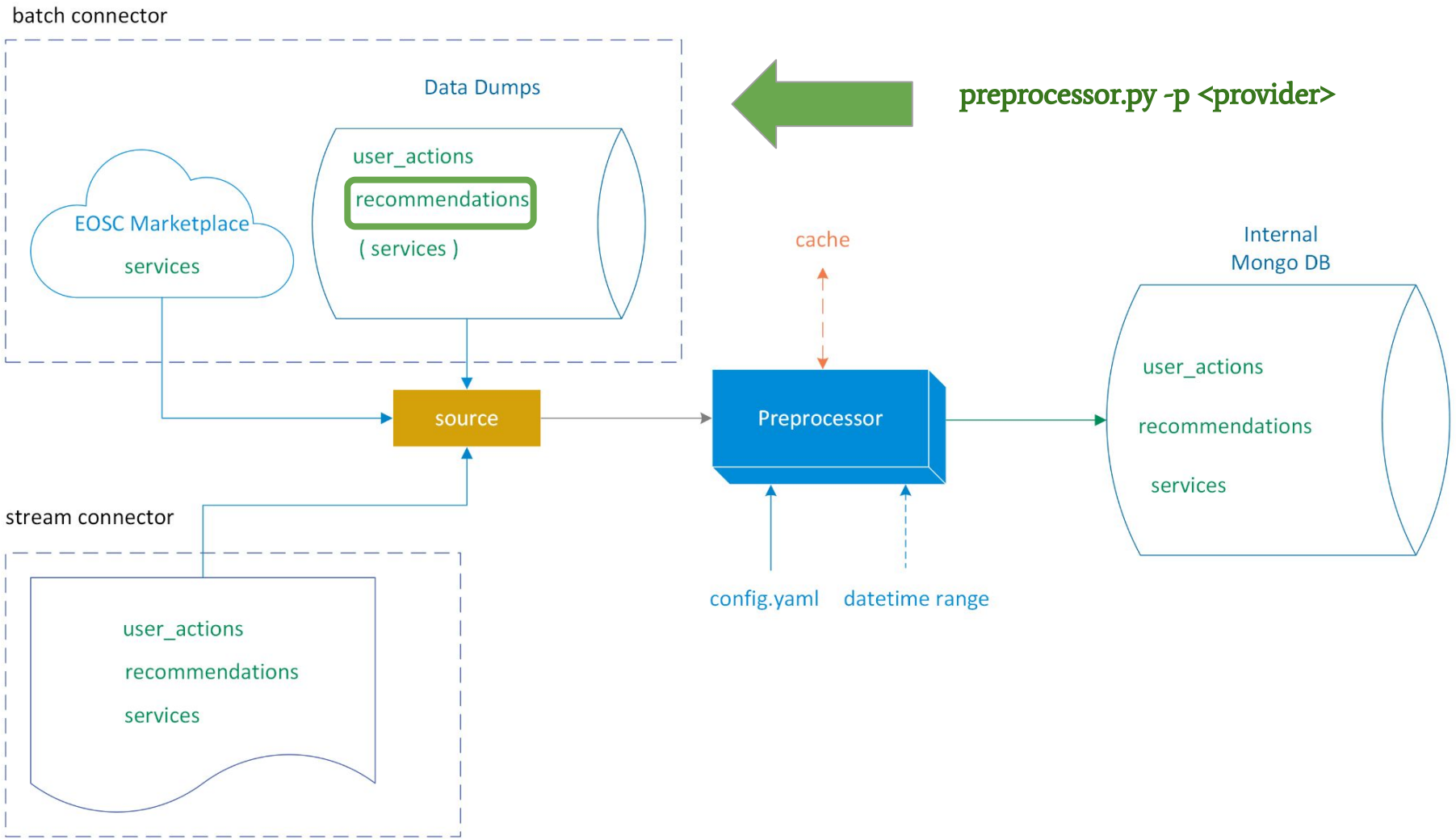
# Preprocessor Unit



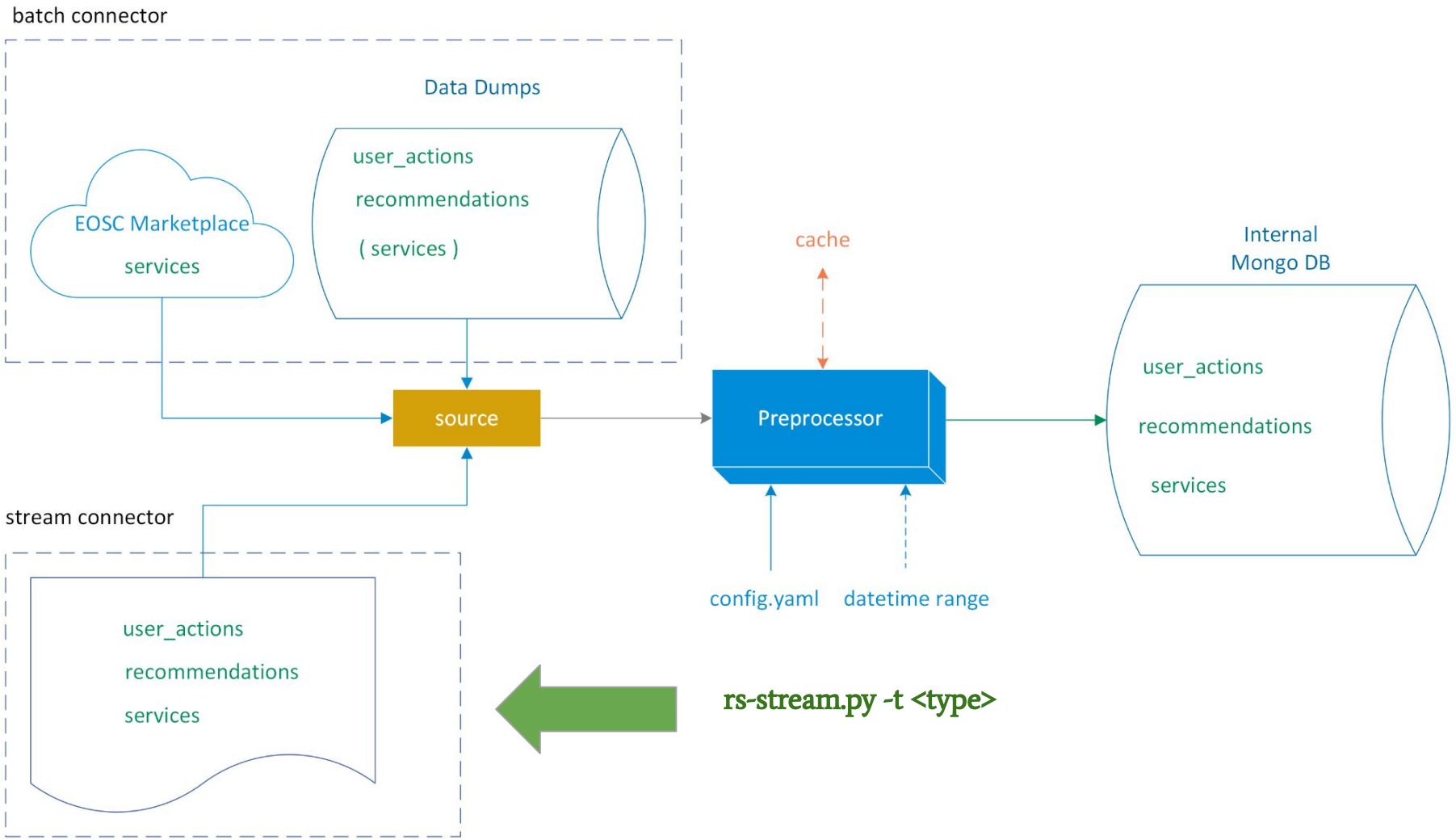
# Preprocessor Unit



# Preprocessor Unit



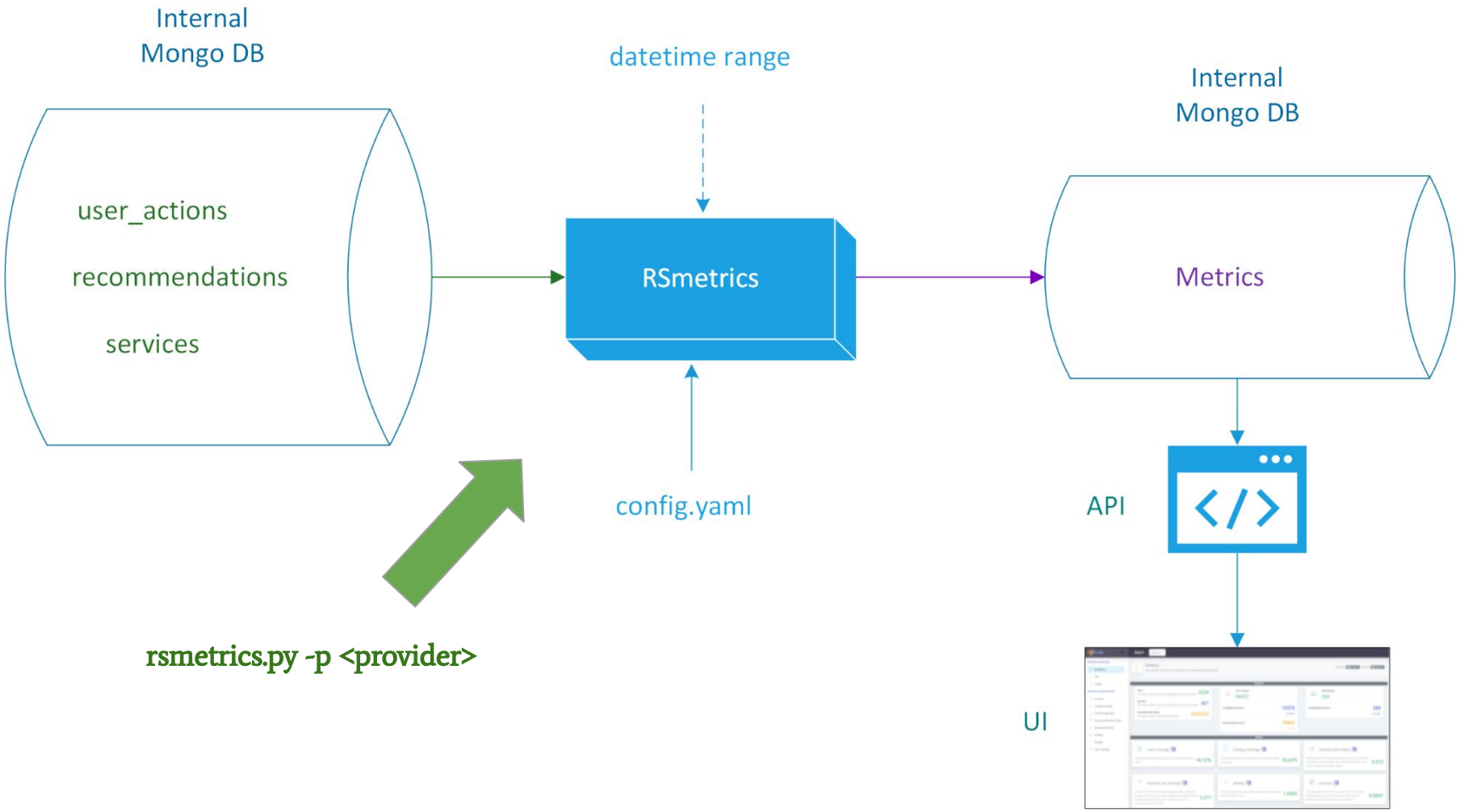
# Preprocessor Unit



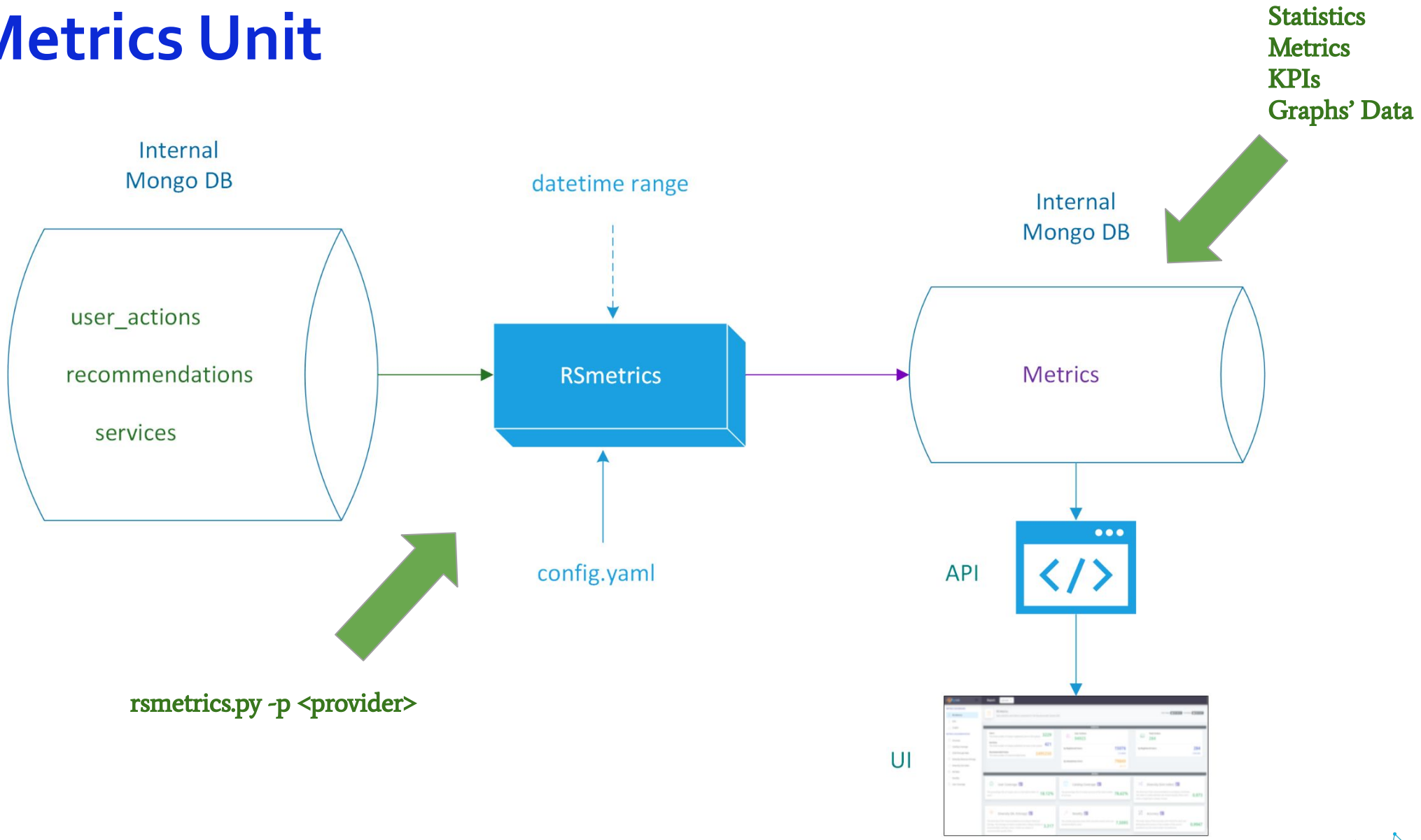
# RMF Components

RS Metrics Unit

# RS Metrics Unit



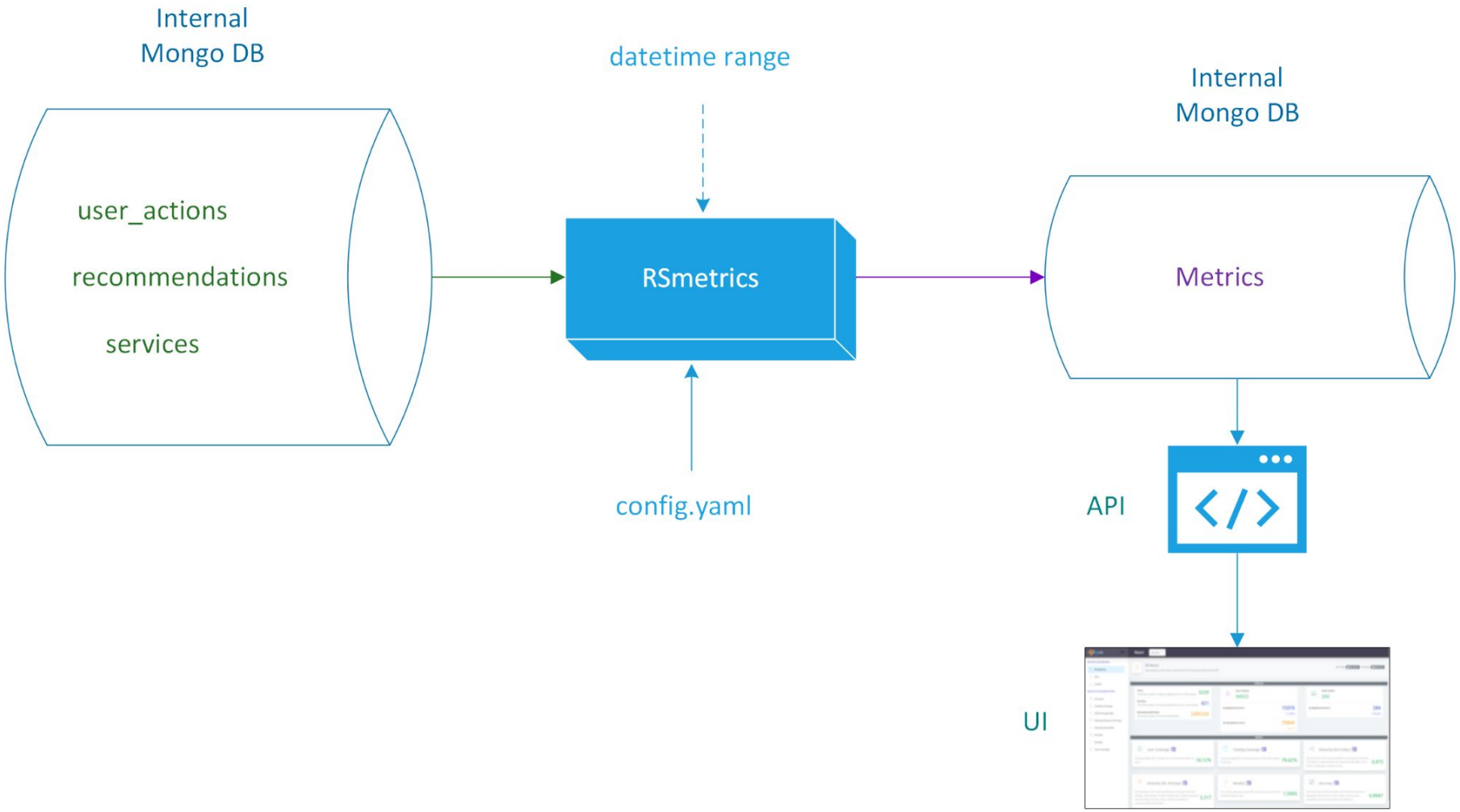
# RS Metrics Unit



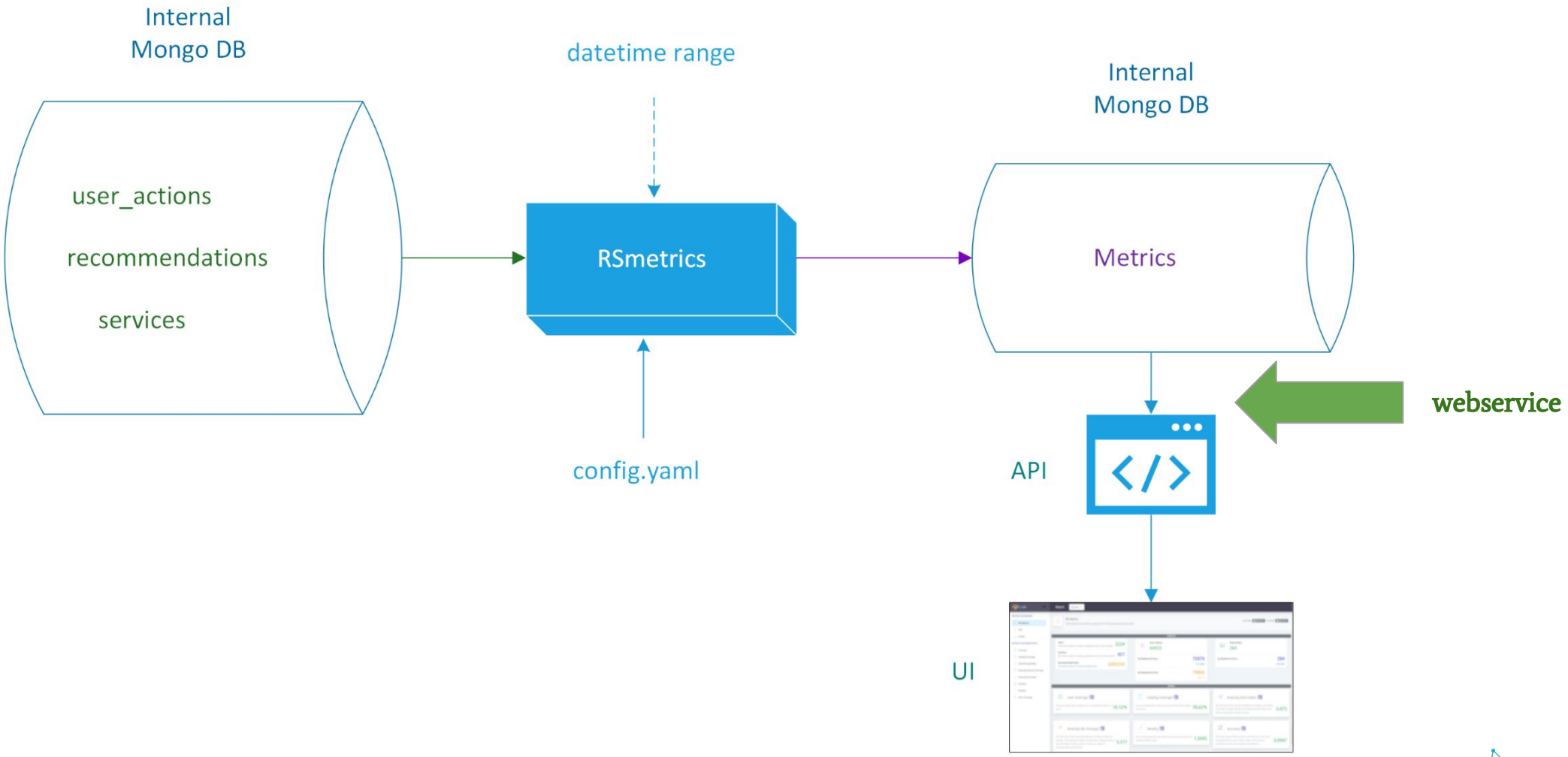
# RMF Components

Rest API/ UI Dashboard Unit

# REST API/UI Dashboard Unit



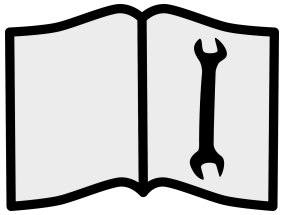
# REST API/UI Dashboard Unit



# More to see ...

overview

https://github.com/ARGOeu/eosc-recommender-metrics



https://argoeu.github.io/eosc-recommender-metrics/docs

# Ready to answer your questions!

# Thank you for your attention

The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement 101017536