A quasi-real Music Recommender System

Pietro Morichetti

What? Why? How?

What is a Recommender System?

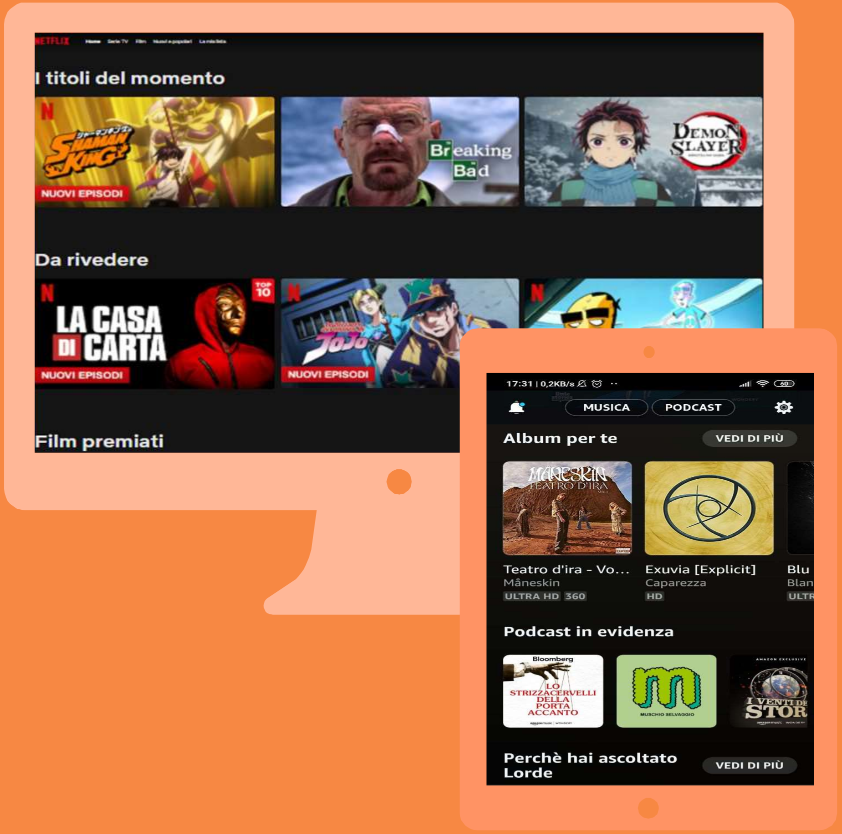
A recommendation system is an algorithm that aims to provide suggestions to users, based on knowledge provided by the users about their personal interests.

Why a Recommender System?

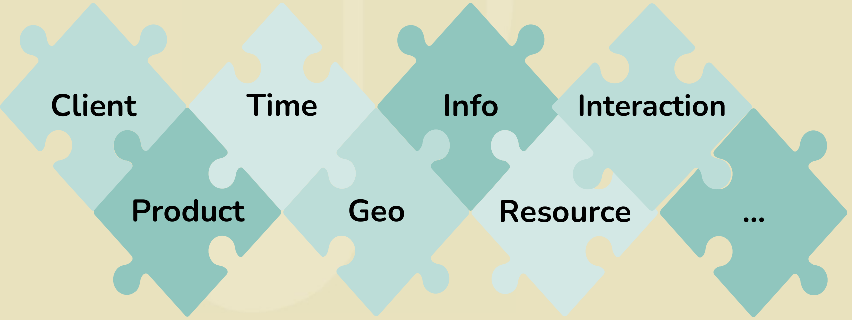
- In the on-line, purchases are exponentially increased
- A great market strategy is based on the customers
- A lot of products... no for real a **LOT** of products

How create a Recommender System?

- It depends on...



. . .The situation


Client

Time

Info

Interaction

Product

Geo

Resource

...

CONTENT RS

Just user/item features are given

COLLABORATIVE RS

Customer are supported considering also interests of the other customers

MODEL BASED

Define a model based on the past user-item interaction

MEMORY BASED

Similarity based

Road Map

01

Introduction

Description of the «situation»

02

Quasi-real

A Quasi-well structured music platform

03

R.S.

One RS it's okay, two RS is great... but four it's awesome!

04

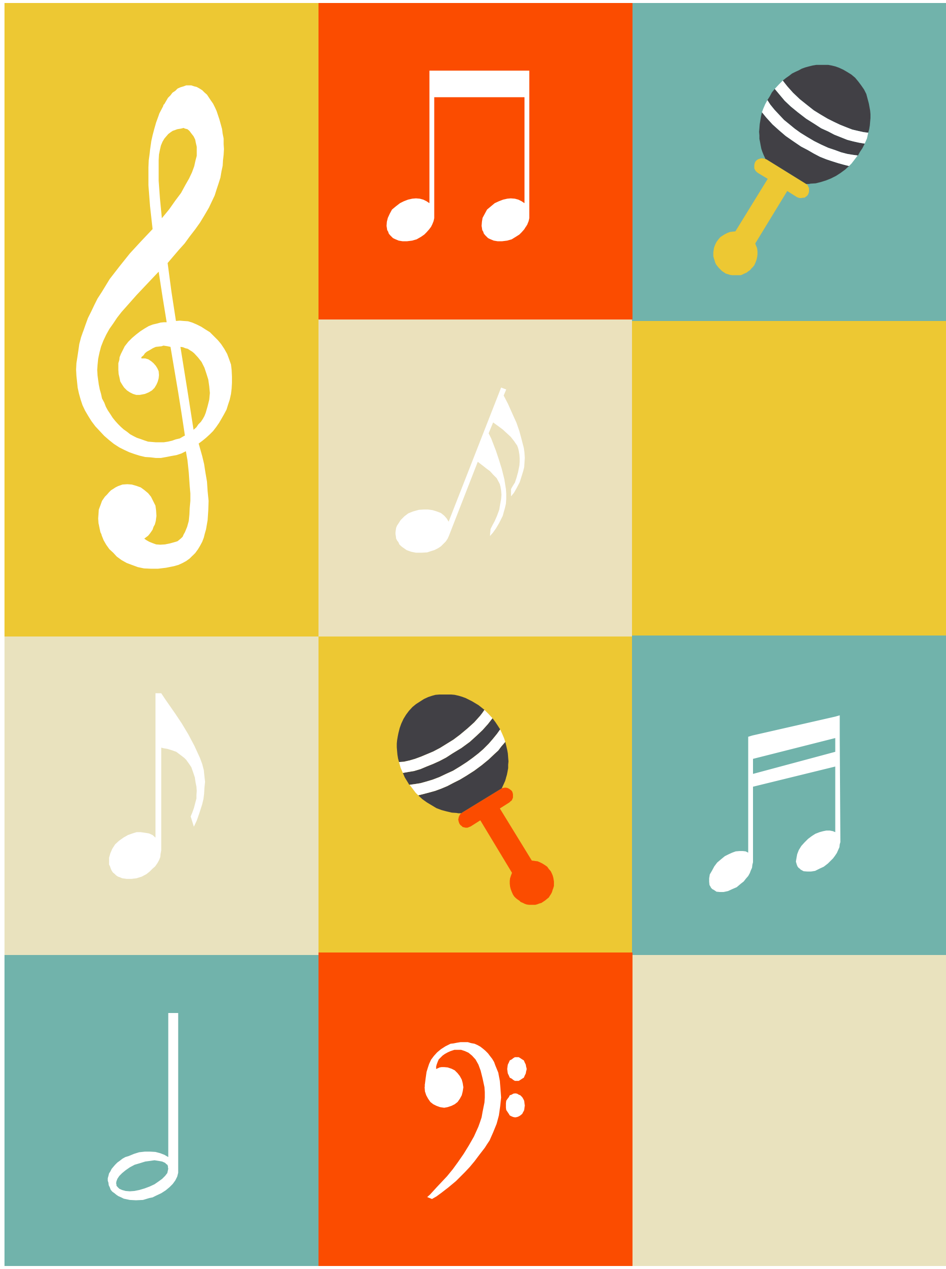
Conclusion

Shall we play the next song?

01

Introduction

- The situation
- The data



Our Situation

To develop this project were considered different information but, unfortunately, by different separated sources...

Kaggle ✓ Github AmazonMusic YahooMusic ✓ LastFM

Let's play a bit of.. Music!

users

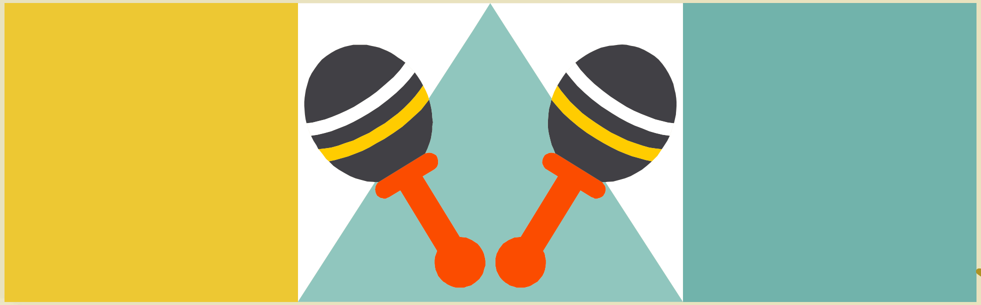
- Set of unique users

songs

- Set of unique songs

ratings

- Range [1, 5]



Preprocessing

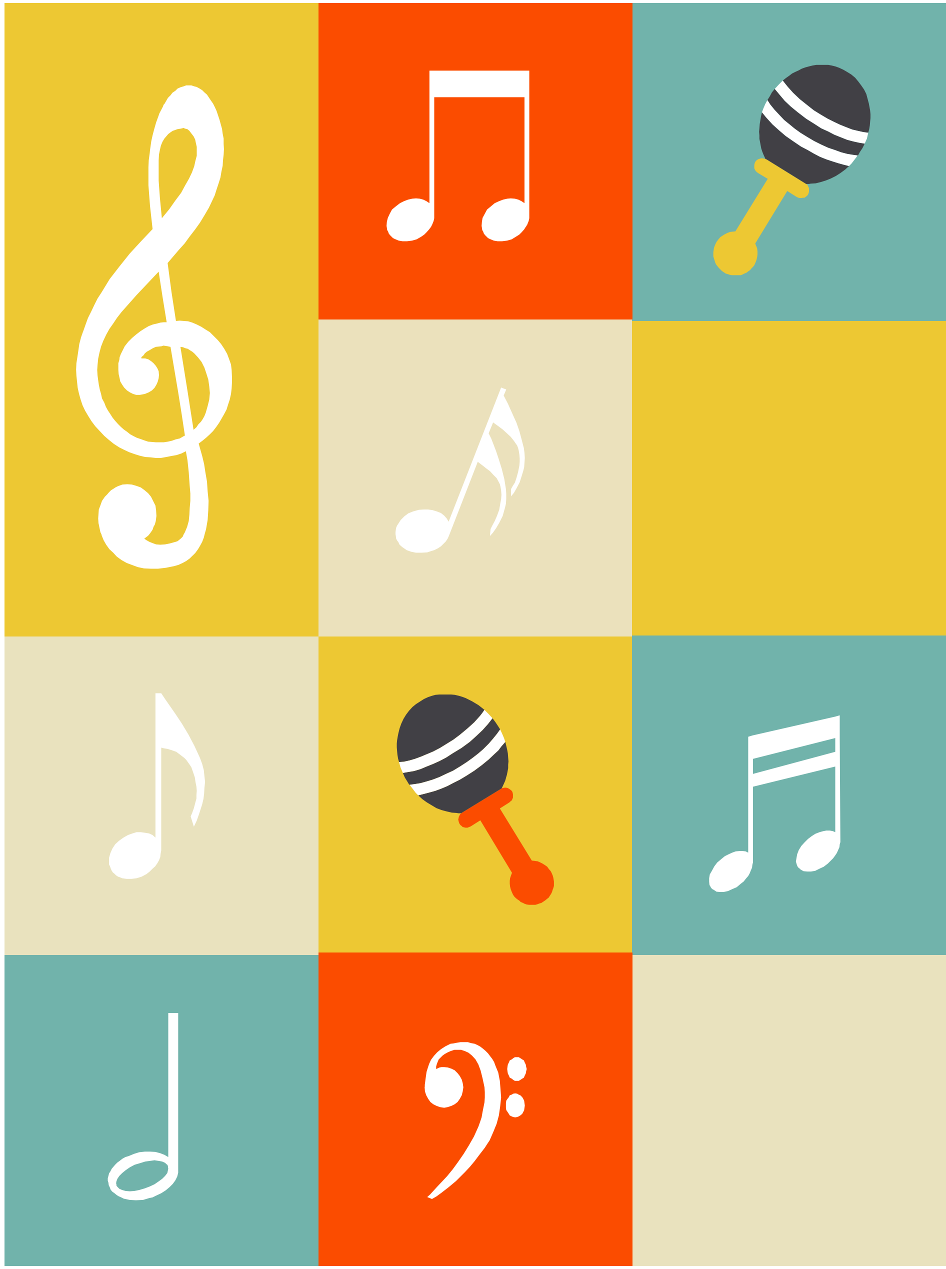
Dataset	Columns			
songs	song_id	author	genre	...
user	user_id	username	password	...
history	user_id	song_id	repetition	...
ratings	user_id	song_id	ratings	-

An artificial composition of datasets.

02

That's …
Quasi-real

- What does it means «quasi»?
- Software design point of view



Not just a RS.. but much more

Design a RS it's cool, but what about all the rest? A **real** music platform is composed by other components...

T

Users, song, history handled by means of tables

D

Persistent, available and consistent, let's use a database

E

Exception occurs everywhere, in music and in this project

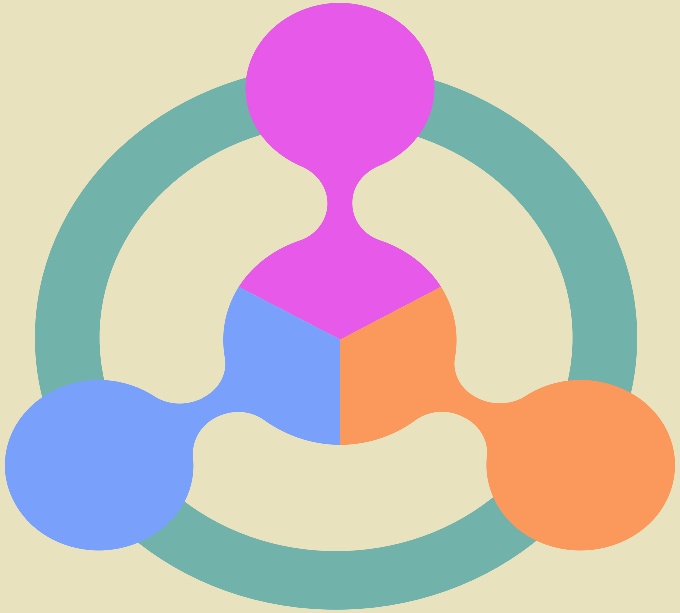
A

You can find «A» in Market as well as Analytics

Not just a RS.. but much more

Analytics

- Error 404 -
Analytics not found

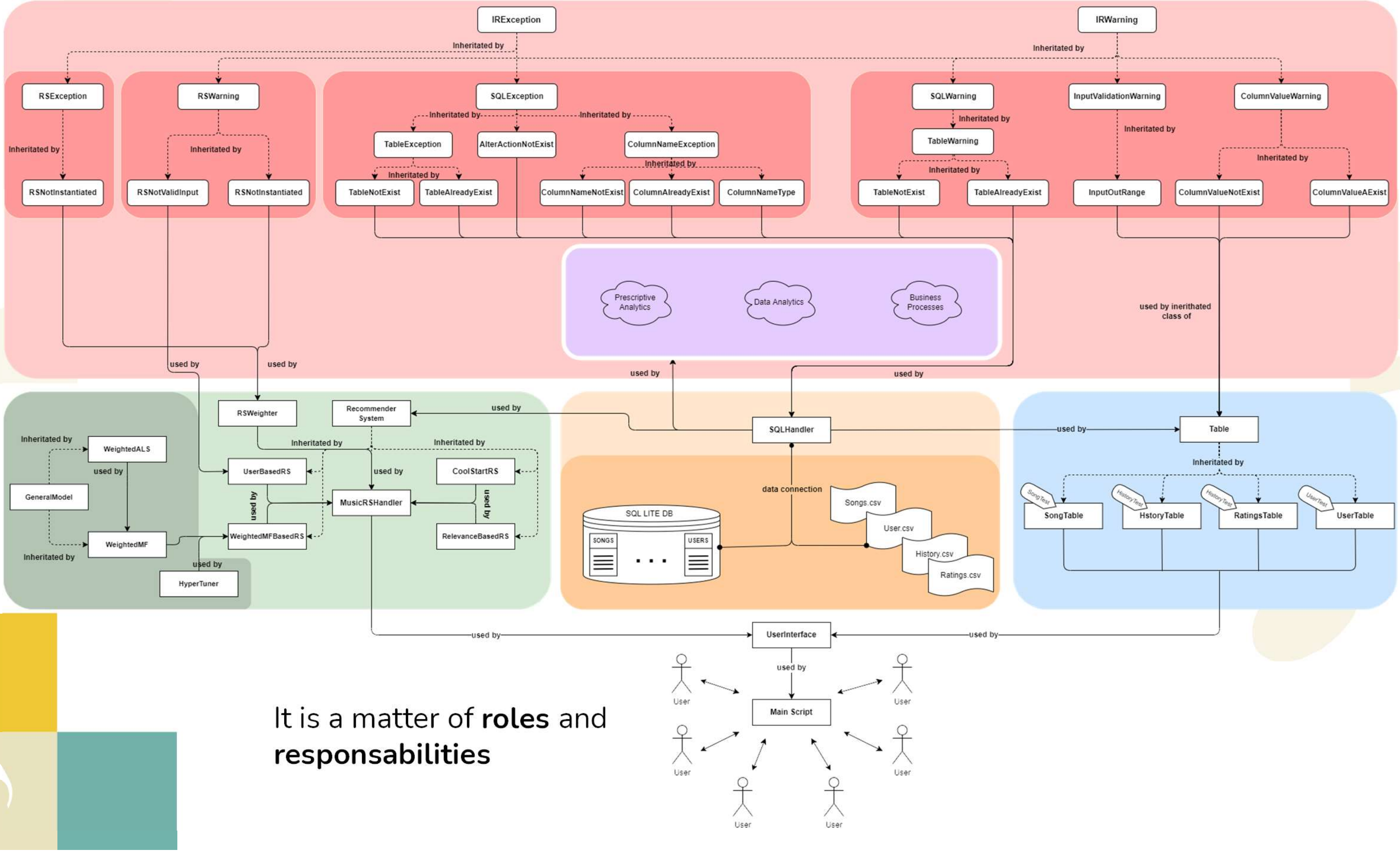


Table

- Data represented by a table
- Table functionality
- No direct connection with datatabe

Exception

- Hierarchical Exception structure
- Exception for database
- Warning for tables



It is a matter of **roles** and **responsabilities**

03

We are the R.S. Band

Relevance RS

CoolStart RS

WeightedMF RS

UserBased RS

The R.S. Band

Relevance RS

- Content Based RS
- Focused on some songs features
- Almost «dynamic»

CoolStart RS

- Content Based RS
- Focused on the songs features
- Mostly «static»

WeightedMF RS

- Collaborative Based RS
- Focused on user-song interaction
- Implemented by means of **WALS**

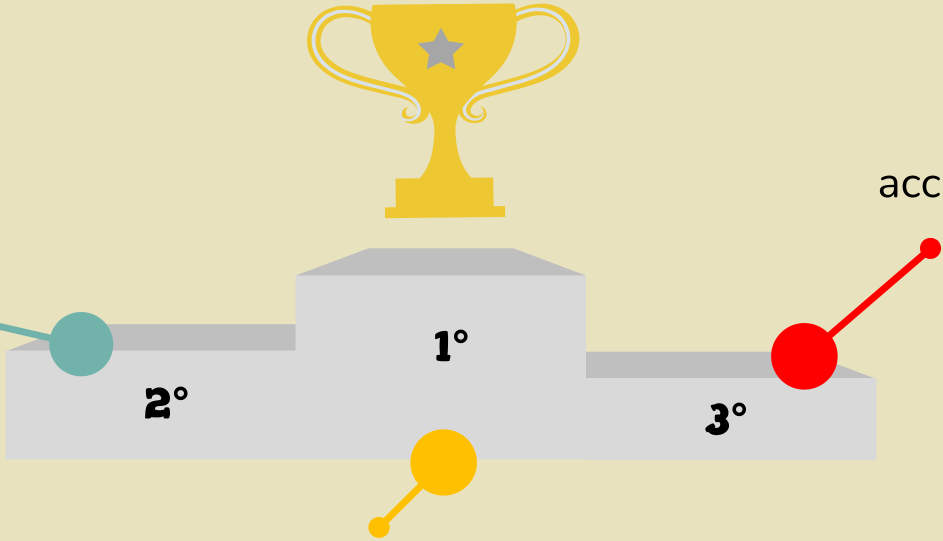
UserBased RS

- Collaborative Based RS
- Focused on user-song interaction

Relevance RS

«On top of the world»

Is it… sthocastic?



Items in descending order

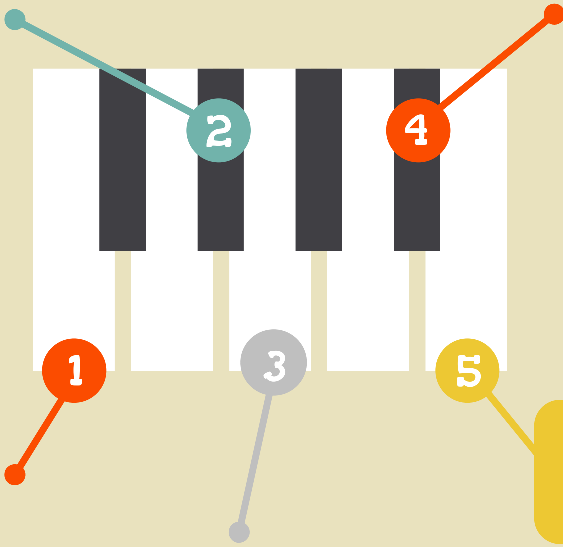
Items are selected according a set of columns

Top N items are

Recommended

CoolStartRS

«Are you a newcomer? Let me introduce you to this rave-world!»



Select

Already viewed items (target is part of this set)

N closest items to the target item

Select

‘artificial’ middle item

Compute

Select

closest items to the artificial item

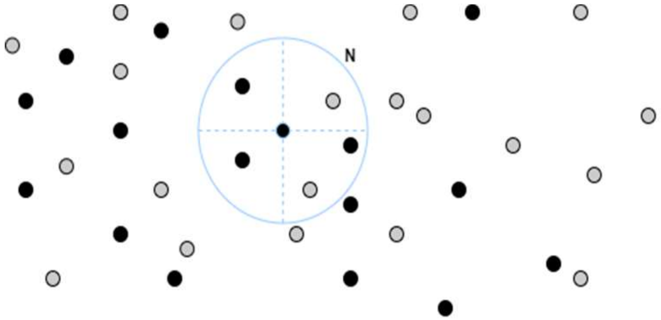
Never viewed N items

Recommended

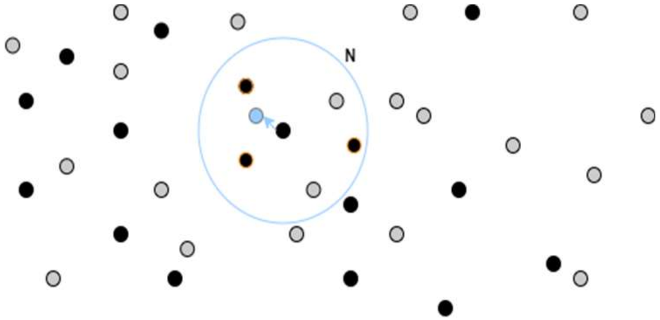
CoolStartRS

«Are you a newcomer? Let me introduce you to this rave-party!»

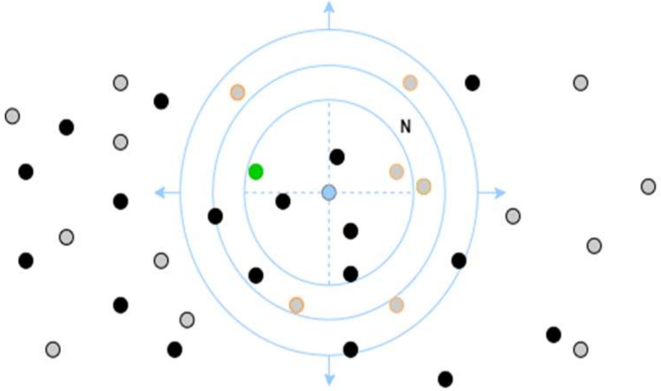
Is it... sthocastic?



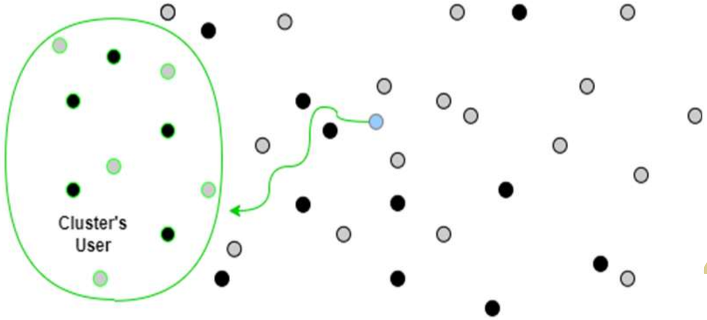
1



2



3



4*

WeightedMF RS

«Ehi you! Are you interested on underground music?»

HyperTuner

Tune of Hyperparameter through Bayesian Optimization

WeightedMF

Feedback Matrix Factorization and top N items are

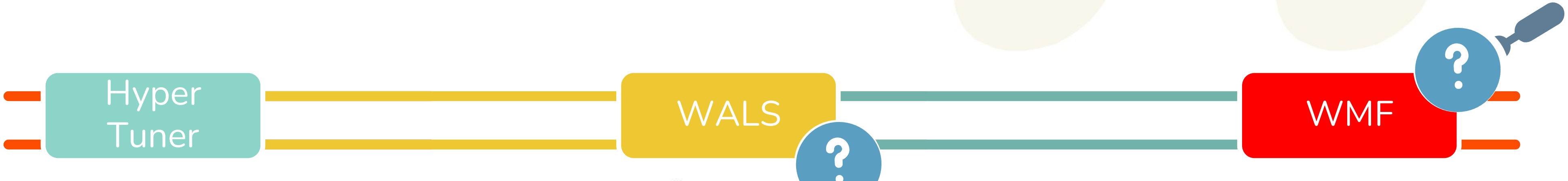
Recommended

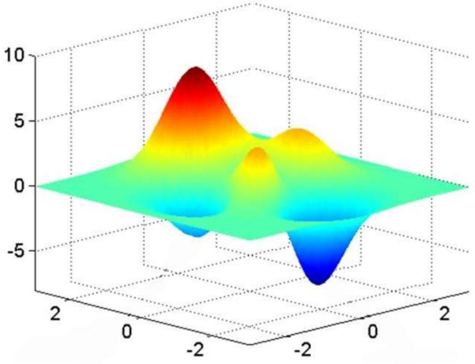
WeightedALS

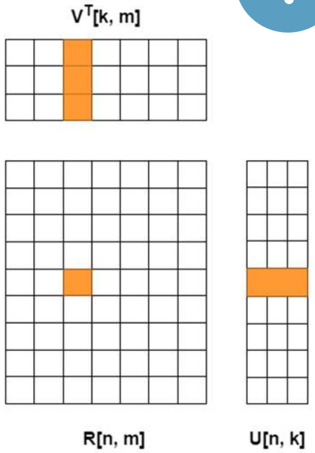
Compute embedded matrices and re-compute the Feedback Matrix

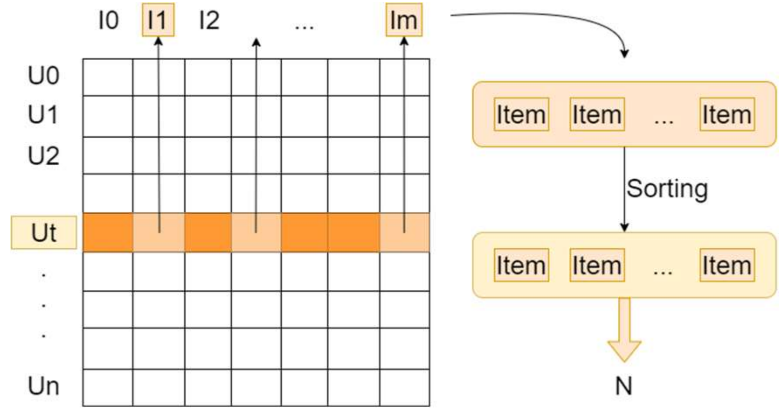
WeightedMF RS

«Ehi you! Are you interested on underground music?»









- Reg in [0,01; 1,0], K in [20; 60], iter in [100; 120]
- K-Fold Cross-Validation

- Matrix R decomposed in U, V
- U, V embeddings matrices
- Iteratively compute U and V

- Compute the approximated R matrix
- Select top N items

1. *The HyperSpace image is for illustration purposes only*
2. *Image credits to ResearchGate*

Is it...
sthocastic?

WeightedMF RS

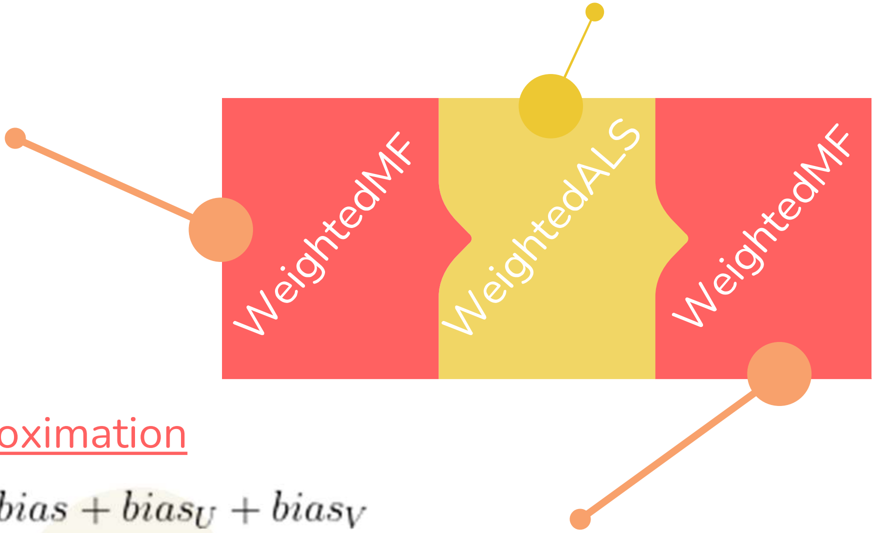
*«Ehi you! Are you interested on **underground** music?»*

Matrix Generation

$$M = W \circ (X - bias - bias_M - bias_{other}) \bullet M_{other} \bullet \left[\left(M_{other}^T \circ W_{transf} \right) \bullet M_{other} + \lambda \cdot \mathbb{I}_{n_{lf}} \right]^{-1}$$

Settings

- $X = original\ matrix$
- $bias = Avg(X)$
- $bias_U = Avg(X_{[1,\ u]})$
- $bias_V = Avg(X_{[v,\ 1]})$



- $X \approx \tilde{X}$
- Add biases
- Weights dependent
- WALS iterative

X Matrix Approximation

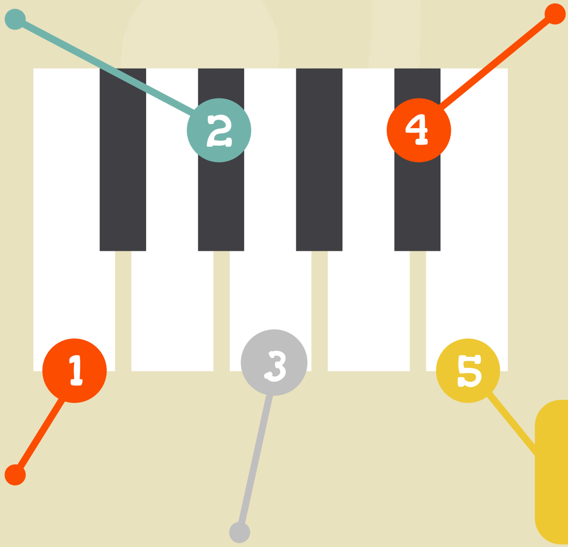
$$\tilde{X} = U \bullet V^T + bias + bias_U + bias_V$$

Loss Function

$$Loss\,(\lambda, U, V) = \left\| W \bullet \left(X - bias_X - bias_U - bias_V U \bullet V^T \right) \right\|^2 + \lambda \cdot \left(\|U\|^2 + \|V\|^2 + bias_U^2 + bias_V^2 \right)$$

UserBased RS

«Fan club means funny music»



Select

Most similar users to the current user

Users who have listened the item

Select

Select

closest items to the listened item

Never viewed N items

Recommended

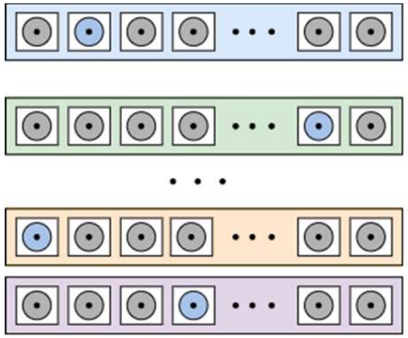
Items of the current user in the history of the other users

Ignore

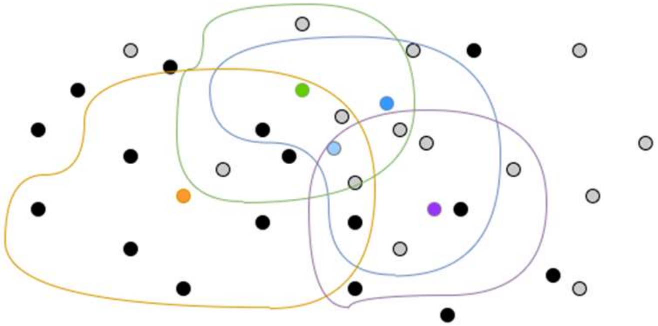
UserBased RS

«Fan club means funny music»

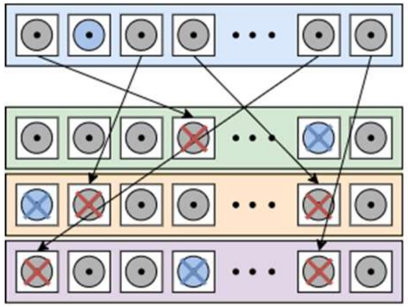
Is it... sthocastic?



1



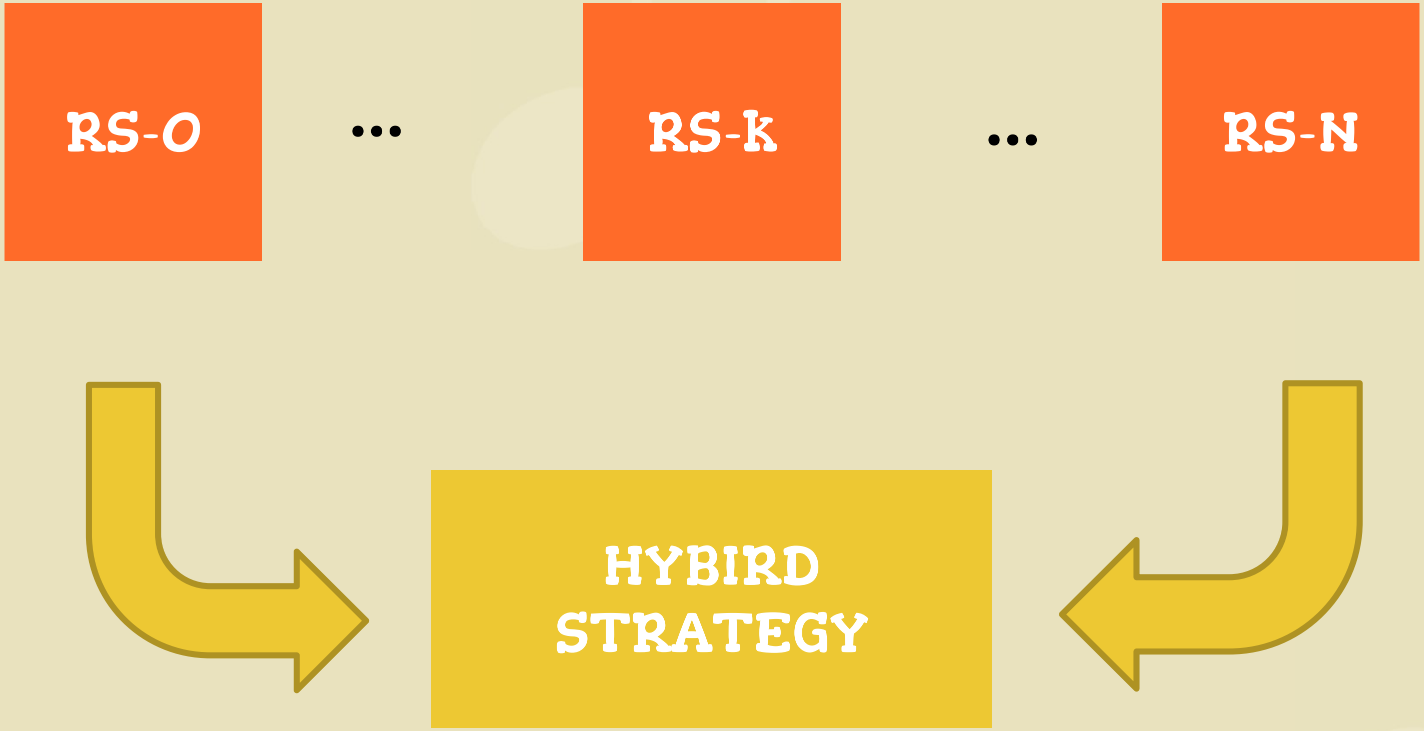
2*



3

4*

Which configuration?



RS-0 ... RS-k ... RS-N

HYBIRD STRATEGY

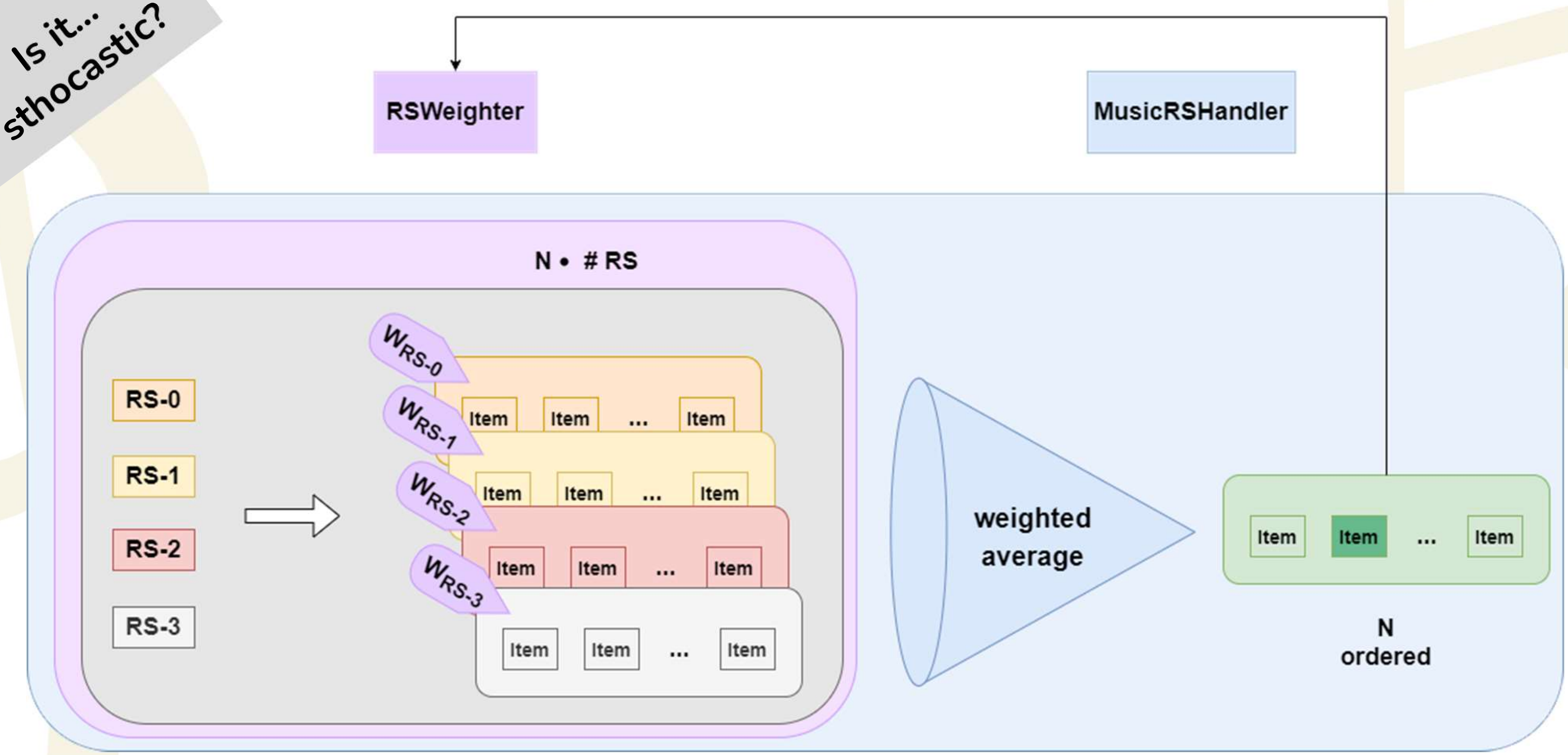
No Music Managers.. No Party

Each RS play its song, but without a music handler we may have just «noice»... a way to **combine** different sets of recommendations is necessary.

Image credits to shutterstock.com

No Music Managers.. No Party

Is it... sthocastic?



- Retro-active system to give more influence to certain(s) RS(s)
- How assign weights to RSs is a critical point!

04

Conclusion

“even the most beautiful
songs have an end”

Problem & Achivment

Problem

- Lack of “proper” datasets
- Big party in a small place
- Refactoring

Achivment

- Nice challenge!
- Quasi-real is a step towards the real
- No one, no two, no three but four RSs

Great Band but.. who is the best?

It is not really correct to ask this question:

- Investigation of different approches
- Different data means different playing fields
- It depends on the **customer**

Image credits to shutterstock.com

Show must go on!

- APIs connection to be always update
- Advanced cryptography for data & Input validation
- User interface
- General improvement with more item features (Geolocalization, ...)
- **Develop Analytics section & develeop fifth member of RS Band**
- Periodic backups of data
- Grant settings for user and staff
- Cloud migration
- Adavanced weight assignment algorithms for MF
- **Weighting RS recommendations on the basis of the accuracy** (Reinforcement Learning, Ensemble semi-unsupervised Learning, Probabilistics weight, ...)

Artists need inspiration too

A set of resources where it is possible to find useful information/references.

Datasets

1. Kaggle Dataset
2. Yahoo Music

How build a Recommender System

1. Build a Recommendation engine collaborative filtering
2. Kmeans Clustering to categorize music by similar audio
3. Comprehensive Guide on item based collaborative Filtering
4. Information Retrieval, lecture 12, Luca Manzoni
5. In-depth guide to how Recommender System work

Weighted Alternate Least Square Algorithm

1. Alternating Least Squares with Weighted Regularization
2. Weighted-SVD: Matrix Factorization with Weights on the Latent Factors
3. Alternating Least Square for Implicit Dataset with code
4. Finding Similar Music using Matrix Factorization
5. Explicit Matrix Factorization: ALS, SGD, and All That Jazz
6. Matrix Factorization for Personalized Recommendation With Implicit Feedback

Thanks!

A quasi-real Music Recommender System

Pietro Morichetti