

Home

TV Shows

Movies

News & Popular

My list

Watchlist



# SEARCH SYSTEM MILESTONE 3



PAULO RIBEIRO, UP201806505  
PEDRO FERREIRA, UP201806506  
PEDRO PONTE, UP201809694



# | TABLE OF CONTENTS

01

## Collection

- Definition
- Document

02

## Indexing

- Preparing Solr

03

## Revisions Introduced

- Work improvements
- New features

04

## Final System Evaluation

- Comparison of the three systems

05

## User Interface

- Functionalities

06

## Conclusion & Future Work

- Final conclusion



Home

TV Shows

Movies

News & Popular

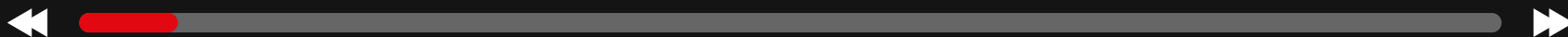
My list

Watchlist



01

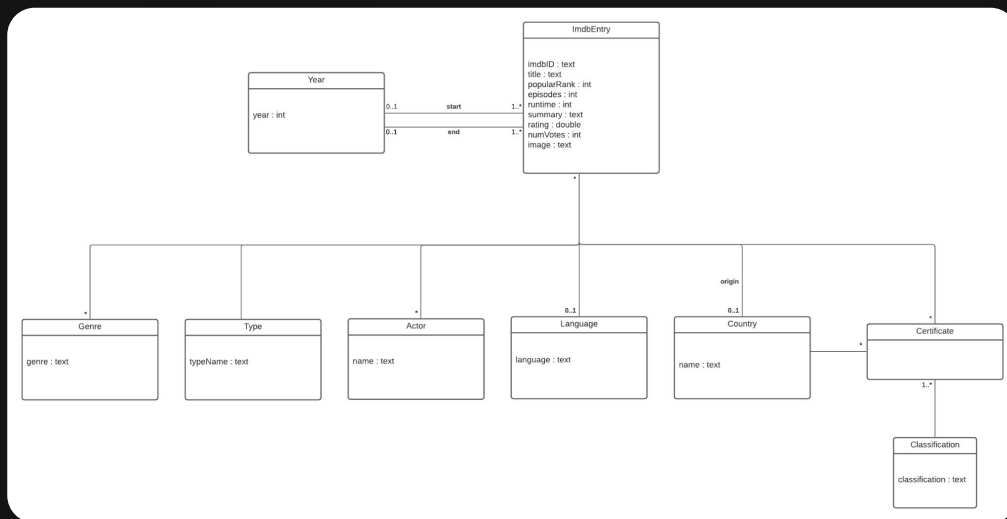
**Collection**





# | Collection

- **Collection:** All shows in the dataset obtained from Milestone 1
- **Document:** A show
- The dataset is a single CSV file which, for convenience, we transform into a JSON file



Home

TV Shows

Movies

News & Popular

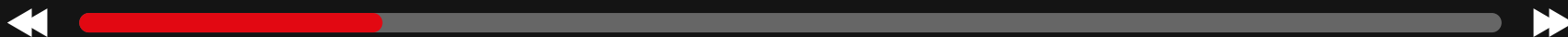
My list

Watchlist



02

**Indexing**





# | Indexing

- Run Docker image based on Solr's container
- Dockerfile to define a custom image and perform all the basic setup operations
  - Added query 5, synonyms file and new config file modified to include spellcheck
- Script to automatically create a core (shows), define our schema and populate the database
  - Copied the new config file to Solr

```
FROM solr:8.10

COPY netflix_list.json /data/netflix_list.json

COPY query1.json /data/query1.json

COPY query2.json /data/query2.json

COPY query3.json /data/query3.json

COPY query4.json /data/query4.json

COPY query5.json /data/query5.json

COPY schema.json /data/schema.json

COPY synonyms.txt /data/mysynonyms.txt

COPY solrconfig.xml /data/config.xml

COPY startup.sh /scripts/startup.sh

ENTRYPOINT ["/scripts/startup.sh"]
```

```
#!/bin/bash

precreate-core shows

# Start Solr in background mode so we can use the API to upload the schema
solr start

sleep 2

# Define Synonyms
cat /data/mysynonyms.txt >> /var/solr/data/shows/conf/synonyms.txt

cat /data/config.xml > /var/solr/data/shows/conf/solrconfig.xml

# Schema definition via API
curl -X POST -H 'Content-type:application/json' \
  --data-binary @/data/schema.json \
  http://localhost:8983/solr/shows/schema

sleep 2

# Populate collection
bin/post -c shows /data/netflix_list.json
# bin/post -c shows /data/query1.json
# bin/post -c shows /data/query2.json
# bin/post -c shows /data/query3.json
# bin/post -c shows /data/query4.json
# bin/post -c shows /data/query5.json

# Restart in foreground mode so we can access the interface
solr restart -f
```



Home

TV Shows

Movies

News & Popular

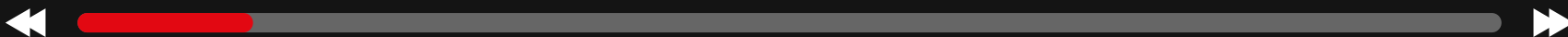
My list

Watchlist



03

# Revisions Introduced





# | Revisions Introduced

## Changes introduced to the system:

- **Metrics:** Modify, improve and add new Metrics. Update Precision, Recall, Average Precision and add MAP
- **Improvements:** Spellcheck and Synonyms
- **Interface:** Creation of an interface for the search system

	Schema-less	Enhanced Schema	Final Schema
MAP	1.0	0.88	1.0





# | Synonyms and Spellcheck

- Search for PRC, the initials for Popular Republic of China

```
{
  "responseHeader":{
    "status":0,
    "QTime":7,
    "params":{
      "q":"title:prc",
      "indent":"true",
      "q.op":"OR",
      "_":"1642450558223"}},
  "response":{"numFound":3,"start":0,"numFoundExact":true,"docs":[
    {
      "imdbID":"tt8983180",
      "title":["Made in China"],
```

- Spellcheck suggestions when we misspell the word "friends" as "fiends"

```
"spellcheck":{
  "suggestions":[
    "fiends",{
      "numFound":5,
      "startOffset":6,
      "endOffset":12,
      "origFreq":0,
      "suggestion":[{
        "word":"friend",
        "freq":49},
        {
          "word":"field",
          "freq":7},
        {
          "word":"find",
          "freq":11}}]}
```



Home

TV Shows

Movies

News & Popular

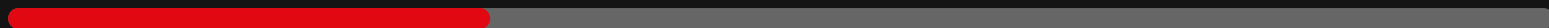
My list

Watchlist



04

# Final System Evaluation





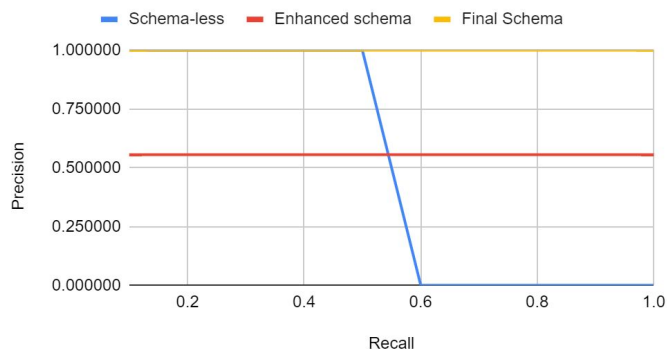
# | Final System Evaluation

## Query 4 (revisited)



- 30 shows
- 5 relevant
- User expects only series and miniseries

Precision-Recall Curve



### Schema-less:

Metric	Value
1 Average Precision	1.0
2 Precision at 9 (P@9)	0.222222
3 Recall at 9 (R@9)	0.4
4 F1 at 9 (R@9)	0.285714

### Enhanced schema:

Metric	Value
1 Average Precision	0.403492
2 Precision at 9 (P@9)	0.555556
3 Recall at 9 (R@9)	1.0
4 F1 at 9 (R@9)	0.714286

### Final system:

Metric	Value
1 Average Precision	1.0
2 Precision at 9 (P@9)	0.555556
3 Recall at 9 (R@9)	1.0
4 F1 at 9 (R@9)	0.714286





# | Final System Evaluation

## Query 5

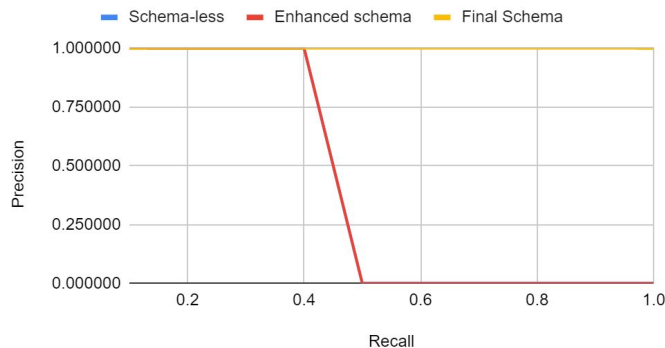


- 30 shows
- 10 relevant
- Shows related to the United Kingdom

### Example of Results:

- Criminal: UK
- Battle of Britain

Precision-Recall Curve



#### Schema-less:

	Metric	Value
1	Average Precision	1.0
2	Precision at 10 (P@10)	0.3
3	Recall at 10 (R@10)	0.3
4	F1 at 10 (R@10)	0.3

#### Enhanced schema:

	Metric	Value
1	Average Precision	1.0
2	Precision at 10 (P@10)	0.3
3	Recall at 10 (R@10)	0.3
4	F1 at 10 (R@10)	0.3

#### Final system:

	Metric	Value
1	Average Precision	1.0
2	Precision at 10 (P@10)	1.0
3	Recall at 10 (R@10)	1.0
4	F1 at 10 (R@10)	1.0



Home

TV Shows

Movies

News & Popular

My list

Watchlist



05

# User Interface



[Home](#)[TV Shows](#)[Movies](#)[News & Popular](#)[My list](#)[Watchlist](#)

# | Search Page

## NETFLIX

[SORT BY: RELEVANT ▾](#)

Search Fields:

- ☒ title
- ☒ originCountry
- ☒ language
- ☒ cast
- ☒ genres
- ☒ summary
- ☒ imdbID

Show Type:

- ☒ animation
- ☒ miniSeries
- ☒ movie
- ☒ series
- ☒ short
- ☒ special

Start Year:

 TO 

Number of Episodes:

 TO 

Rating:

 TO 

End Year:

 TO 

Runtime:

 TO 

Number of Votes:

 TO [SUBMIT](#)

No results found





[Home](#)[TV Shows](#)[Movies](#)[News & Popular](#)[My list](#)[Watchlist](#)

# | Spell Check

## NETFLIX

Search

frend

SORT BY: RELEVANT ▾

Search Fields:

☒ title  
☒ originCountry  
☒ language  
☒ cast  
☒ genres  
☒ summary  
☒ imdbID

Show Type:

☒ animation  
☒ miniSeries  
☒ movie  
☒ series  
☒ short  
☒ special

Start Year:

TO

Number of Episodes:

TO

Rating:

TO

End Year:

TO

Runtime:

TO

Number of Votes:

TO

SUBMIT

68 results found

Did you mean: [friend](#) [frenc](#) [frenz](#)



[Home](#)[TV Shows](#)[Movies](#)[News & Popular](#)[My list](#)[Watchlist](#)

# | Show Info

Search Fields

☒ title  
☒ originCountry  
☒ language  
☒ cast  
☒ genres  
☒ summary  
☒ imdbID

Show Type:

☒ animation  
☒ miniSeries  
☒ movie  
☒ series  
☒ short  
☒ special

Start Year:

TO

Number of Episodes:

TO

Rating:

TO

Number of Votes:

TO

walk

SORT BY: RELEVANT

Self Made: Inspired by the Life of Madam C.J. Walker

INFO

CAST


CERTIFICATES

Biography

Drama

History

1465



Rating: 7.3

Description

The inspiring story of trailblazing African American entrepreneur Madam C.J. Walker who built a haircare empire that made her America's first female self-made millionaire.

Type:

miniSeries

Language:

English

Origin Country:

United States

Runtime:

191

Year of Release:


2020

Episodes:

4

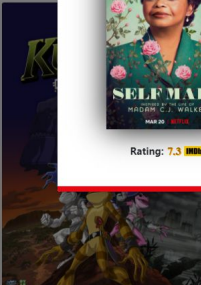
Number of Votes:

9922




Self Made: Inspired by the Life of Madam C.J. Walker

MORE INFO




Kulipari: Dream Walker

MORE INFO




Walk. Ride. Rodeo.

MORE INFO



He Walked by Night

MORE INFO



Takers

MORE INFO

◀

▶

[Home](#)[TV Shows](#)[Movies](#)[News & Popular](#)[My list](#)[Watchlist](#)

# | Show Cast

Search Fields:

- ☒ title
- ☒ originCountry
- ☒ language
- ☒ cast
- ☒ genres
- ☒ summary
- ☒ imdbID

Show Type:

- ☒ animation
- ☒ miniSeries
- ☒ movie
- ☒ series
- ☒ short
- ☒ special

Start Year:

TO

Number of Episodes:

TO

Rating:

TO

Number of Votes:

TO

walk

SORT BY: RELEVANT

Self Made: Inspired by the Life of Madam C.J. Walker

INFO

CAST


CERTIFICATES

Biography

Drama

History

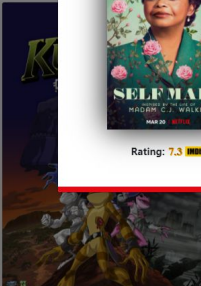
1465



Rating: 7.3


Cast

Octavia Spencer, Tiffany Haddish, Carmen Ejogo, Kevin Carroll, Blair Underwood, Garrett Morris, J. Alphonse Nicholson, Sydney Morton, Karen Glave, Mouna Traor, Cornelius Smith Jr, Bill Bellamy, Rothafor Gray, Sophia Deerr, Zahra Bentham, Sheila Isaro, Kiki Hammill, Keeya King



Kulipari: Dream Walker


MORE INFO



Walk. Ride. Rodeo.


MARCH 8  
NETFLIX

MORE INFO




He Walked by Night

MORE INFO



Takers

MORE INFO



Self Made: Inspired by the Life of Madam C.J. Walker

MORE INFO



[Home](#)[TV Shows](#)[Movies](#)[News & Popular](#)[My list](#)[Watchlist](#)

# | Show Certificates

Search Fields:

- ☒ title
- ☒ originCountry
- ☒ language
- ☒ cast
- ☒ genres
- ☒ summary
- ☒ imdbID

Show Type:

- ☒ animation
- ☒ miniSeries
- ☒ movie
- ☒ series
- ☒ short
- ☒ special

Start Year:

TO

Number of Episodes:

TO

Rating:

TO

Number of Votes:

TO


Sort By: RELEVANT

Self Made: Inspired by the Life of Madam C.J. Walker

INFOCASTCERTIFICATES

BiographyDramaHistory

1465



Rating: 7.3

### Certificates

Argentina: 13, Australia: M, Brazil: 16, Germany: 12, Netherlands: 12, Singapore: M18, South Korea: 15, Spain: 13, United Kingdom: 15, United States: TV-MA

Self Made: Inspired by the Life of Madam C.J. Walker

MORE INFO

Kulipari: Dream Walker

MORE INFO

Walk. Ride. Rodeo.

MORE INFO

He Walked by Night

MORE INFO

Takers

MORE INFO



Home

TV Shows

Movies

News & Popular

My list

Watchlist



06

# Conclusion & Future Work





# THANKS!

## Do you have any questions?

CREDITS: This presentation template was created by Slidesgo,  
including icons by Flaticon and infographics & images by Freepik

