

Manual

1 Initial Setup

First, clone the repository from [Github](#).

Then, extract the data generation files into the db-generation directory and run `python3 genTable_mongoDB10G.py` (The videos should be in mp4 format to be playable in the webbrowser). The resulting file tree should look similar like this:

```
./
├── backend
│   ├── static
│   ├── templates
│   └── backend.py
├── configs
│   ├── nginx.conf
│   └── storage_nginx.conf
└── db-generation
    ├── articles
    ├── bbc_news_texts
    ├── article.dat
    └── genTable_mongoDB10G.py
...
├── .gitignore
├── calcTables.py
├── docker-compose.yml
├── import_data_to_mongo.py
├── import_map_to_mongo.py
├── init.sh
├── README.md
├── slice_table.py
├── task_splitting.md
├── update_file_to_dfs.py
├── update_file_to_dfs.sh
└── utils.py
```

This project requires Docker and Docker-compose. To pull docker image, you may need to replace the resources with mirror resources.

To initialize the database system run

```
bash init.sh
```

This creates and starts the databases and the FastDFS server as docker-compose services, fragments the data and imports it to the respective databases, and uploads the images, texts and videos to the FastDFS storage server.

The output is like

```
...
```

```

2024-12-27T14:11:26.960+0000    connected to: mongodb://localhost/
2024-12-27T14:11:26.984+0000    22 document(s) imported successfully. 0
document(s) failed to import.
Line 27: Fri Dec 27 22:11:27 CST 2024 - Command took 41 seconds
{ "acknowledged" : true, "deletedCount" : 0 }
2024-12-27T14:11:26.960+0000    connected to: mongodb://localhost/
2024-12-27T14:11:26.984+0000    22 document(s) imported successfully. 0
document(s) failed to import.
Line 27: Fri Dec 27 22:11:27 CST 2024 - Command took 41 seconds
led to import.
Line 27: Fri Dec 27 22:11:27 CST 2024 - Command took 41 seconds
Successfully copied 3.58kB to storage1:/etc/fdfs_buffer/
Uploading files..
File upload complete. Mapping stored in /etc/fdfs_buffer/mapping_records.jsonl
Line 38: Fri Dec 27 22:33:28 CST 2024 - Command took 1321 seconds
Line 40: Fri Dec 27 22:33:28 CST 2024 - Command took 0 seconds
Line 42: Fri Dec 27 22:33:28 CST 2024 - Command took 0 seconds
...

```

If the file is uploaded to fdfs correctly, we can access to it by visiting http://localhost:8888/group1/M00/00/00/rBIABmcg4nyAbwISAAJYI_w6cw260.txt

To run the Flask App navigate to ./backend and run

```
python3 backend.py
```

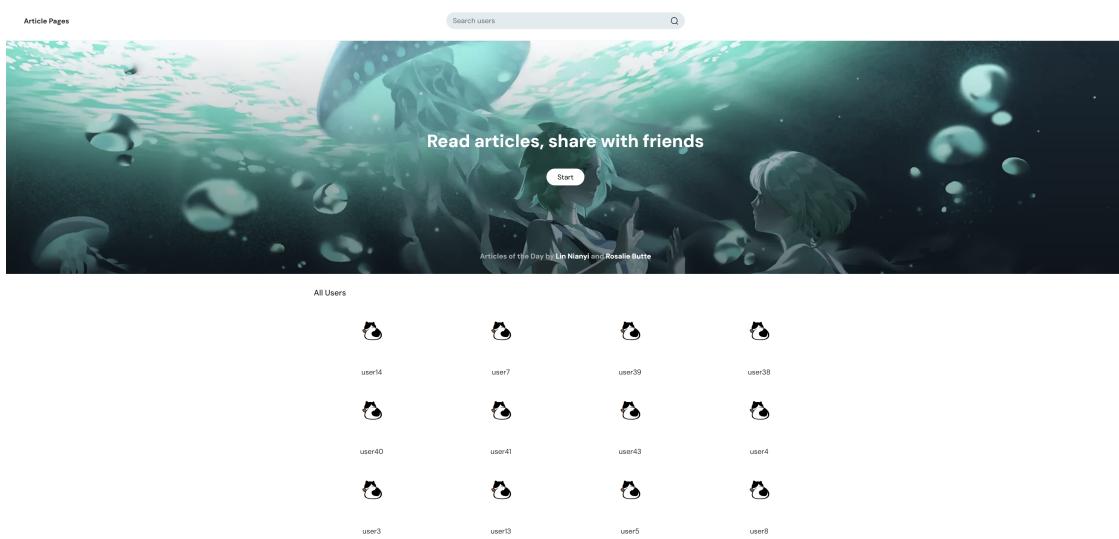
With this the system is ready to receive clients.

2 Usage

We created several APIs to query and display the data:

- http://localhost:8080/frontend/user_list/1

Lists all users, able to click to view a user in detail, able to search for a specific user, can switch to view the article list



- <http://localhost:8080/frontend/user/1>

Possible to replace id 1 with another user id. Shows a users detailed personal information as well as their reading history and for each article if they commented/liked/shared it.

user1
Email: user1@example.com
Dept: design
Grade: grade4
Tags: tag1, tag2, tag3, tag4, tag5, tag6, tag7, tag8, tag9, tag10, tag11, tag12, tag13, tag14, tag15, tag16, tag17, tag18, tag19, tag20, tag21, tag22, tag23, tag24, tag25, tag26, tag27, tag28, tag29, tag30, tag31, tag32, tag33, tag34, tag35, tag36, tag37, tag38, tag39, tag40
Obtained Credits: 77

Reading Records

Title	Date
title3729	2017-09-25 20:41
title3099	2017-09-26 04:37
title7855	2017-09-27 10:38
title6540	2017-09-27 19:43
title3220	2017-09-30 21:29
title6717	2017-10-04 02:39
title5022	
title6484	
title3801	

- http://localhost:8080/frontend/article_list/1

Lists all articles, able to view an article in detail, able to search for an article, able to search for popular ranked articles, can switch to view the user list

User Pages

Search articles

Read articles, share with friends

Articles of the Day by Lin Nianyi and Rosalie Butte

Start

Search Popular Articles

Select granularity: Daily

Select date: 年 / 月 / 日

Submit

All Articles

Title	Date
title1	2017-09-21 21:20
title9	2017-09-21 21:20
title0	2017-09-21 21:20
title1	2017-09-21 21:20

- <http://localhost:8080/frontend/article/1>

Possible to replace id 1 with another article id. Shows the article text with all its images, videos and comments if it has any. Also shows statistical information, such as number of views/likes/shares.

Views: 43 Comments: 6 Likes: 8 Shares: 0

Published on 2017-09-21 21:20

Abstract:
abstract of article 1

UK 'world's biggest music buyer' UK consumers are the biggest music buyers in the world, according to new figures released on Tuesday by the British Phonographic Industry (BPI). But global sales of recorded music were flat in 2004, with growing sales in DVD music videos offset by online and physical piracy. "The UK market is the only one that has continued to grow over the last few years, which is mainly due to its robust albums market. However, world music sales declined by 13% to \$33.6 billion (Â£17.7 billion). The UK CD album market grew by 4.5% in 2004 with a record 714.6 million units sold. On average, there were 10.5 million unique UK album buyers per month in 2004, up from 9.5 million in the UK each year, making Britain second only to the US in terms of the number of releases. Overall sales were driven by new UK-signed artists such as Keane, Katie Melua and Scissor Sisters, whose debut album sold almost 1.5 million copies in the UK. The BPI said that the UK had the highest level of demand for great music," said BPI chairman Peter Jamieson. The best-selling album globally was Usher's Confessions - one of eight albums to sell more than five million copies in 2004. The growth in the DVD market has been fuelled by the success of the Harry Potter films. The BPI said: "The total number of tracks downloaded last year were up more than tenfold on 2003. Digital sales in the US in the first half of 2004 were up 15% compared with the same period in 2003. However, some markets in Central Europe and Asia - notably Sweden, France, Italy, Netherlands and Spain - have been drastically hit by internet piracy. Australia, Italy, and the Netherlands also saw a decline. Online sales of digital music in the US increased by 10% in 2004, according to the International Federation of Phonographic Industries (IFPI). However, commercial piracy and illegal file-sharing are continuing to depress our markets. 'The priority in the coming year is to step up the advance of legitimate digital sales and sustain our anti-piracy efforts.' The main choice today's music fan now has to make is whether to get music legally or illegally," added Mr Jamieson.

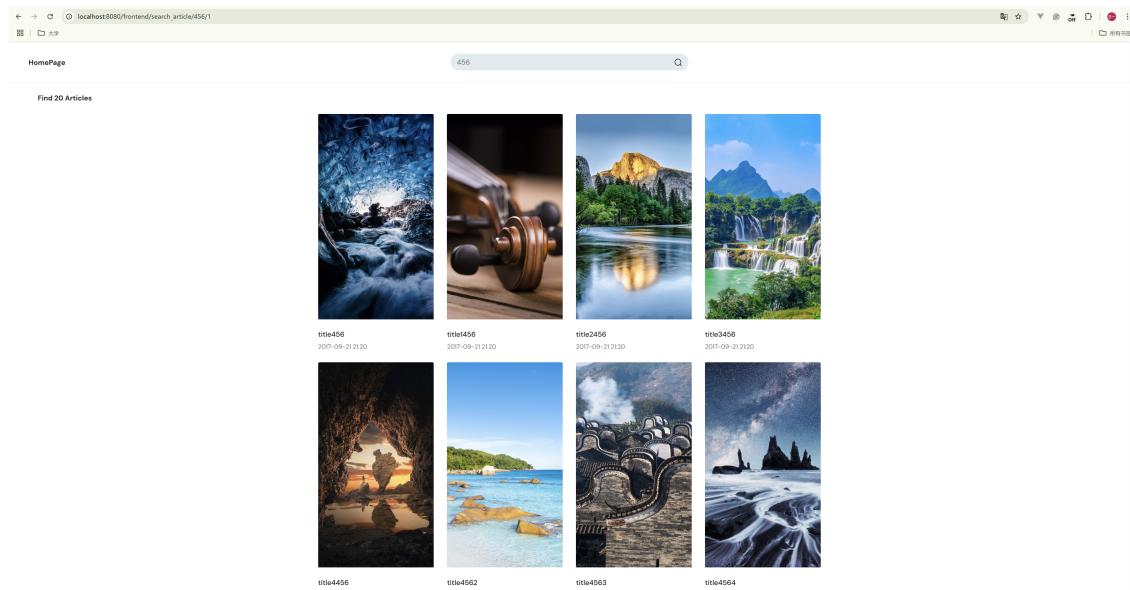
- http://localhost:8080/frontend/search_user/{search_text}/{pageid}

Filters users that match with the search_text. Search_text can be the user id, name, email or phone number.



- http://localhost:8080/frontend/search_article/{search_text}/{pageid}

Filters articles that match with the search_text. Search_text can be the title or the article id.



- http://localhost:8080/frontend/popular_rank/{grainaty}/{date}

Finds the top 5 popular articles based on granularity and date. Where granularity is either daily, weekly or monthly.

localhost:8080/frontend/popular_rank/weekly/2017-12-28

大学

HomePage

Search articles

Search Popular Articles
Select granularity: Daily Select date: 年 / 月 / 日 Submit

Top 5 Articles from Week 52, 2017


title9211
Number of Views: 14


title3722
Number of Views: 13


title3700
Number of Views: 12