

# **SGDB Project Report**

## **Books**

## **Recommendation**

## **And Sharing**

**SHA-FI-RE-DO**

---

### **GROUP MEMBERS :**

Bhumika Suresh Kothwal : 191081014

Neha Bhausaheb Shinde : 191081072

Nishtha Rajiv Sainger : 191081064

15th December, 2021



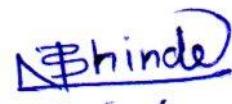
**Project Under the guidance of Prof. Sowmiya Raksha**

## **Statement By Candidate**

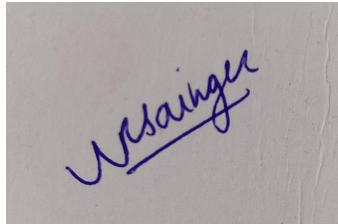
We wish to state that work embodied in this special topic titled "Spatial Graph Database" forms my own contribution to the work carried out under the guidance of Prof.Somiya Raksha at the 'Veermata Jijabai Technological Institute', Matunga, Mumbai – 19. This work has not been submitted for any other Degree or Diploma of any University / Institute. Wherever references have been made to previous works of others, it has been clearly indicated.



Ms. Bhumika Kothwal



Ms. Neha Shinde



Ms. Nishtha Sainger

## Table Of Contents

- Problem Statement
- Objectives
- Introduction
- Background Research
  - Recommendation Models
- Database
  - Database Model : neo4j
- Graph Database
  - Graph Database Tables
  - Graph DatabaseModel

- 
- Spatial Database
    - Spatial Database Tables and ER diagram
  - Features Applied
  - Implementation
  - Tech Stack
  - References
  - Future Scope
  - Conclusion

## **Problem Statement**

Creating a web portal for recommending books based on content filtering and collaborative filtering.

Also providing user option to buy book online (using web scraping) and offline with location of the seller (using spatial analysis)

## **Objectives**

To incorporate spatial and graph databases in our web application and enhance features to improve user experience.

## **Introduction**

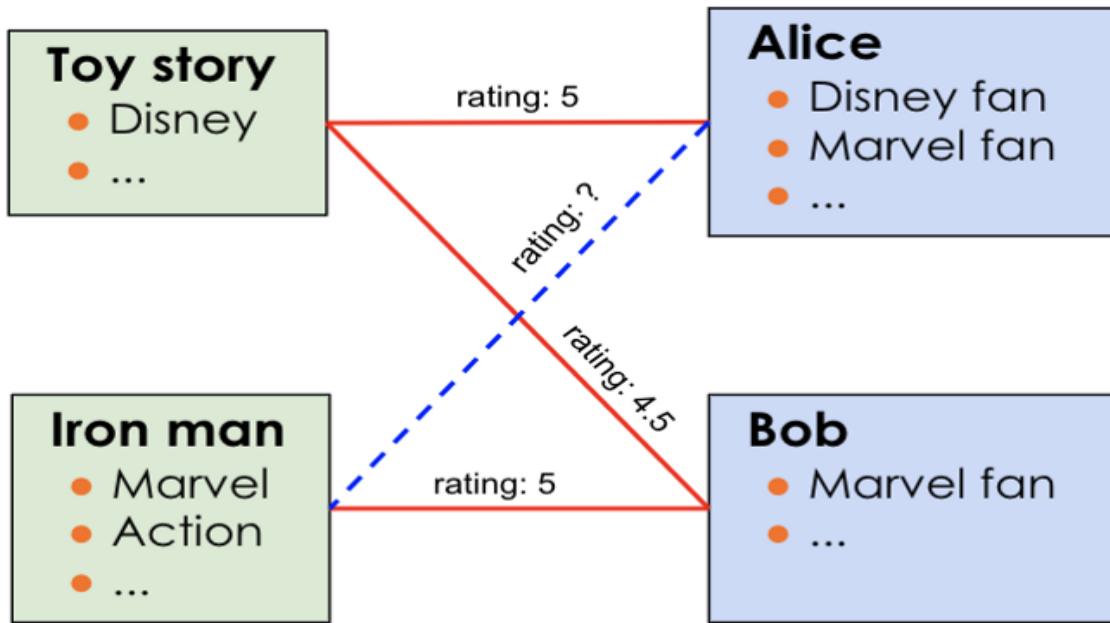
What is a recommendation system?

A recommendation system is any rating system which predicts an individual's preferred choices, based on available data.

Recommendation systems are utilized in a variety of services, such as video streaming, online shopping, and social media.

Typically, the system provides the recommendation to the users

based on its prediction of the rating a user would give to an item. Recommendation systems can be categorized by two aspects, the utilized information and the prediction models.



## Which book to read next?

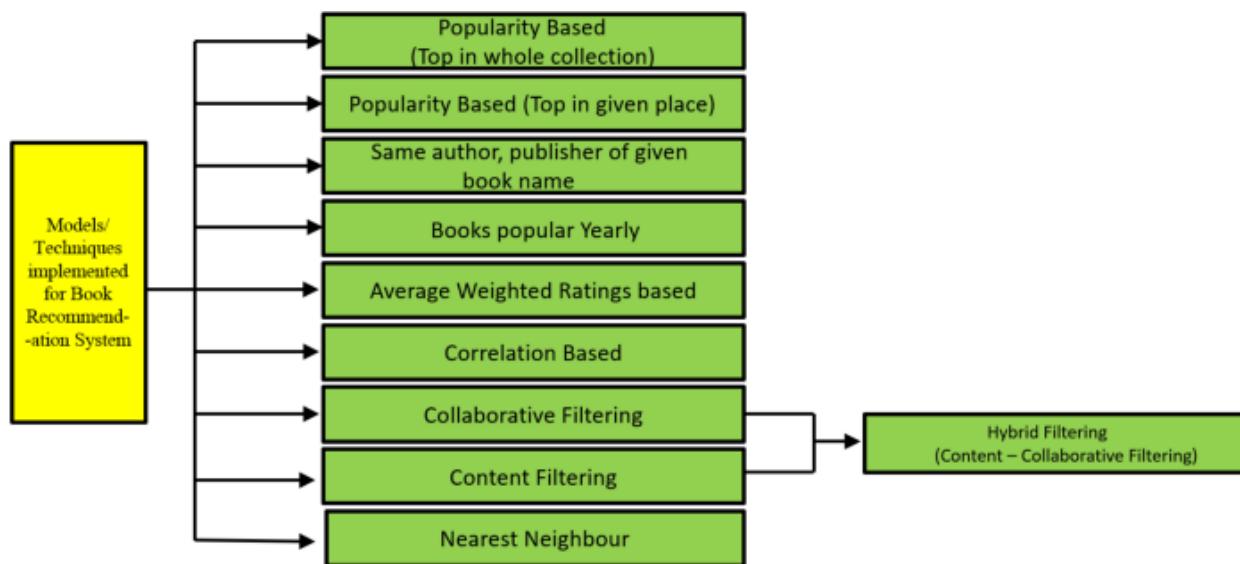
When we want to read a new book we generally ask our friends or classmates or may search all the books available in a library (as if we can). After all asking and searching, we may still not find any book of our preference as not everyone has the same interests. For such situations, we need a system which takes our choices into consideration and suggests to us some good books.

“ A good **recommender system** has to consider how users interact with the **recommendations.**”

A recommendation system broadly recommends items to the user best suited to their tastes and traits. It uses the user's previous data and other user's data to give new recommendations

## Recommendation Models

We started building some basic recommendation systems and then implemented collaborative and content-based filtering methods as well.



---

## **Books by the same author, publisher of the given book name**

For this model, we have sorted the books by rating for the same author and same publisher of the given book and recommended top n books.

# **TYPES OF RECOMMENDATION SYSTEMS**

## **Correlation Based**

For this model, we have created the correlation matrix for which we needed to reduce the dataset (because of limited resources). So we have considered only those books which have total ratings of more than 50. Then from this data, we have created a user-book rating matrix. For the input book using the correlation matrix, top books are recommended.

## **Nearest Neighbours Based**

To train the Nearest Neighbours model, we have created a compressed sparse row matrix taking ratings of each Book by each User individually. This matrix is used to train the Nearest Neighbours model and then to find n nearest neighbors using the cosine similarity metric

### **Collaborative Filtering (User-Item Filtering)**

Collaborative Filtering Recommendation System works by considering user ratings and finds cosine similarities in ratings by several users to recommend books. To implement this, we took only those books' data that have at least 50 ratings in all (because of limited resources).

### **Content-Based Filtering**

We have implemented a content-based recommendation system that recommends books by calculating similarities in Book Titles. For this, TF-IDF feature vectors are created for unigrams and bigrams of Book-Titles where only those books' data has been considered which are having at least 80 ratings (because of limited resources)

### **Hybrid Recommendation System**

We have built a hybrid recommendation system using both content-based filtering and collaborative filtering systems. A percentile score is given to the results obtained from both content and collaborative filtering models and is combined to recommend top n books.

## **Performing recommendations : the overall strategy**

In what follows, we will implement three complementary collaborative-filtering strategies in the so-called “user-based” paradigm. This means we’ll focus on the past behaviour of users (“what did they read ?”, “what ratings did they attribute?”) to highlight subsets of users who are similar to one another. To perform a recommendation for one given user  $u_1$ , the global behaviour of all its similar users will be studied. Here are the different steps :

- Define a way to evaluate the similarity between two users
- Retain users who are the most similar to the user you want to recommend books to,  $u_1$
- Define a way to rank books among the ones seen by the similar users

- 
- Recommend to  $u_1$  the books with best ranking, that he has not already seen.

For each of the three strategies, we asked our guinea pig to attribute ratings afterwards to the top 20 suggestions in order to qualitatively evaluate their performances.

### **Strategy 1 : « Books that have been read by most of the similar users »**

The first strategy consists in considering the similarity between two users  $u_1$  and  $u_i$  as the ratio of the number of books they have in common, over the number of books  $u_1$  has read. As regards books seen by similar users, they are ranked from most popular to least popular. This simple method can be implemented even if ratings are not known.

- Given  $R_{ui}$ , the set of books rated by  $u_i$ , the similarity between two users  $u_1$  and  $u_i$  can be expressed as :

$$\text{sim}(u_1, u_i) = \frac{\text{card} [R_{u_1} \cap R_{u_i}]}{\text{card} [R_{u_1}]}$$

- Let  $U_{\text{sim}}$  be the set of users who are most similar to  $u_1$ . We define a threshold  $t$  (in our case  $t$  is arbitrarily set to 0.5) such that :

$$U_{\text{sim}} = \bigcup_{\text{sim}(u_1, u_i) > t} u_i$$

••

The threshold  $t$  can be optimized to tune the selectivity of a subset of similar users, and at the same time the average number of users in  $U_{\text{sim}}$ .

- The set of books rated by the similar users, but not by  $u_1$ , can be expressed as :

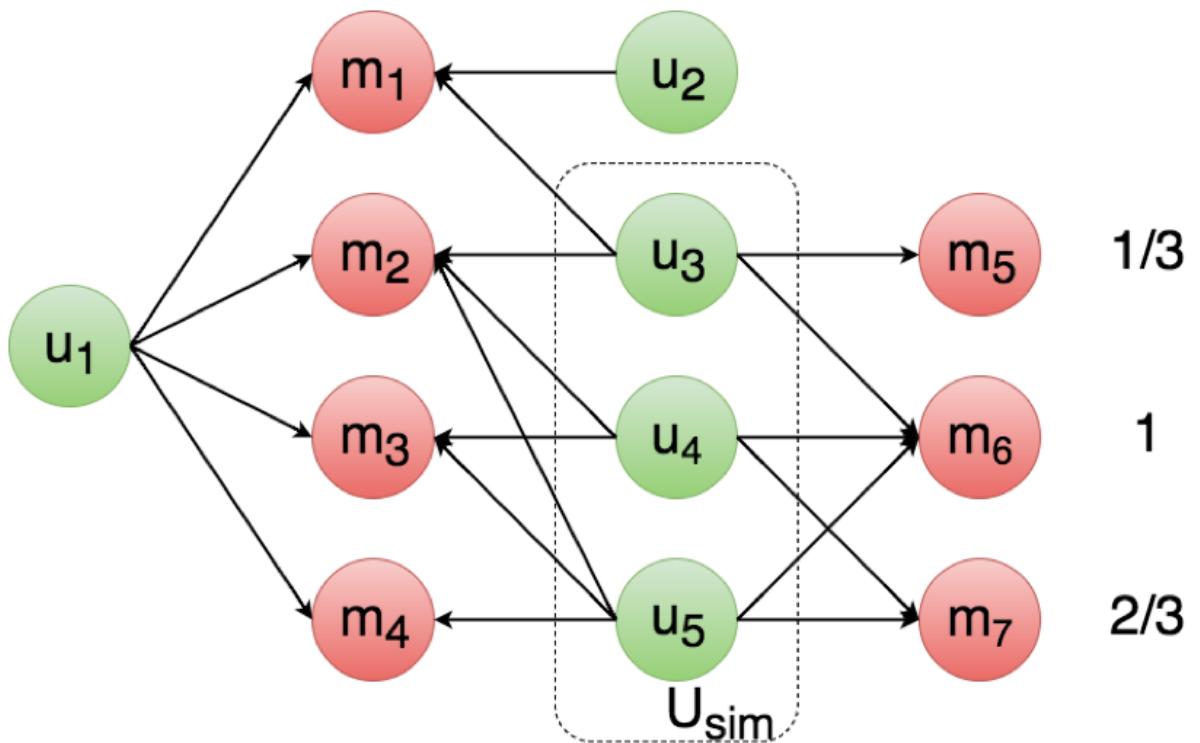
$$M_{\text{sim}} = [ \bigcup_{u_i \in U_{\text{sim}}} R_{u_i} ] \cap \overline{R_{u_1}}$$

- The ranking of each book  $m_j$  of  $M_{\text{sim}}$  is then computed as the proportion of similar users who have rated  $m_j$ . In this case, books will be ranked from « read by most similar

users » to « read by few similar users »:

$$s(m_j) = \frac{\text{card} \left( \bigcup_{\substack{u_i \in U_{\text{sim}} \\ m_j \in R(u_i)}} u_i \right)}{\text{card}(U_{\text{sim}})}$$

- The 20 books  $m_j$  with largest ranking  $s(m_j)$  will be recommended to user  $u_1$ .



## **DATABASES :**

**Our book.csv dataset is downloaded from the [link](#).**

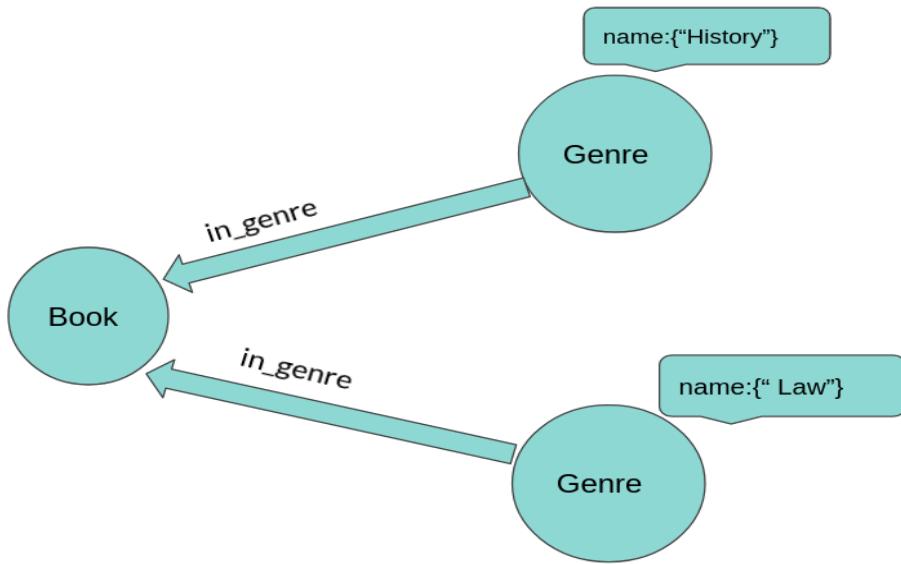
### **Graph Database Model : neo4j**

One of the earliest decisions you may encounter is whether to model something as a property on a node or as a relationship to a separate node. Consider the following model, the data below modeling a book genre as a property on the Book node.

To write a query finding the genre(s) of a particular book is very simple. It would find the Book node it wants to know about, then return the values listed in the genre property.

However, to find out which books share genres, you would need a much more complex query to find each Book node, loop through each of the genres in the property array, and compare with each value in the second book's property array of genres.

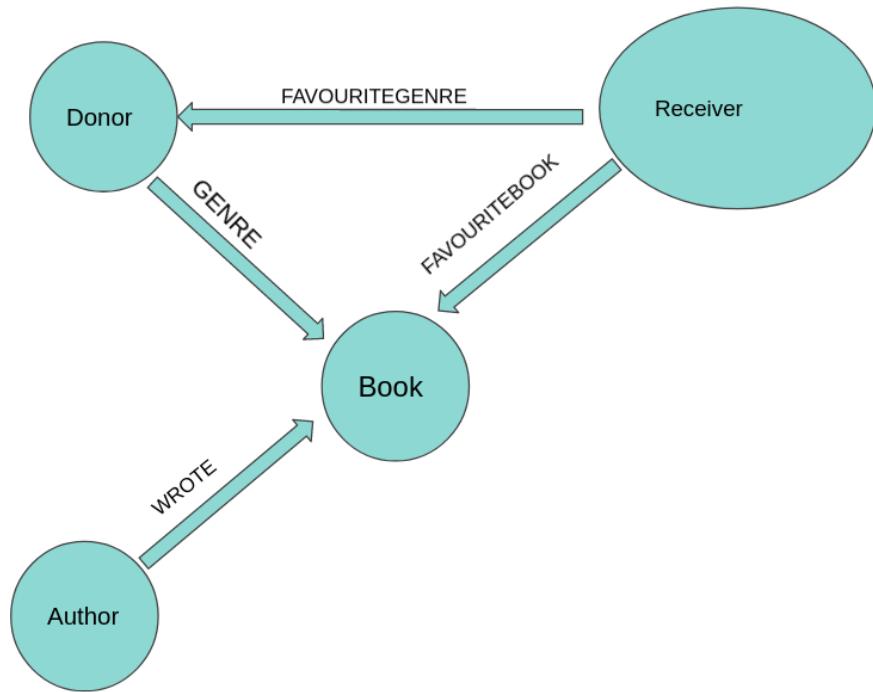
This would take a toll on performance (nested looping and comparison of node properties), and the query would be much more complicated, as well.



Now, instead, if we were to model our books and genres as separate nodes and create a relationship between the two, we would come up with a model something like the image below.

This creates a completely separate entity (node) for the genre, allowing you to connect all the books with a shared genre to that Genre node. Let us see how this changes our queries. To find the genres of a particular book, it first needs to find the Book node it is looking for (in this case, 'The Matrix'), then find the node that is connected to that book through the IN\_GENRE relationship.

## Graph database model - Book Recommendation



## Nodes And their Properties

Nodes	Properties
Genre	name, genre_id
Book	Title, img_url
UserProfileInfo	first_name, last_name, email, username, address, pincode, latitude, longitude, phone

Author

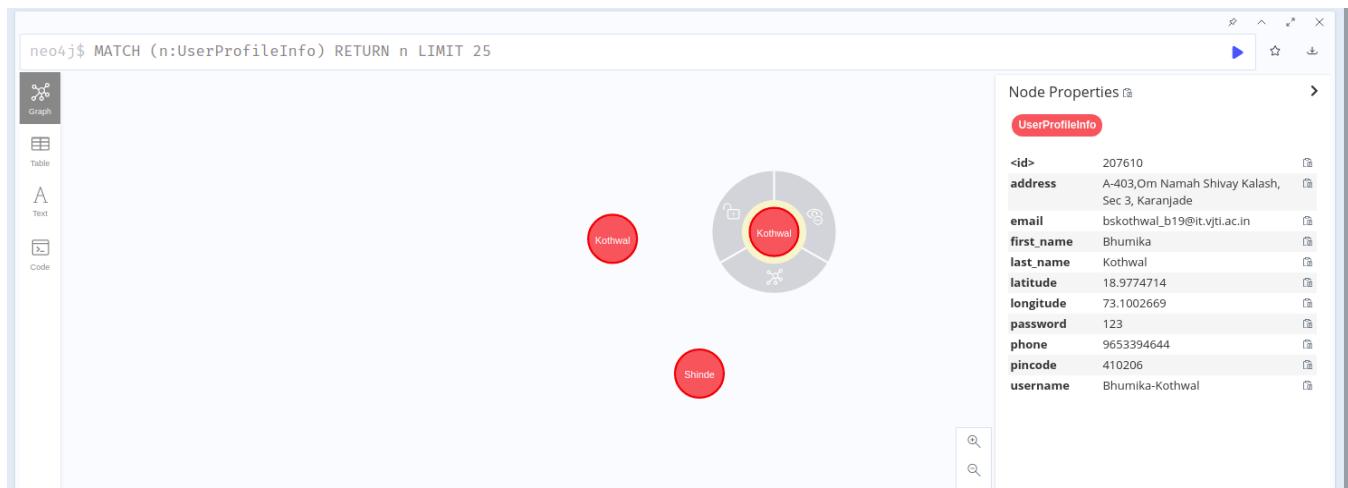
name

## Relationships Between nodes

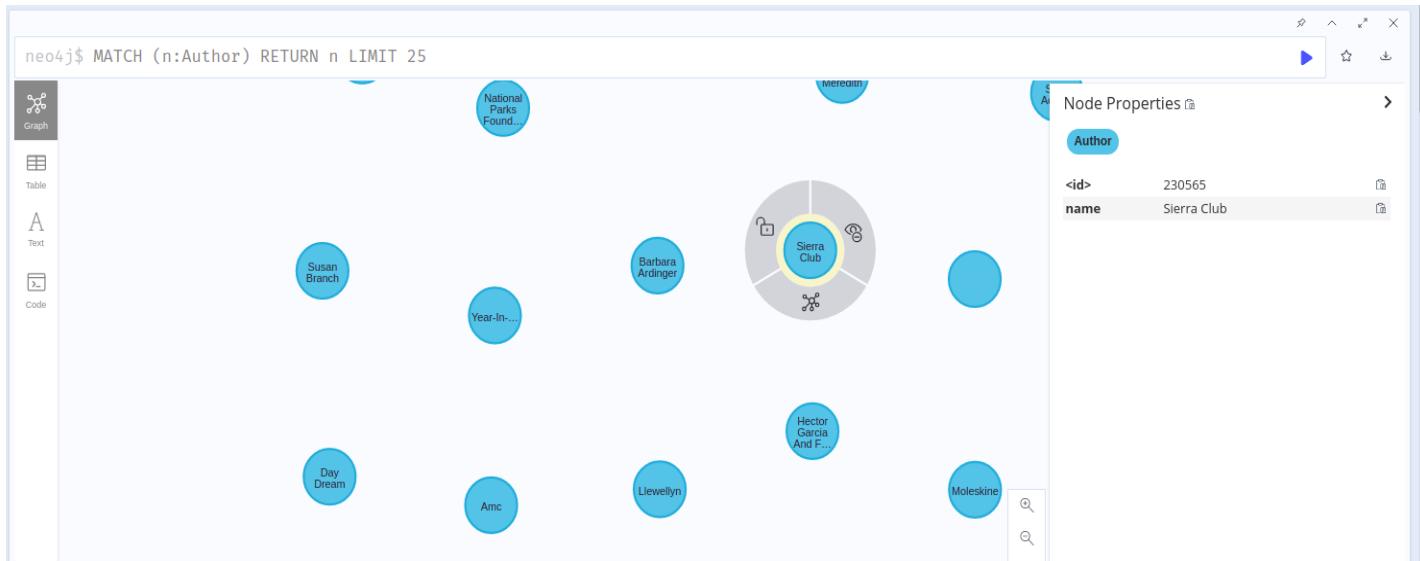
Relations	RelationFrom	RelationTo
FAVOURITEGENRE	UserProfileInfo	Genre
FAVOURITEBOOK	UserProfileInfo	Book
RATING	UserProfileInfo	Book
GENRE	Genre	Book
WROTE	Author	Book

## Graph Database

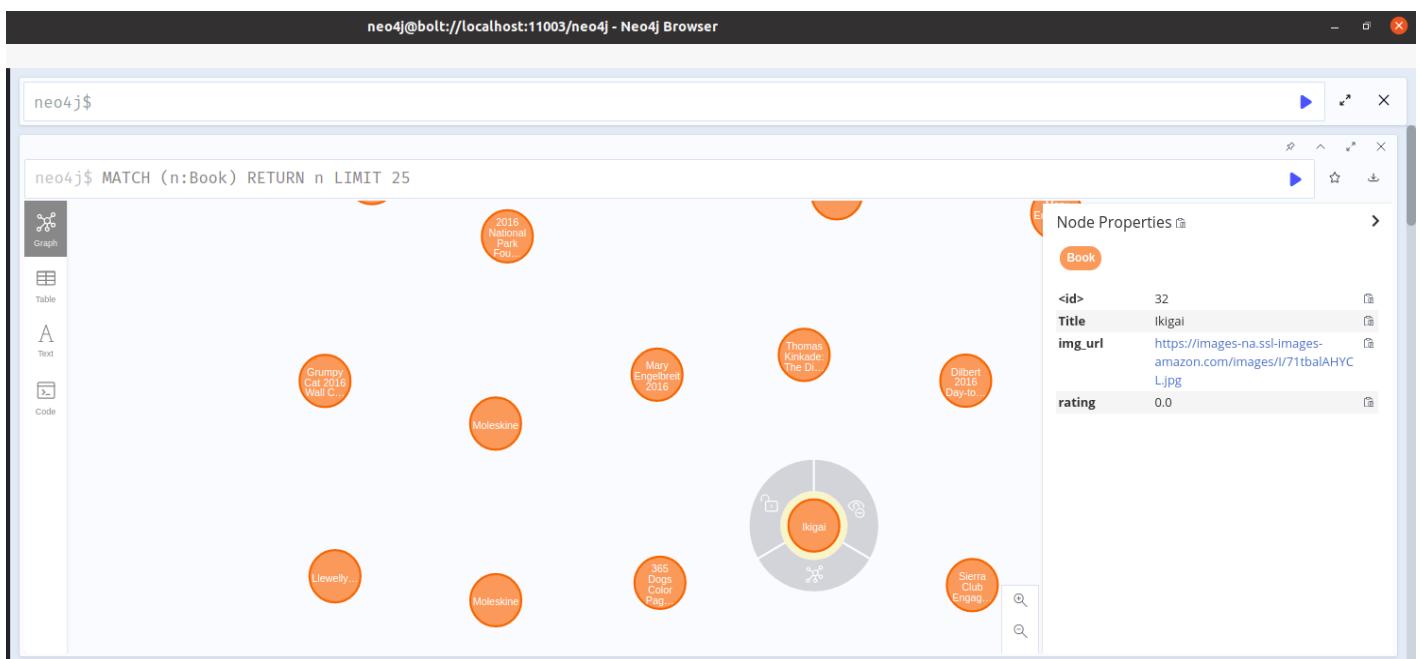
### Nodes : UserProfileInfo Node



## Nodes : Author Node



## Nodes : Book Node

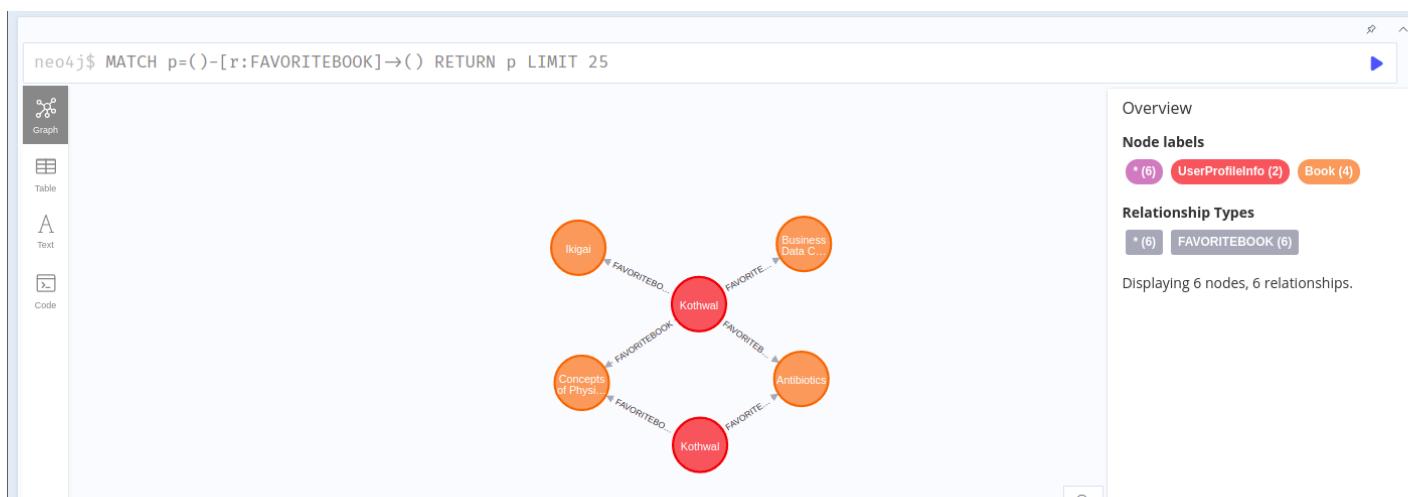


## Node : Genre Node



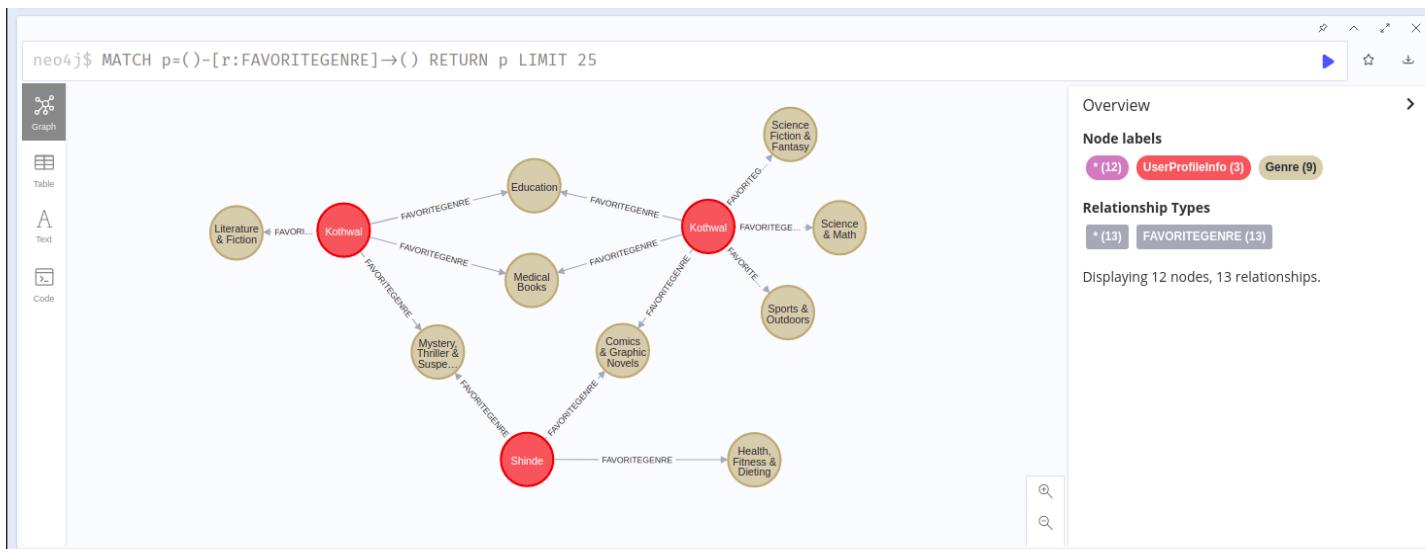
## Relationship :

FavoriteBook Relationship => UserProfileInfo Node to Book Node



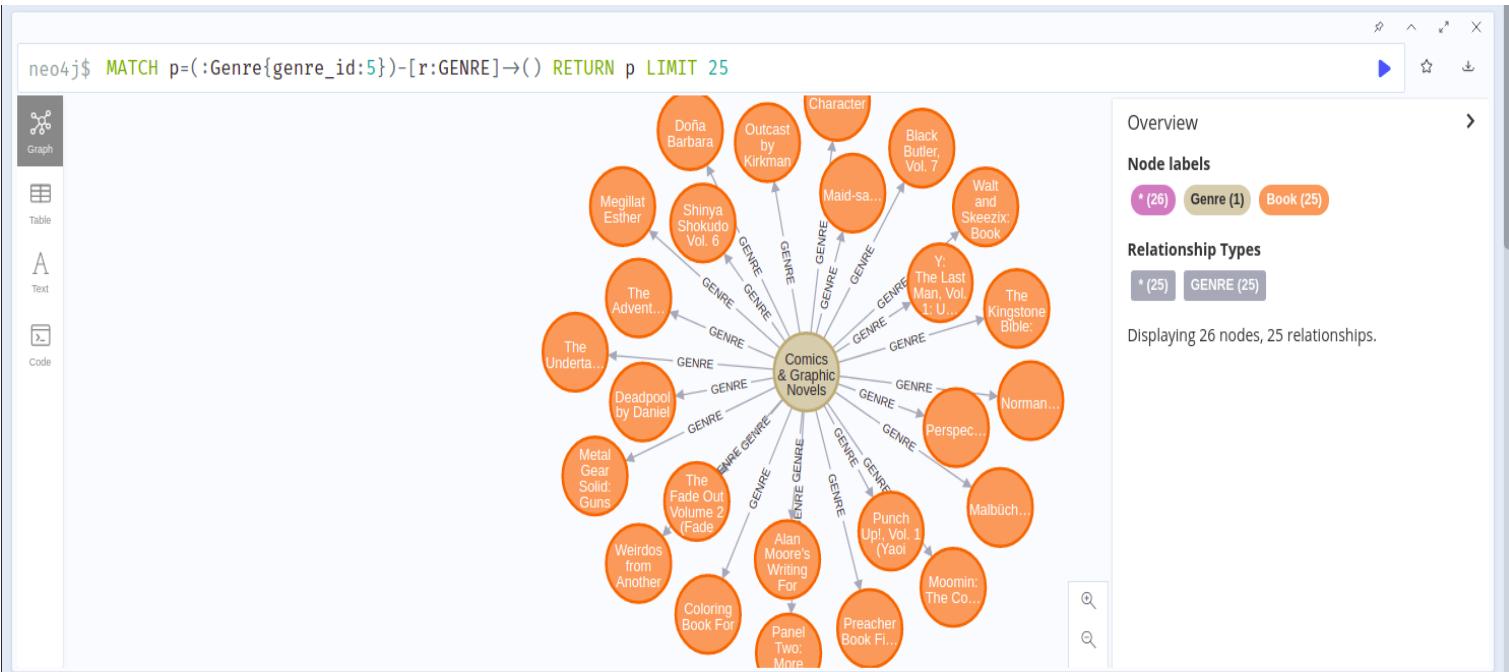
## Relationship :

FavoriteGenre Relationship => UserProfileInfo Node to Genre Node

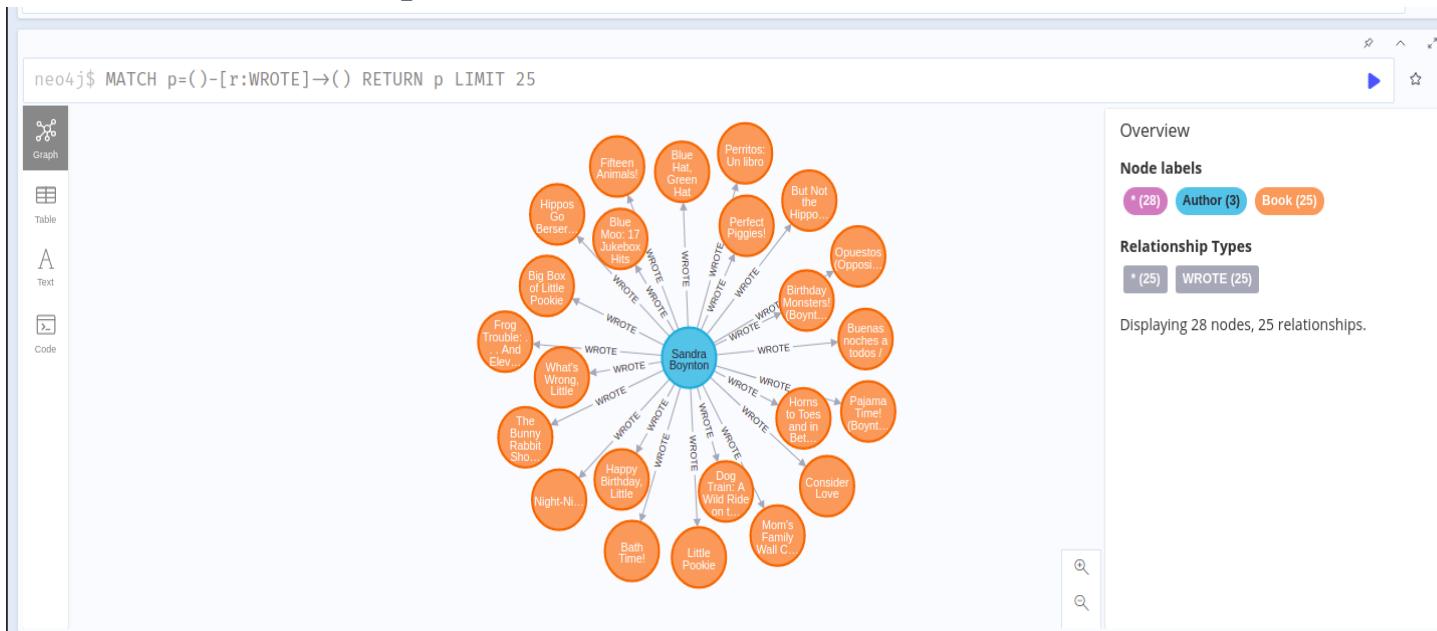


## Relationship :

GENRE Relationship => From Genre Node to Book Node



**Relationship :**  
WROTE Relationship => From Author Node to Book Node



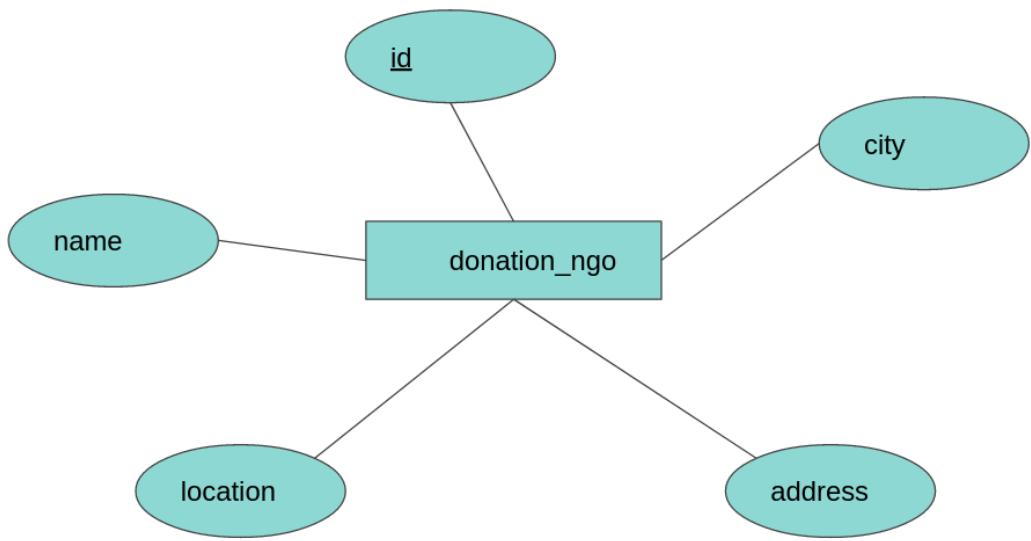
## SPATIAL GRAPH DATABASE TABLES

---

## Table Name : NGOs

Name of column	Data Type
id	integer
name	character
location	geometry(point)
address	character
city	character

## ER DIAGRAM

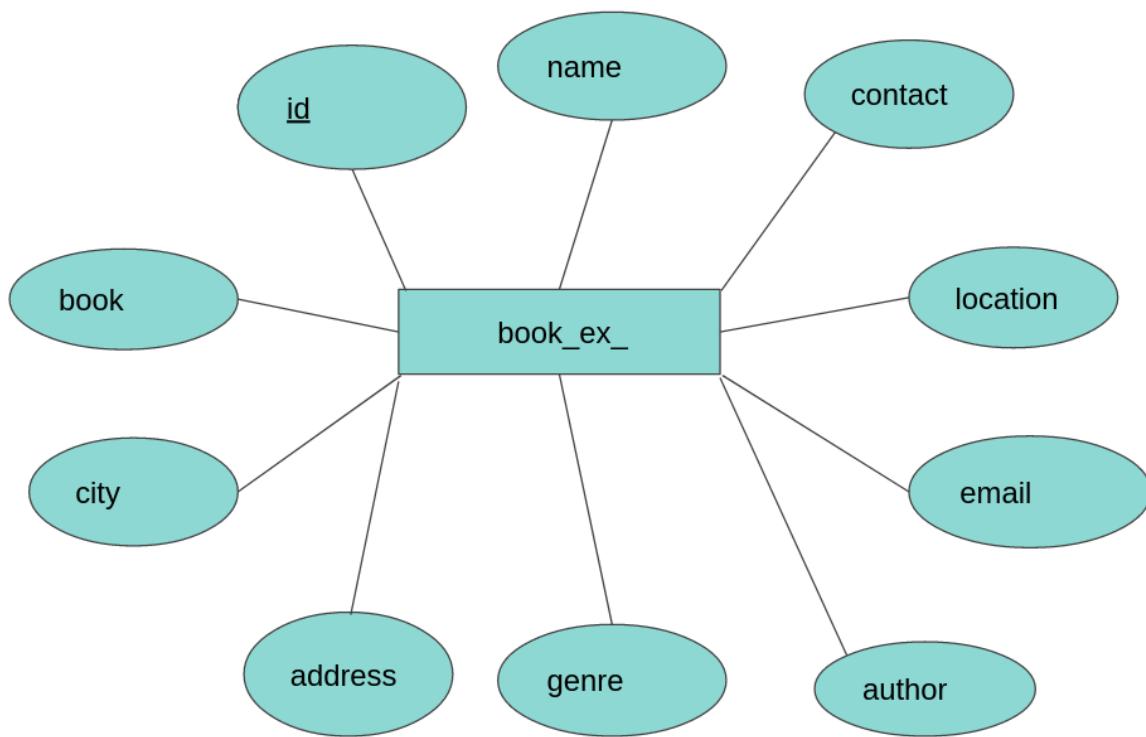


## Table Name: book\_ex

Name of column	Data Type Of Column
id	integer
name	character
location	geometry(point)
address	character
city	character
book	character
author	character

email	character
contact	character
genre	character

## ER DIAGRAM



## Features Applied

### ⌚ Registration for Users

Users can register on our platform to explore all the features .

---

## ◎ **Search**

In this feature users can look for the book as per their interest.

- **Search by name**
- **Search by author**
- **Search by Genre**

## ◎ **Get Recommendation**

In this feature, we recommend books to users based on what other similar users have read and chosen as their favorite book.

## ◎ **Online Price**

You can check the online price of books from different sites and navigate to that site to buy the books online

## ◎ **Connect**

You can connect with other users and exchange the book you want from them. For this purpose we have used the spatial queries and attributes.

- This has options like finding the users within 1km, 5km, 10km, 15km

## ◎ **Save**

Here users can save their favorite genre and favorite books

## ◎ **Exchange Books**

---

Option to fill a form to exchange book you have wherein latitude, longitude, book name and book info are taken and stored in the database, so that when someone wants to exchange that book, he can do the same with you.

### ◎ **Registration For NGOs**

This feature is added to help NGOs so that they can get donated books from our users.

### ◎ **Donate**

- One can donate books to Schools and NGOs to help others to gain maximum knowledge.
- In this feature users are allowed to find the Schools and NGOs nearby their locations. For this purpose we have used the spatial queries and attributes.
- This has options like finding the ngos and schools within 1km, 5km, 10km, 15km

# Implementation

## 1. Home Page



**WELCOME TO SHA-FI-REC-DO**

**Leaders Are Readers**

We are students of Veermata Jijabai Technological Institute, Matunga(VJTI Mumbai) are building this website to enable readers find what they love to read in an efficient and economic way.

[Find Book Now](#)

**Hey Bookworm,**  
Let's Explore The Treasure !!!  
Novel ! Biography ! And Many More !

[LOOKING FOR YOUR FAVOURITE AUTHOR](#)

**Hey Bookworm ,**  
We Have A Tremendous Collection Of Awesome Novels ,.....

[LOOKING FOR YOUR FAVOURITE GENRE](#)



## WHAT WE HAVE ?



Let's Come Together To Make Some Difference !



✉ Connect

### Connect

Connect with other users who can add some great things in your knowledge

[Connect](#)



👤 Donate

### Donate

Let's donate some of your books collection !

[Donate](#)



📍 Find

### Find

It is a long established fact that a reader will be distracted by the readable.

[Find](#)



❤ Share

### Share

Find a book benefactor near you and read your favourite book economically

[Find](#)



👤 Register NGOs

### Register NGOs

Register for your NGOs so that you can connect with the great donors

[Register](#)



📍 Get Recommendation

### Get Recommendation

Get Recommendation of books that suit you the best !

[Learn More](#)



#### Quick links

- » Home
- » About
- » FAQ
- » Contact
- » About

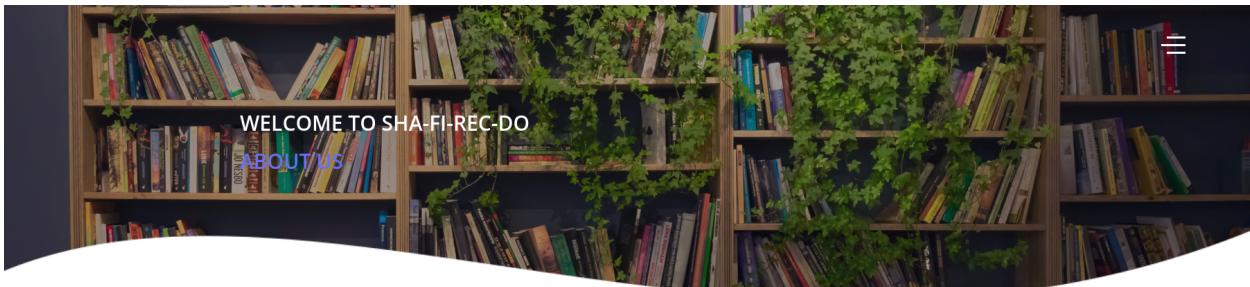
#### Contact Us

- 📞 +91 1234567890  
✉ inspiredcodes@gmail.com

#### Address

Mumbai, India

## 2. About Us



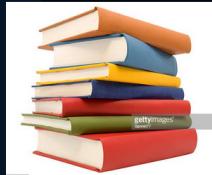
### WE ARE HERE TO BRING THE BEST !

We are students of Veermata Jijabai Technological Institute, Matunga(VJTI Mumbai) are building this website to enable readers find what they love to read in an efficient and economic way.



### WE ARE HERE TO BRING THE BEST !

We are students of Veermata Jijabai Technological Institute, Matunga(VJTI Mumbai) are building this website to enable readers find what they love to read in an efficient and economic way.



OUR MINDSET

- ◉ User-Friendly Book Sharing
- ◉ Best Recommendations
- ◉ Support Educational NGOs
- ◉ Reuse of books

### 3. Register Users



Create an account.

Location

Give access to your location

Latitude

Longitude

Personal Details

First name:

Enter First Name

Last name:

Enter Last Name

Email:

Enter Email

Username:

Enter username

Password:

Enter password

Address:

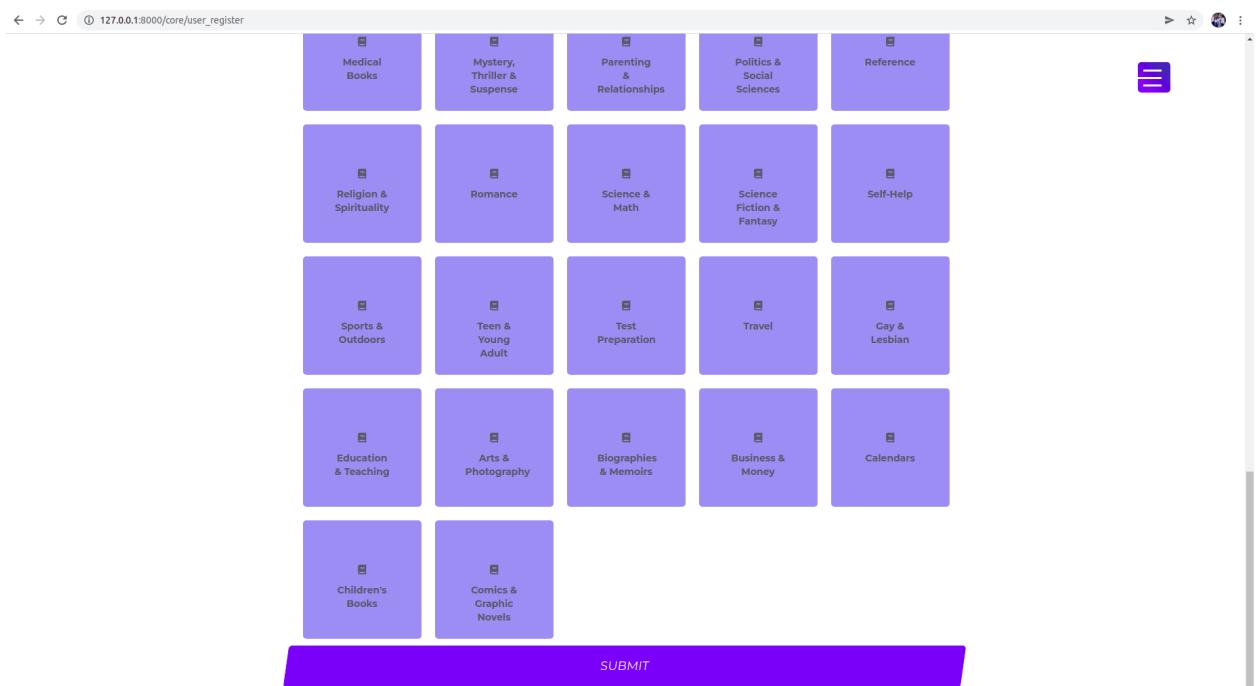
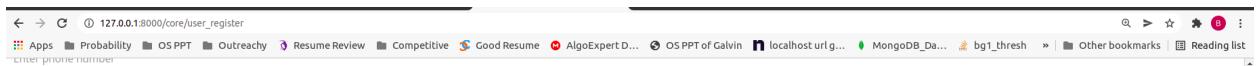
Enter address

Pincode:

Enter pincode

Phone:

Enter phone number



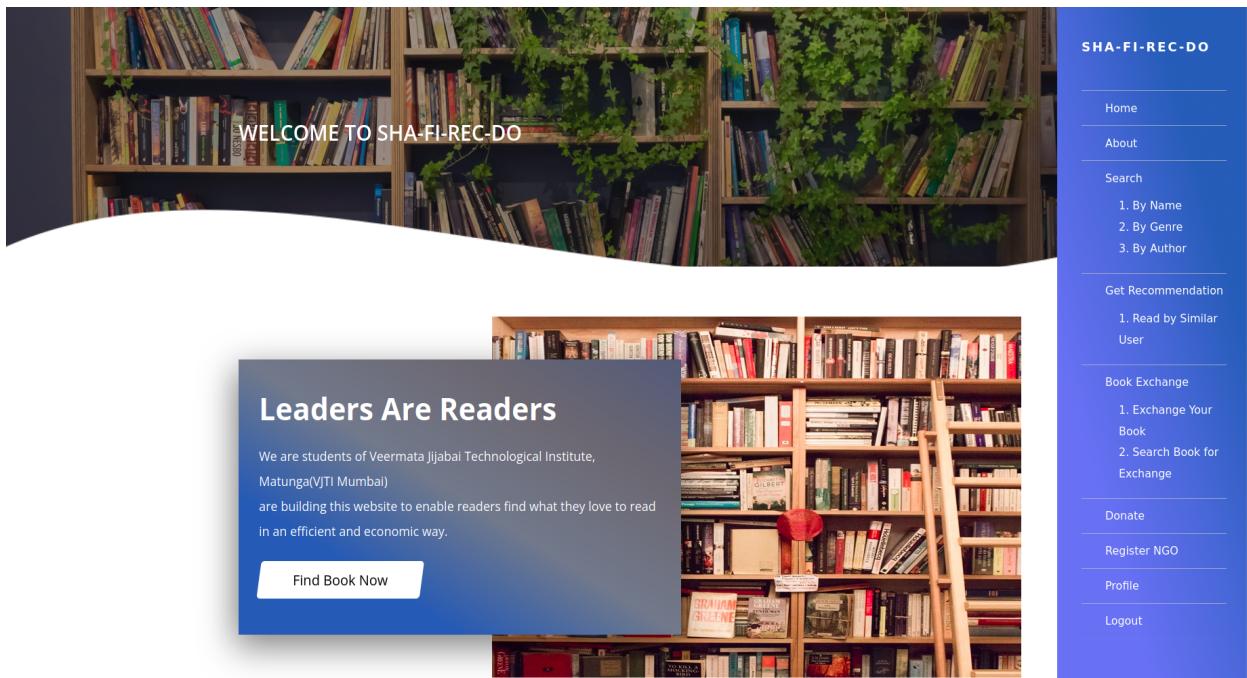
## 4. Login



Username  
 Username

Password  
 password

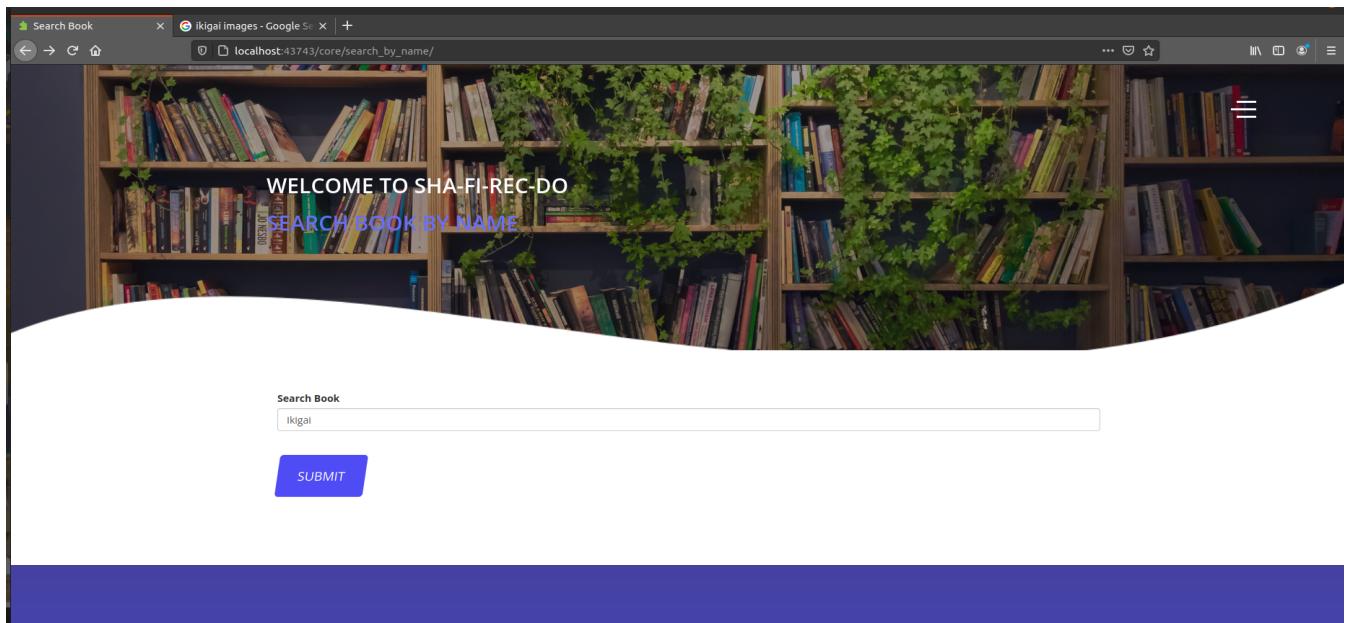
**On Successful Login available features on sidebar get unlocked for user -**



## 5. Search Books

### A. Search Books by Name :

For demo - Searched book name : Ikigai



## Result of searched book -

Search Book

ikigai

SUBMIT

**Searched Book**

**Ikigai**

Authors:

Hector Garcia And  
Frances Miralles

Genres:

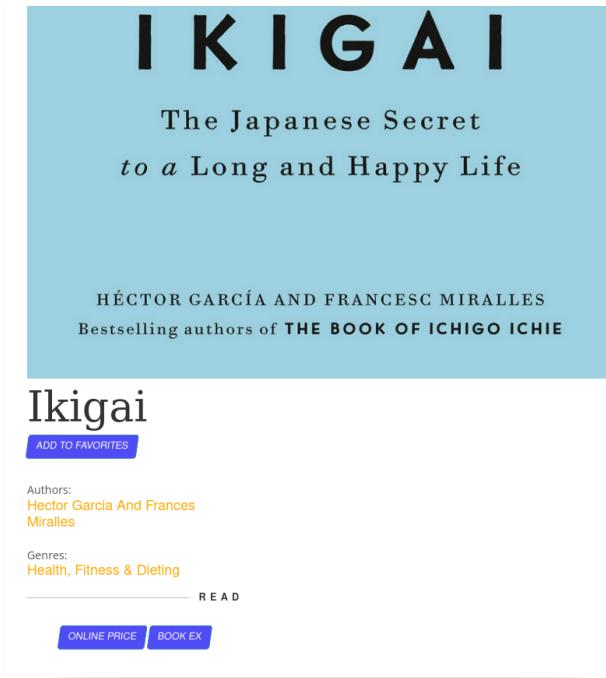
Health, Fitness &  
Dieting

**DETAILS**

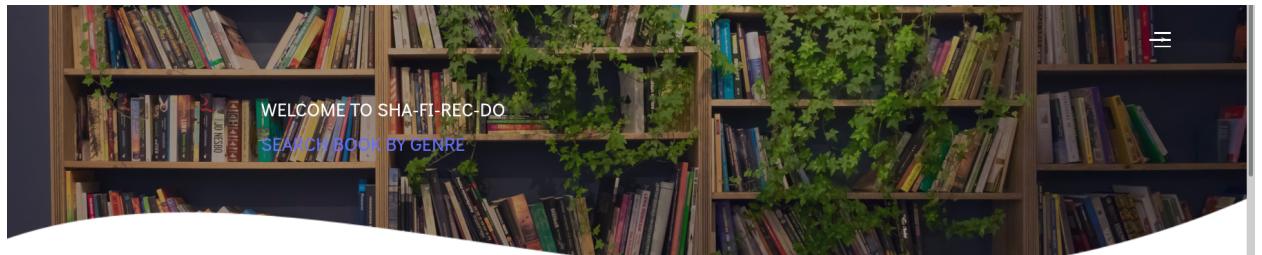
## After clicking on details -



After clicking on details -



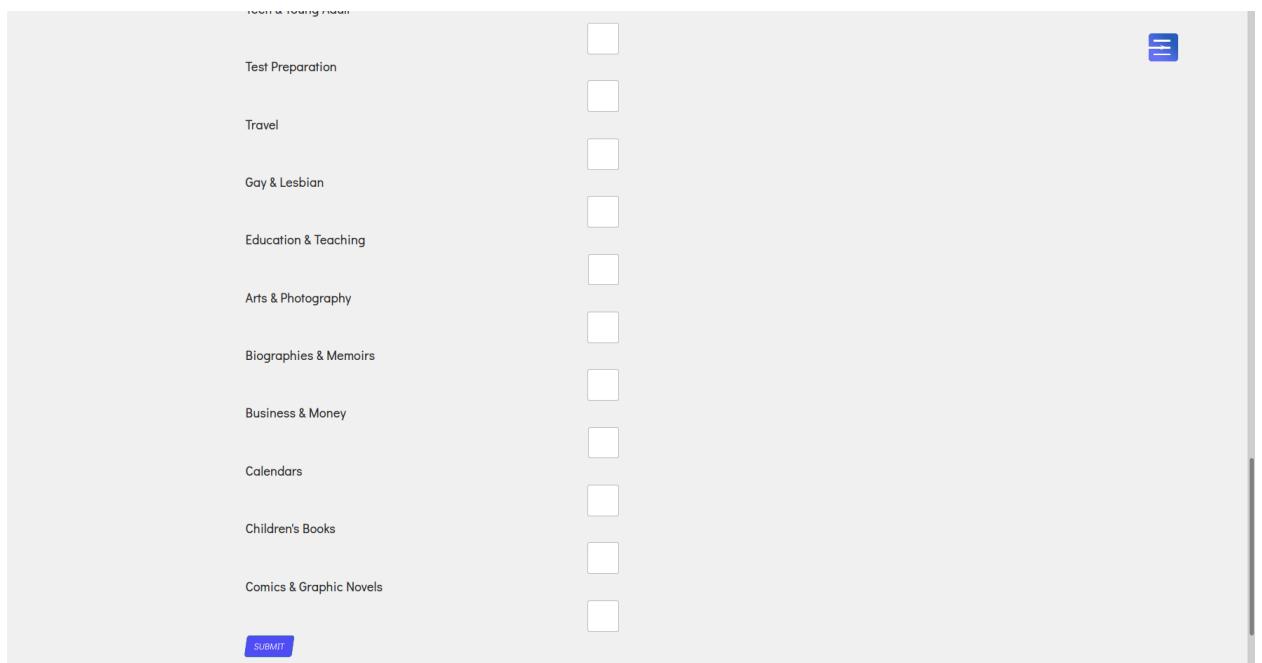
## B. By Genre



WELCOME TO SHA-FI-REC-DO  
SEARCH BOOK BY GENRE

Select the genres:

Computers & Technology	<input type="checkbox"/>
Cookbooks, Food & Wine	<input type="checkbox"/>
Crafts, Hobbies & Home	<input type="checkbox"/>
Christian Books & Bibles	<input type="checkbox"/>
Engineering & Transportation	<input type="checkbox"/>
Health, Fitness & Dieting	<input type="checkbox"/>



Test Preparation

Travel

Gay & Lesbian

Education & Teaching

Arts & Photography

Biographies & Memoirs

Business & Money

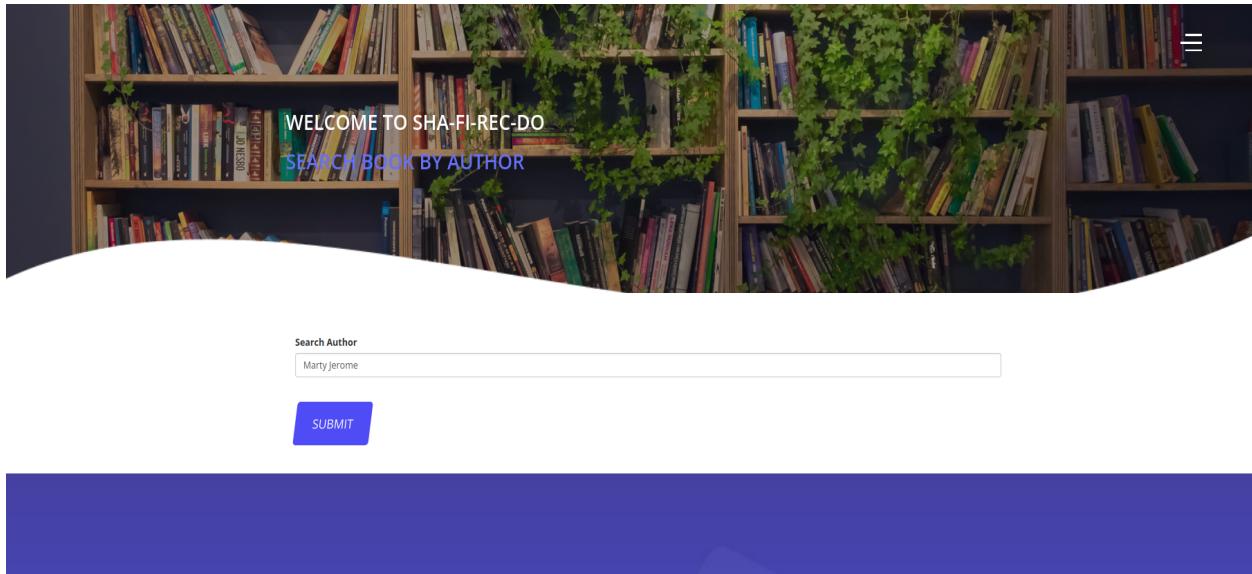
Calendars

Children's Books

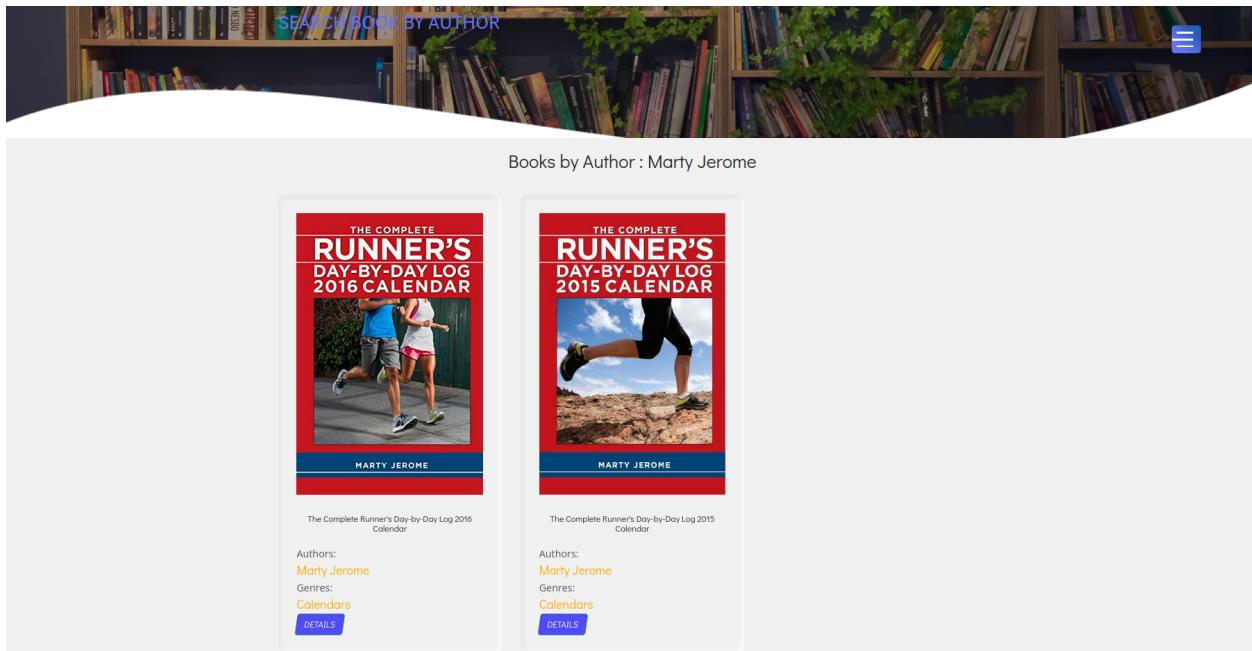
Comics & Graphic Novels

## C. By Author :

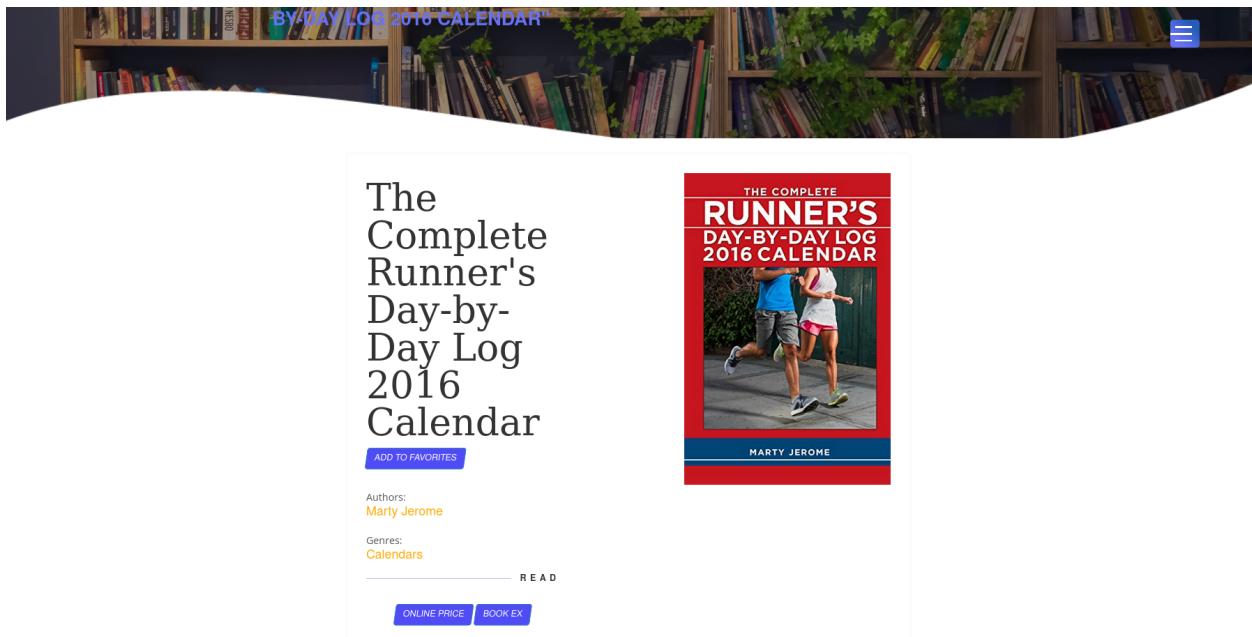
For Demo - Search For Author- Marty Jerome



## Result of searched book :

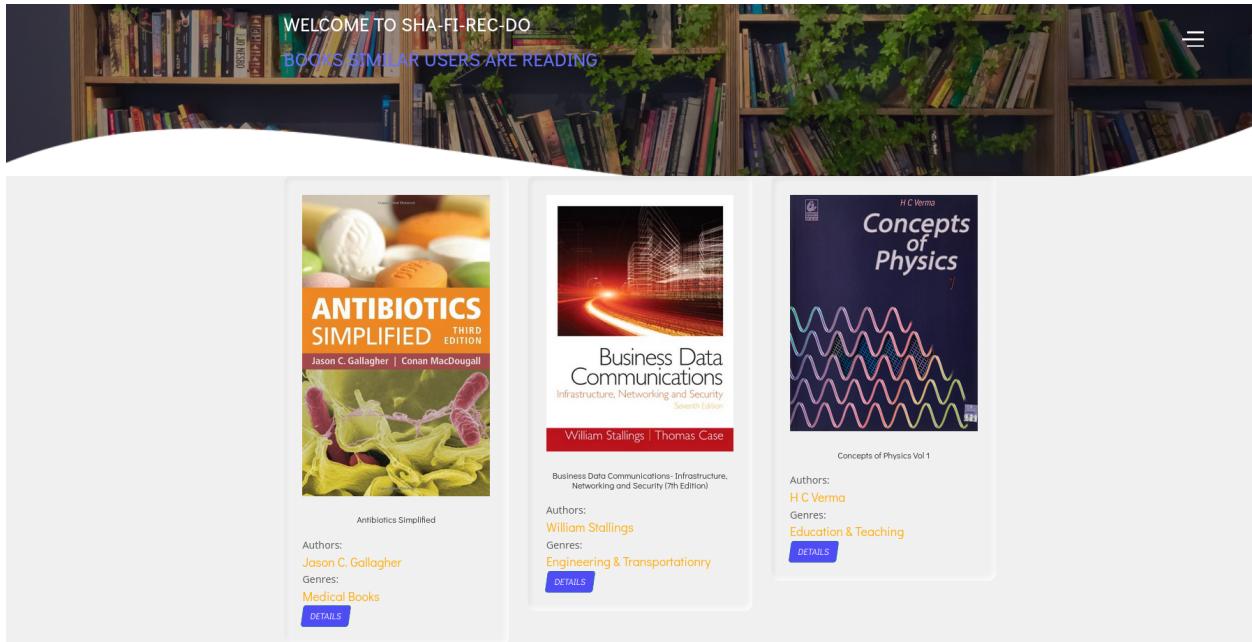


## On clicking Details :



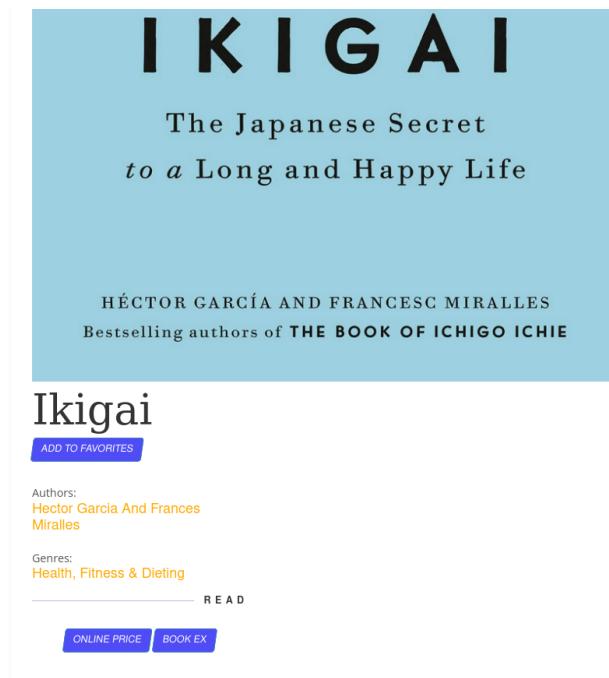
## 6. Get Recommendations :

On selecting feature 'Get Recommendations' on sidebar - Result

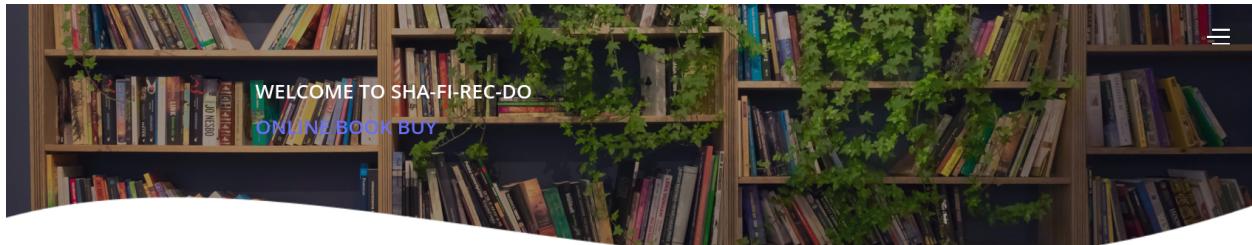


## 7. Checking Online Price :

After clicking on details -



## On clicking online price - (result)



**Amazon**

**Ikigai: The Japanese secret to a long and happy life**

₹285

3.8 out of 5 stars

[CLICK TO BUY](#)

**Flipkart**



**Ikigai-International Best Seller**

₹300

4.3 out of 5 stars

[CLICK TO BUY](#)

**Snapdeal**



**IKIGAI (English, Hardcover, Garcia Hector)**

Rs. 160

4.1 out of 5 stars

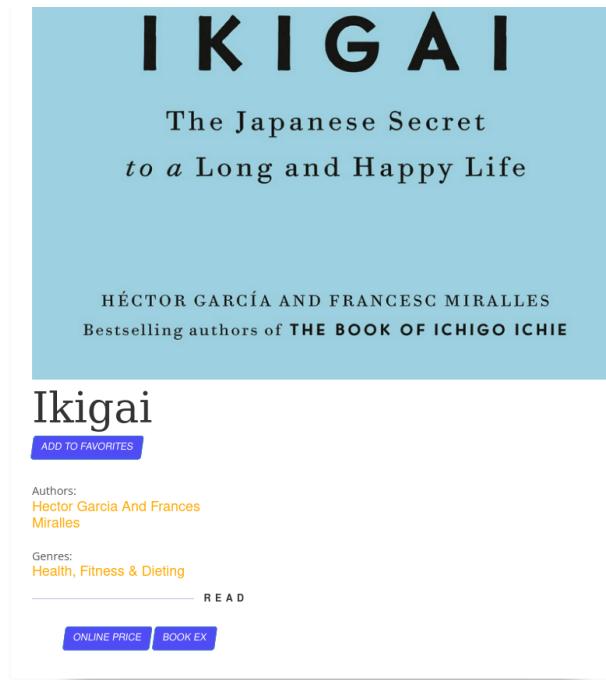
[CLICK TO BUY](#)

## On clicking 'click to buy' - redirects to the shopping website

The screenshot shows the product page for 'Ikigai: The Japanese secret to a long and happy life' on Amazon.in. The main content includes the book cover, title, authors (Héctor García and Francesc Miralles), a 4.5-star rating from 27,913 reviews, and a price of ₹285.00. It also mentions a 5% extra discount offer. Below the main product info, there's a section for 'Save Extra' with three offers, including a 10% instant discount on SBI Credit Card Non-EMI Transactions. The sidebar on the right lists other sellers: Cloudtail India (₹285.00), AmazeBooks (₹358.00), iRead-Store (₹360.00), and another seller (₹375.26). Each listing includes an 'Add to Cart' button.

## 8.Find other users nearby to exchange the book selected :

After clicking on details -



**On clicking 'Book Exchange' - we are headed to find the users to exchange book**

FIND PEOPLE NEAR YOU TO GET YOUR FAVOURITE BOOK!

Find nearest users for exchanging books!

Click me to get your location

Latitude  
18.9774813

Longitude  
73.1001814

Select a Distance Range  
less than 10 km

SUBMIT

On clicking 'Submit' - we get results as list of users wanting to exchange book chosen

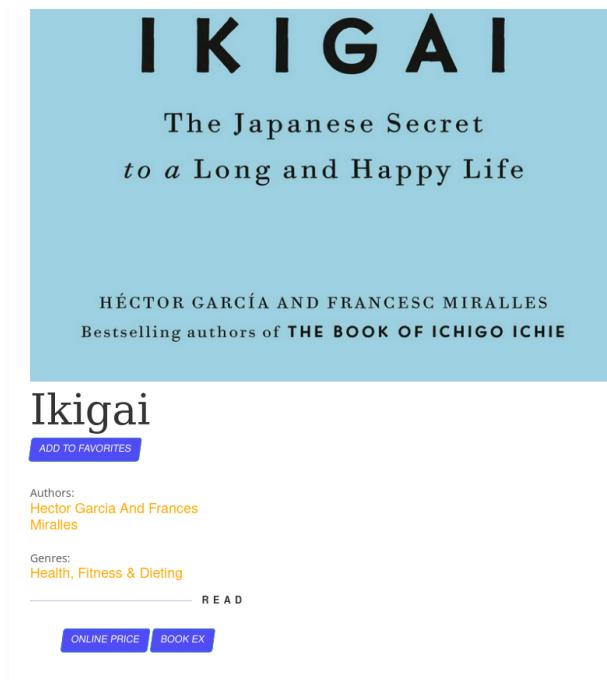
The screenshot shows a web browser window with the URL `127.0.0.1:8000/book_ex/near_book/`. The page has a decorative header featuring a bookshelf and plants. The main content area is titled "Find nearest users for exchanging books!". It contains fields for "Latitude" and "Longitude", a dropdown menu for "Select a Distance Range" (set to "less than 1 km"), and a "SUBMIT" button. Below this, a section titled "Your Results!" displays a table with one row of data. The table columns are: Name, Address, Book Title, Email, Contact, Distance, and Directions. The data in the table is:

Name	Address	Book Title	Email	Contact	Distance	Directions
Bhumika Kothwal	A-403, Jai Sambhav LK Planet, Sector 6, Karanjade	Ikigai	kothwalbhumika@gmail.com	9653394644	0.0037 km	

Information of users wanting to exchange book

## 9. Adding books to favorites :

After clicking on details -



On clicking 'Add to favourites' - get added to your favourite books collection

#### **Favorite Books**

Title	Authors	Genres	Remove from favorites	
Ikigai	Hector Garcia And Frances Miralles	Health, Fitness & Dieting	<b>REMOVE</b>	<b>DETAILS</b>
Antibiotics Simplified	Jason C. Gallagher	Medical Books	<b>REMOVE</b>	<b>DETAILS</b>
Business Data Communications- Infrastructure, Networking and Security (7th Edition)	William Stallings	Engineering & Transportation	<b>REMOVE</b>	<b>DETAILS</b>
Concepts of Physics Vol 1	H C Verma	Education & Teaching	<b>REMOVE</b>	<b>DETAILS</b>

**ADD BOOKS TO FAVORITES**

Ikigai is now added in favourite books list -(Can be seen in User Profile )

## 10. Enter Information to exchange the book you have with other users

WANT TO EXCHANGE YOUR BOOK ?

Give your details to be eligible for book exchange!

Click me to get your location

Latitude  
18.9774502

Longitude  
73.1001957

Name  
Bhumika Kothwal

City  
Panvel

Address  
A-403, Jai Sambhav LK Planet, Sector 6, Karanjade

Book Title  
Ikigai

Book Image URL  
http://127.0.0.1:8000/book\_ex/bookex\_form/

**City**  
Panvel

**Address**  
A-403, Jai Sambhav LK Planet, Sector 6, Karanjade

**Book Title**  
Ikigai

**Book Image URL**  
<https://images-na.ssl-images-amazon.com/images/I/71tbalAHYCL.jpg>

**Author**  
Hector Garcia And Frances Miralles

**Genre**  
Health, Fitness & Dieting

**Email**  
kothwalbhumi@gmail.com

**Contact**  
9653394644

**SUBMIT**

**Quick links**

- » Home
- » About
- » Search

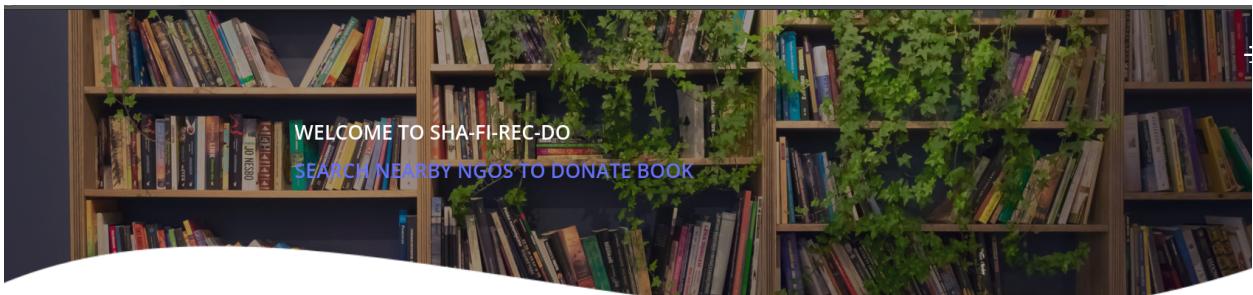
**Contact Us**

📞 +91 1234567890  
✉️ shafiredo@gmail.com

**Address**

Mumbai, India

## 11.Donate



**Find nearest NGOs and Educational Institutions for donating books!**

**Latitude**  
18.5204303

**Longitude**  
73.8567437

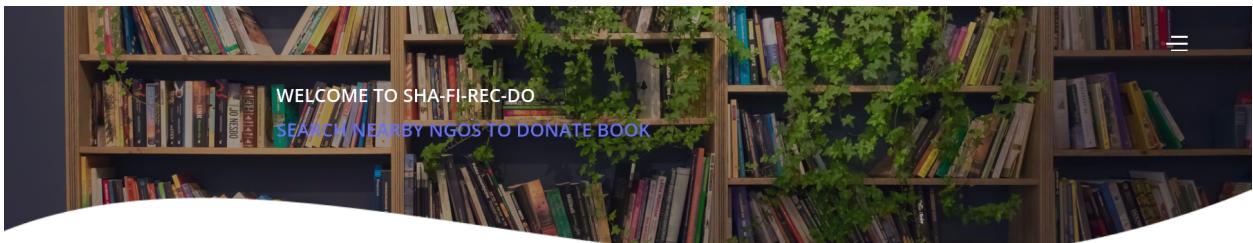
**Select a Distance Range**  
less than 1 km

Quick links

Contact Us

Address

**Result for nearest NGOs :**



**Find nearest NGOs and Educational Institutions for donating books!**

**Latitude**

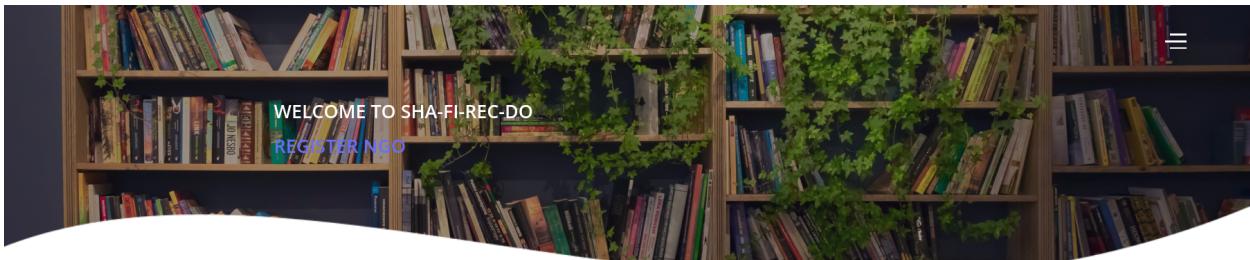
**Longitude**

**Select a Distance Range**  
less than 1 km

Your Results!

Name	Address	Distance	
LiveLaughLove	Visawa naka ,Satara	0.0153 km	<a href="#">Get Directions</a>

## 12. Register NGOs to get donation from other users:



Give your details to be eligible for receiving book donations!

Name of Organization	LiveLaughLove
City	Satara
Address	Visawa naka, Satara
Latitude	17.5951
Longitude	73.9595

**SUBMIT**

## 13. Profile :



### Personal Details

Full Name	UserName
Neha Shinde	nbshinde
Email	Phone
nehashinde3399@gmail.com	9067264451

### Address

Sundara Gardens, Flat No.304, Bldg.No-5, Visawa Naka, Satara 415002

### Location

Latitude 18.5204303      Longitude 73.8567437

## Favorite Books

Title	Authors	Genres	Remove from favorites	
Ikigai	Hector Garcia And Frances Miralles	Health, Fitness & Dieting	<button>REMOVE</button>	<button>DETAILS</button>

[ADD BOOKS TO FAVORITES](#)



## Favorite Genres

Genre	Remove from Favorites
Comics & Graphic Novels	<button>REMOVE</button>
Mystery, Thriller & Suspense	<button>REMOVE</button>
Health, Fitness & Dieting	<button>REMOVE</button>

## Other genres

Select Genre to add to Favorites

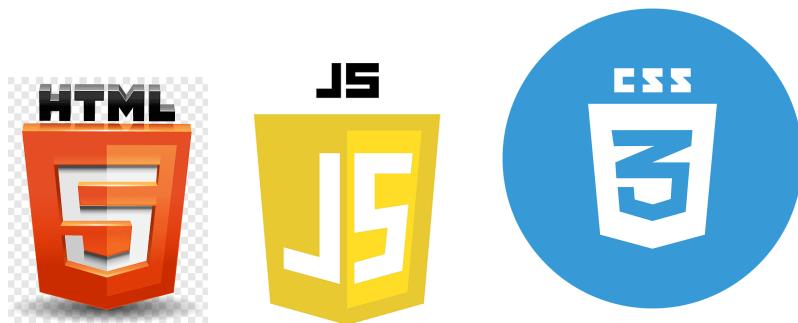
[ADD](#)

## Tech Stack

### Backend :



### Frontend:



### Database:



---

## 1 . Django

Django is a high-level Python web framework that enables rapid development of secure and maintainable websites. Built by experienced developers, Django takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel.

## 2 . Python

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

## 3 . HTML

HTML (Hypertext Markup Language) is the code that is used to structure a web page and its content. For example, content could be structured within a set of paragraphs, a list of bulleted points, or using images and data tables.

## 4 . CSS

Cascading Style Sheets is a style sheet language used for describing the presentation of a document written in a markup

---

language such as HTML. CSS is a cornerstone technology of the World Wide Web, alongside HTML and JavaScript.

## 5 . JS

JavaScript (JS) is a lightweight, interpreted, or just-in-time compiled programming language with first-class functions. While it is most well-known as the scripting language for Web pages, many non-browser environments also use it, such as Node.js, Apache CouchDB and Adobe Acrobat. JavaScript is a prototype-based, multi-paradigm, single-threaded, dynamic language, supporting object-oriented, imperative, and declarative (e.g. functional programming) styles.

## 6 . Postgres

PostgreSQL is used as the primary data store or data warehouse for many web, mobile, geospatial, and analytics applications. The latest major version is PostgreSQL 12.

## 7 . Postgis

PostGIS is a spatial database. Oracle Spatial and SQL Server (2008 and later) are also spatial databases.

---

## 8 . Pgadmin

PgAdmin is a web-based GUI tool used to interact with the Postgres database sessions, both locally and remote servers as well. You can use PGAdmin to perform any sort of database administration required for a Postgres database

## 9 . Neo4j

Neo4j facilitates personal data storage and management: it allows you to track where private information is stored and which systems, applications, and users access it. The graph data model helps visualize personal data and allows for data analysis and pattern detection

---

## References

- <https://www.ethanrosenthal.com/2015/11/02/intro-to-collaborative-filtering/>
- <https://www.analyticsvidhya.com/blog/2021/06/build-book-recommendation-system-unsupervised-learning-project/>
- <https://www.oracle.com/in/big-data/what-is-graph-database/#how-graph-database-works>

## Future Scope

Although this web application has been well-established to provide a platform for designers to evaluate recommender systems comprehensively using different evaluation metrics, there is still some work to do in the future.

- Expand to support more algorithms such as user-based, item-based and biased matrix-factorisation. The application now only supports collaborative filtering method. As the application is readily extensible for

recommendation algorithms, it is expected to include more other algorithms, such as hybrid algorithms, content-based algorithms, demographic-based algorithms, etc.

- Include more evaluation criteria. In literature, there are so many evaluation metrics proposed to measure the performance of recommender systems. Also, there are also so many different variations of a given metric, for example, many approaches for measuring coverage.

## Conclusion

We successfully completed the project by adapting knowledge of Spatial and Graph Database management systems .

**THANK YOU !**

....