

Semesterprojekt Konzept

Titel & Intro - **Bookommender**

Bookommender ist eine Buchempfehlungsplattform, die auf Basis von Ratings und Metadaten dem Benutzer Bücher empfiehlt.

Screens zur Erklärung, Userführung inkl. Beschreibung

<https://www.figma.com/file/nNdUxhhrwFVdLQVt1iS9iW/Bookomender?t=2TAbZzNq0GWR4!rx-0>

Login: Alter, Location, Metadaten (Lieblingsautor, Verlag) angeben
(optional UserId eingeben um sich als bestehender User auszugeben zum Debuggen)

Home: Overview Seite
Rating von X Büchern -> Vorschlägen
Most Popular + Least Popular (für Cold Start Problem), Suche

Predictions:

- top buch von user im selben land ganz oben
- if(ratings < 5) most + least popular
- else if (search) search results + predictions
- else predictions UBCF (Alter, Location, <3)

Details: Buch Details

Predictions: item-based (Author, Verlag, Jahr)

Favoriten: Zeigt die Favoriten des Users an + Empfehlungen
Predictions: Item-based Favoriten: ähnliche Bücher wie die Favoriten des Users

Findings / Zusammenfassung Datenexploration

280 000 **User** mit Ort und Alter (Alter nicht bei jedem)

1.5 Mio **explizite Ratings** 1-10

270 000 **Bücher** mit Titel, Autor, Publikationsjahr, Verlag, Links zum Cover

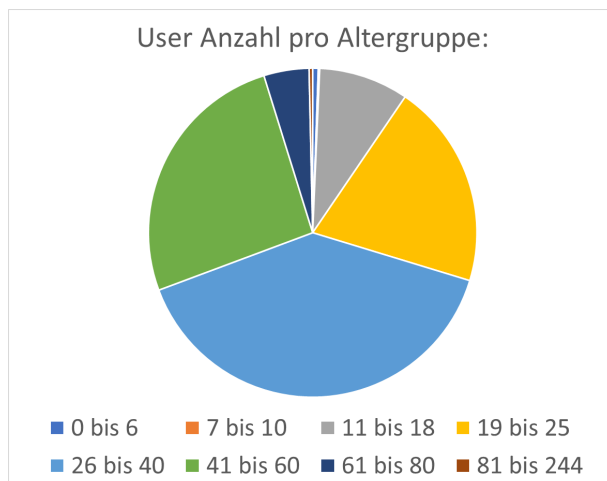
Average Rating (ohne 0): 7.6

Rating hat viele 0 (nicht bewertet)

16780 User haben ein Alter angegeben (bei den anderen NULL)

Alter von 0-244 -> ungültige Werte müssen behandelt werden

User Altersverteilung:



```

NULL      110762
24         5687
25         5618
26         5547
23         5456
...
168         1
132         1
159         1
186         1
127         1
Name: Age, Length: 166, dtype: int64
    
```

Ratingverteilung:

```

0.0      651330
8.0       95037
10.0      71814
7.0       70018
9.0       62101
5.0       46953
6.0       33753
4.0        8031
3.0        5445
2.0        2492
1.0         1601
Name: rating, dtype: int64
    
```

User ID mit Rating Anzahl:
User haben sehr viele Ratings

```

11676      13602
198711     7550
153662     6109
98391      5891
35859      5850
212898     4785
278418     4533
76352      3367
110973     3100
235105     3067
230522     2991
16795      2948
234623     2674
36836      2529
52584      2512
245963     2507
204864     2504
55492      2459
185233     2448
171118     2421
Name: user, dtype: int64
    
```

Additional Information APIs: (sinnvoll sind Preis und Genre, optional)

- OpenLibrary: <https://openlibrary.org/api/books?bibkeys=ISBN:9789513114725&jscmd=data&format=json> (Subjects: Fantasy, Fiction, school stories, Juvenile Fiction, Magic,)

Features

- Buch-Bewertung mit Sternen 1-10
- Autor, Verlag und Jahr in Vorschläge einbeziehen (**IBCF**)
- Alter und Ort in Vorschläge einbeziehen (**UBCF**)
- Suche nach Buchtitel -> wird bei Vorschlägen berücksichtigt + Suchergebnisse
- Favoriten setzen -> Rating wird stärker gewichtet (item-based mit Author)
- Geratete Bücher nicht empfehlen

Ansatz um Generalisierung & Performance sicherzustellen

Cold-Start: erstmal die beliebtesten Vorschläge (**Most Popular + Least Popular**)

Datensplit: k-Fold Cross Validation

Evaluierung: Precision & Recall, Rangmetrik (MRR (Mean Reciprocal Rank))

Algorithmus/Methode: Faktorisierungsmaschine FM

TechStack & Architekturansatz

Backend: FastApi <https://fastapi.tiangolo.com/>

Frontend: React mit Cookie Session

Docker + Docker Compose zum Starten und Containerisierung

Recommender Library:

- Spotlight: for matrix faktorisierung <https://github.com/maciejkula/spotlight> (python last updated 3 years ago)
- Implicit: CF for implicit datasets, logistic matrix faktorisierung, bayesian personalized ranking <https://github.com/benfred/implicit> (python last updated 3 months ago)
- CaseRecommender: for both explicit and implicit Feedback, lots of algorithmsn <https://github.com/caserec/CaseRecommender> (python last updated 1 year ago)
- TensorFlow: Decision Forest
https://www.tensorflow.org/recommenders/examples/basic_retrieval
https://www.tensorflow.org/decision_forests
- Pytorch/Keras

Grobe Aufwandsschätzung der Hauptkomponenten

in Fibonacci Zahlen [0,1,2,3,5,8,13,21,34,55]

Infrastructure (Docker): ~5h

Frontend: ~ 8h

Backend REST Api: ~ 5h

Recommender System:

Evaluierung: ~ 13h

Datenaufbereitung: ~ 13h

Baseline (Most Popular, UBCF, IBCF): ~ 8h

extra Features (Favoriten): ~ 13h

Welcome

Enter your age:

Enter your city:

Enter your country:

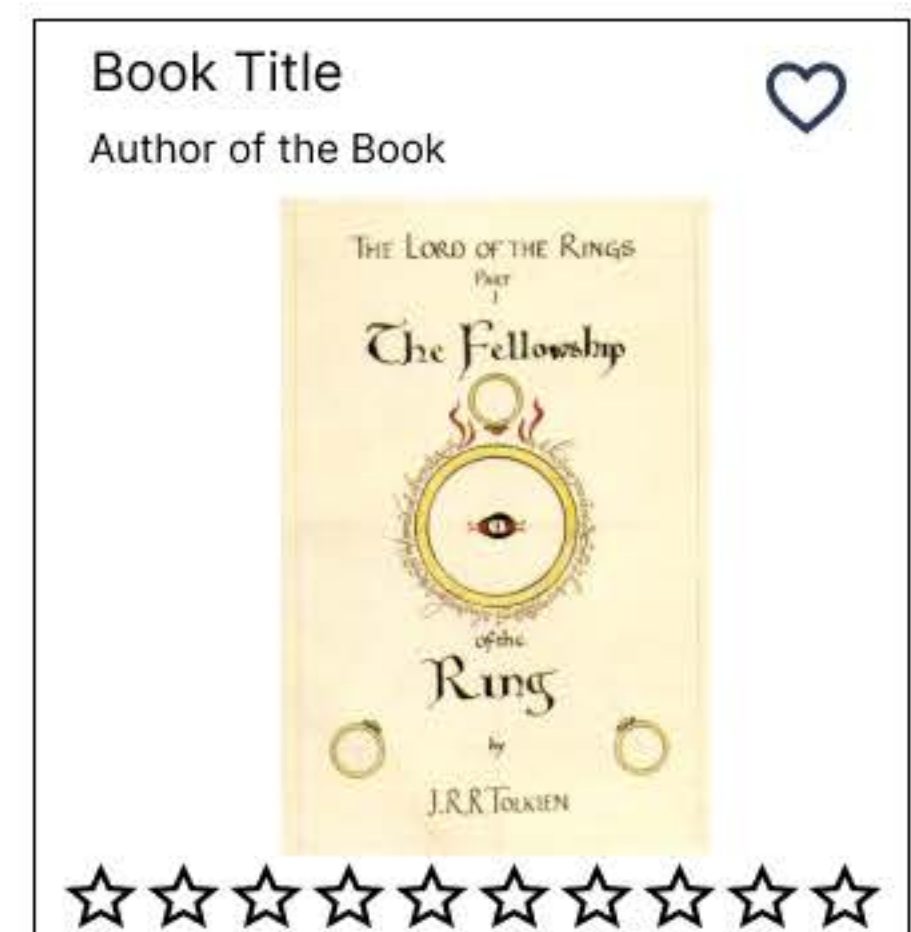
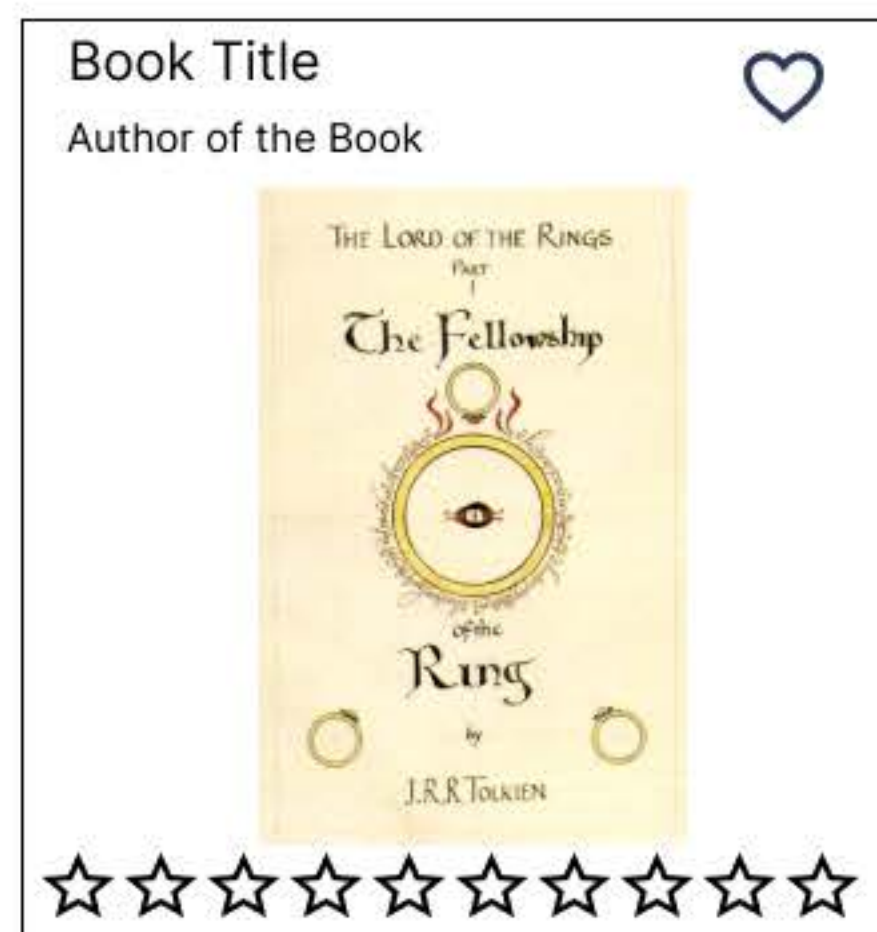
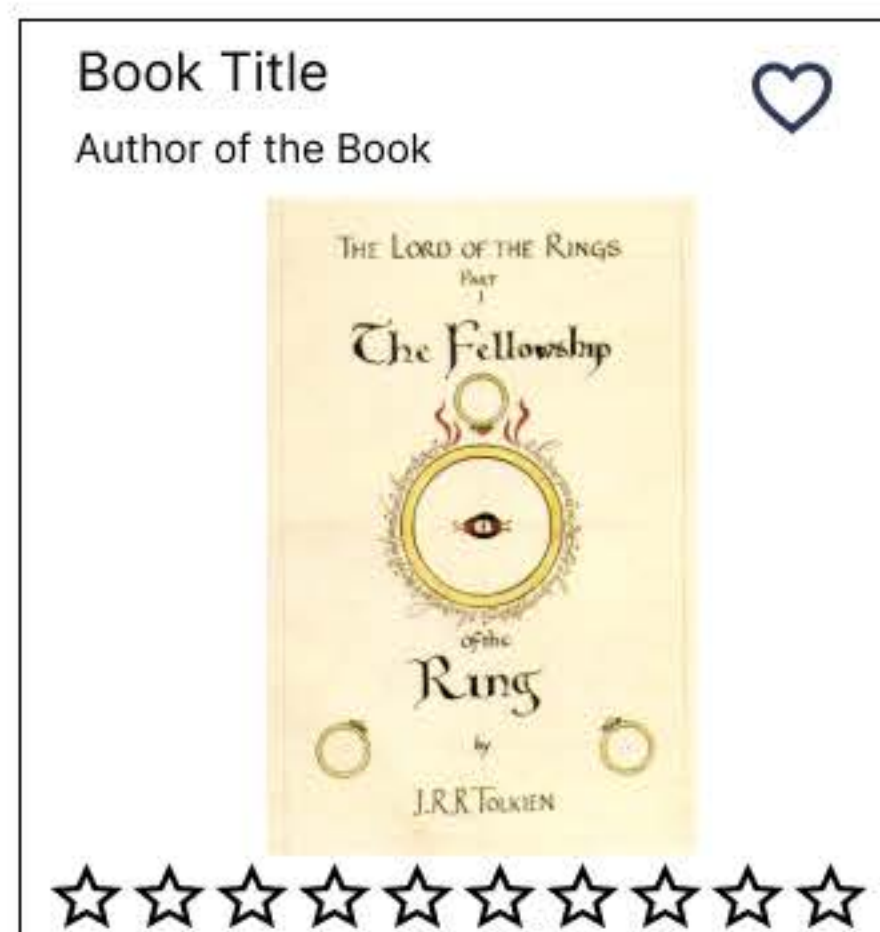
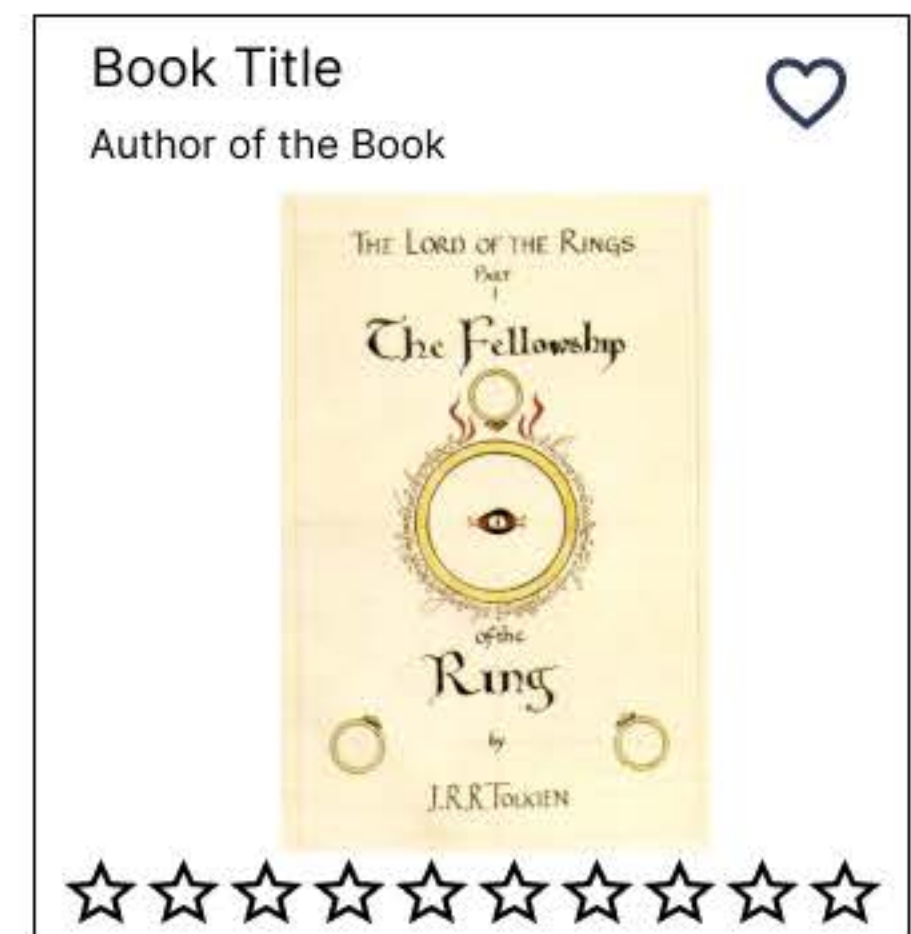
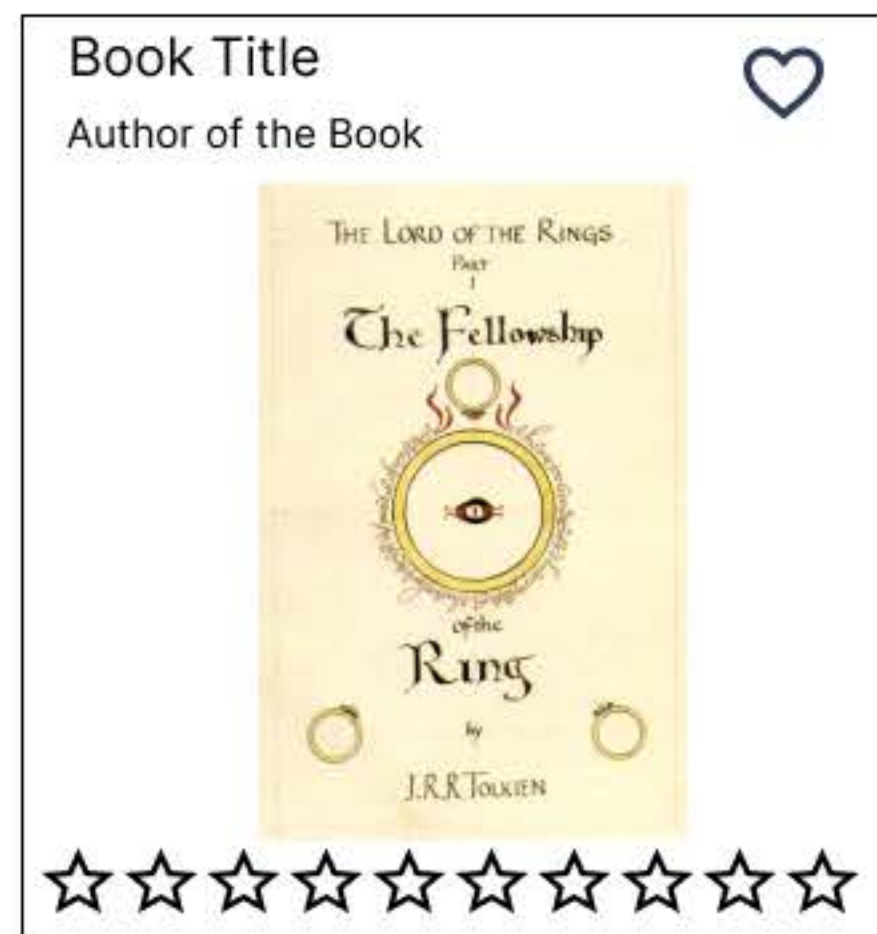
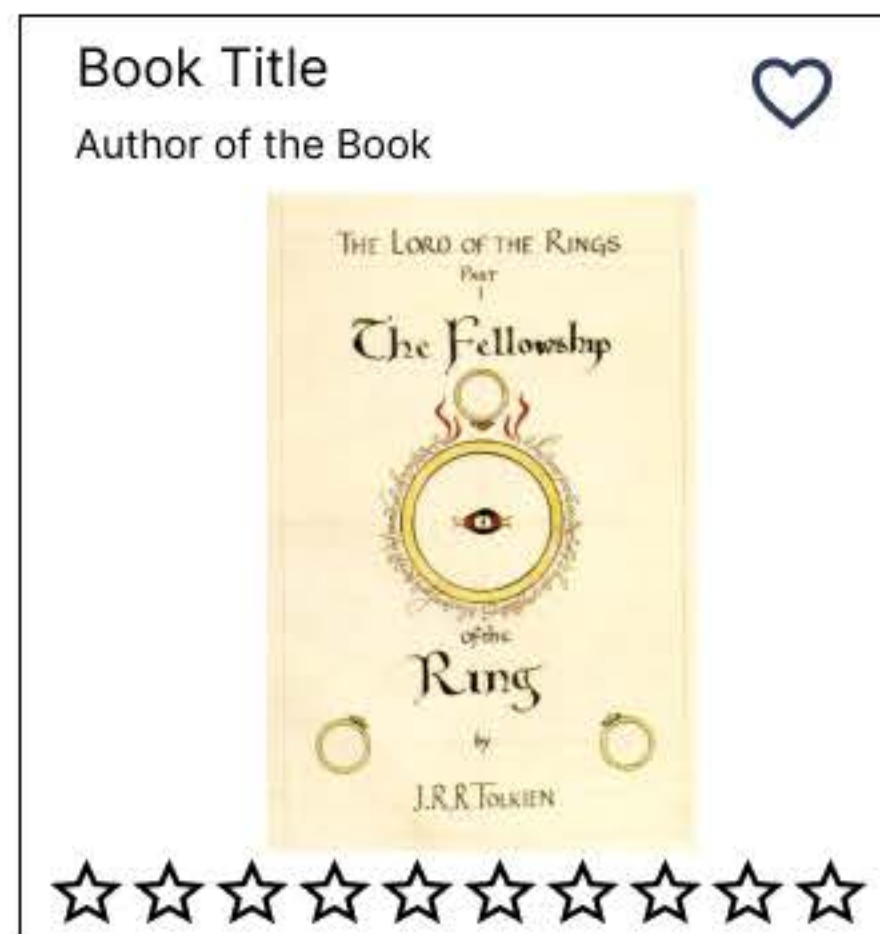
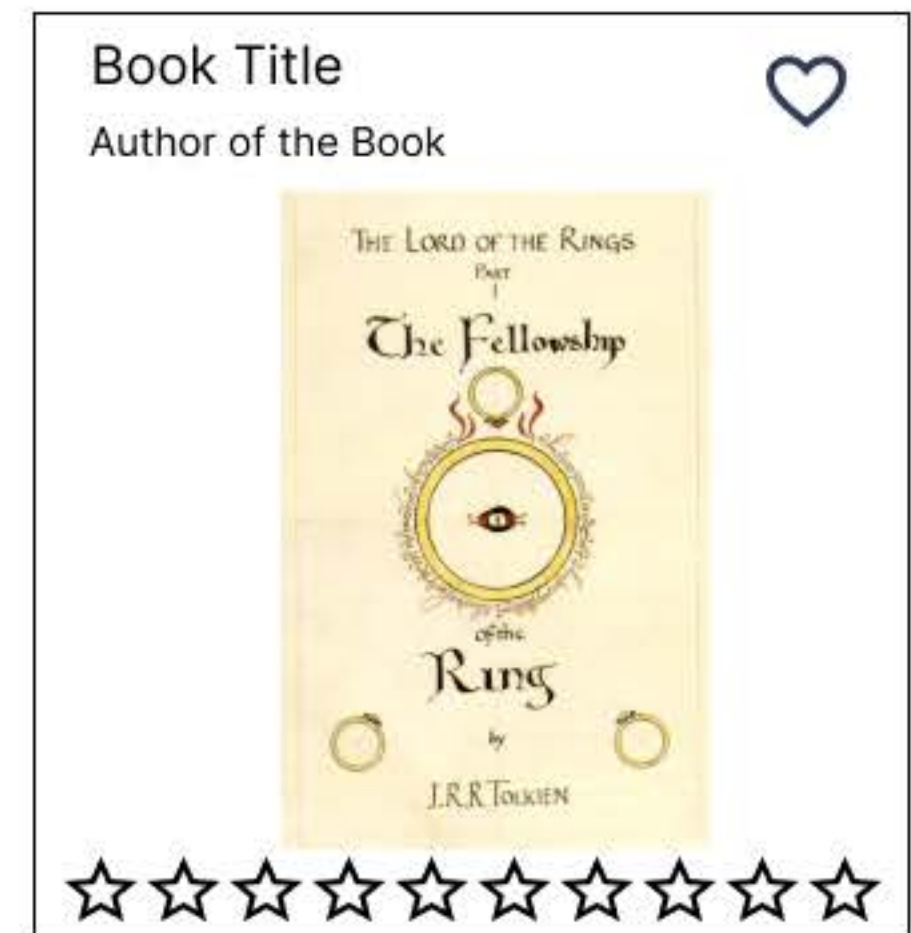
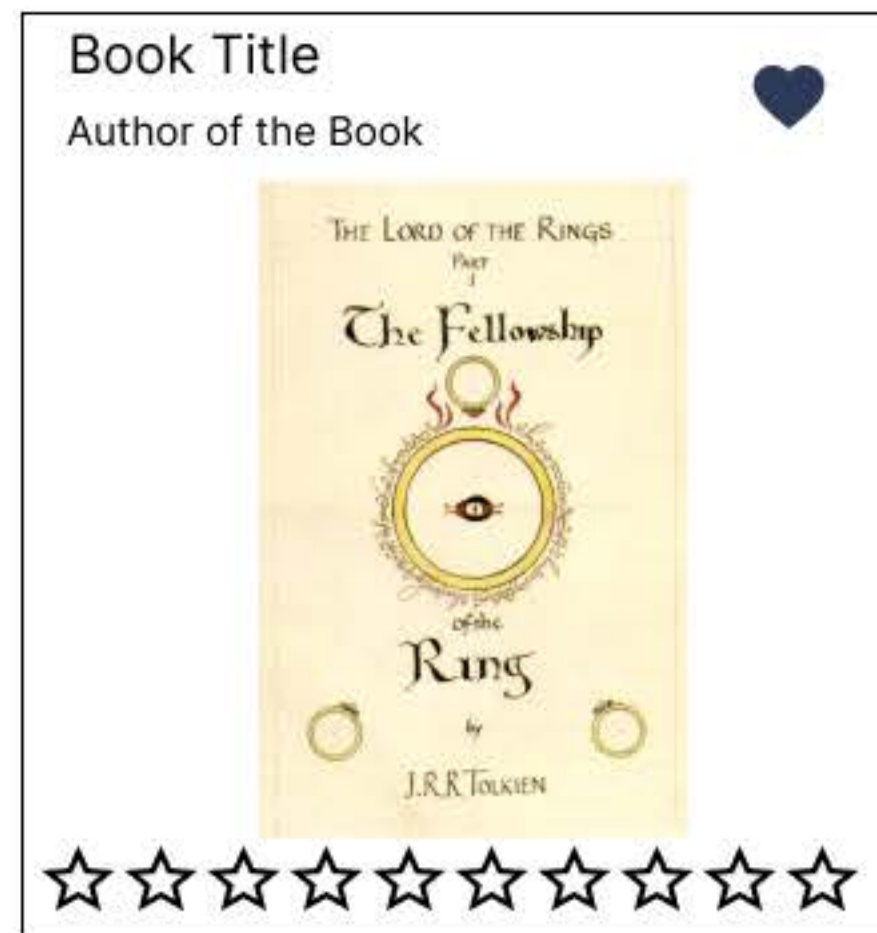
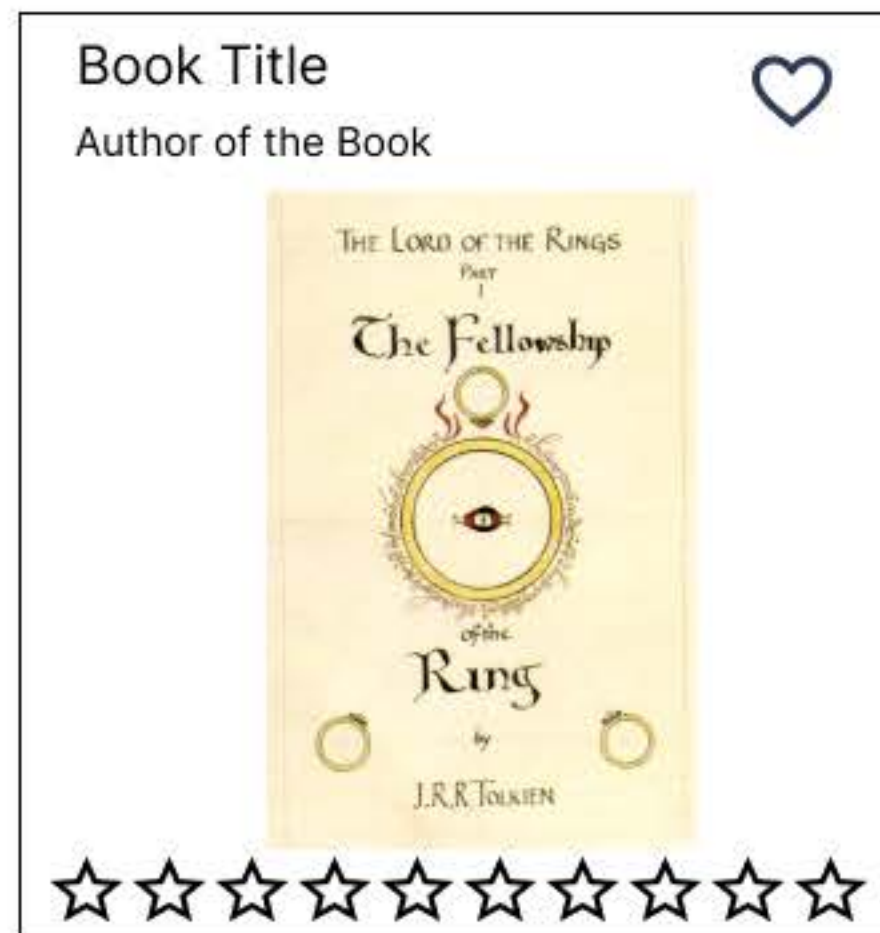
Enter your state:

Enter your favourite author:

Enter your favourite publisher:

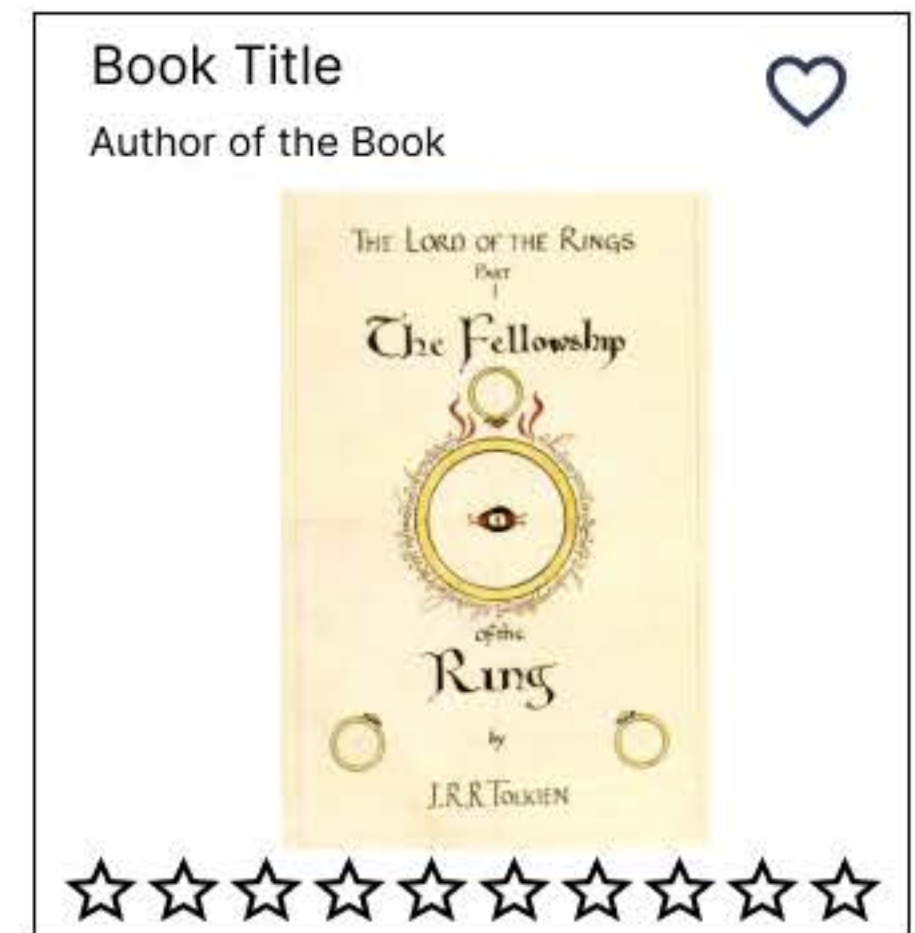
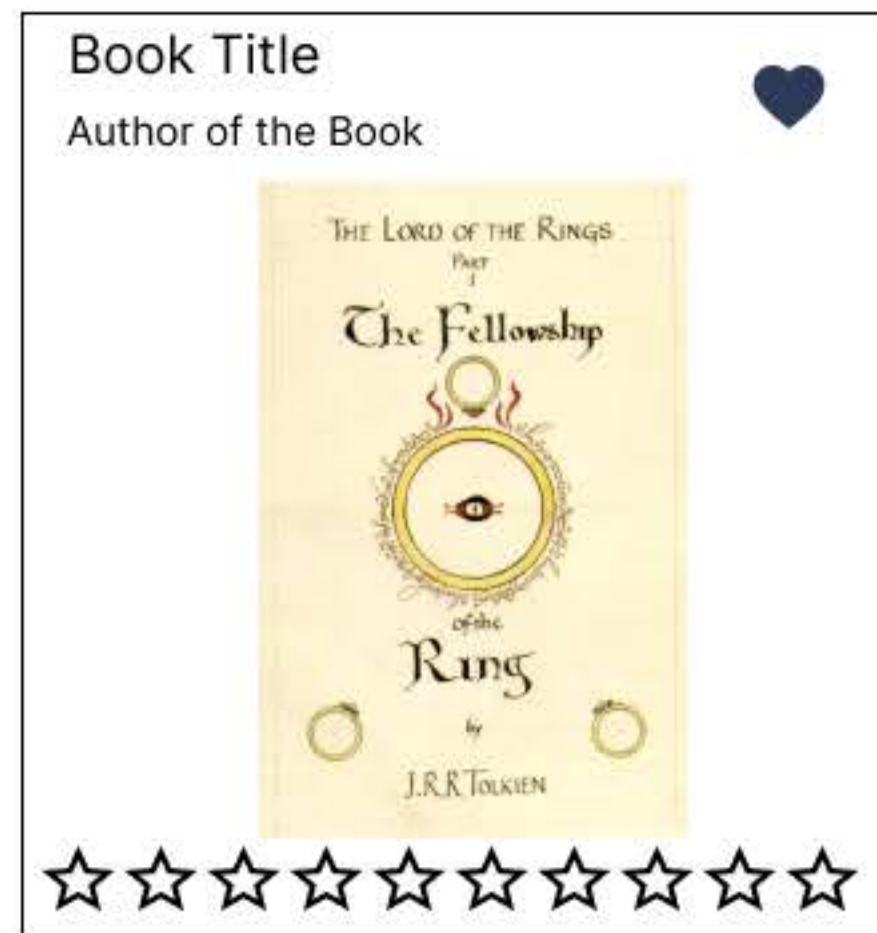
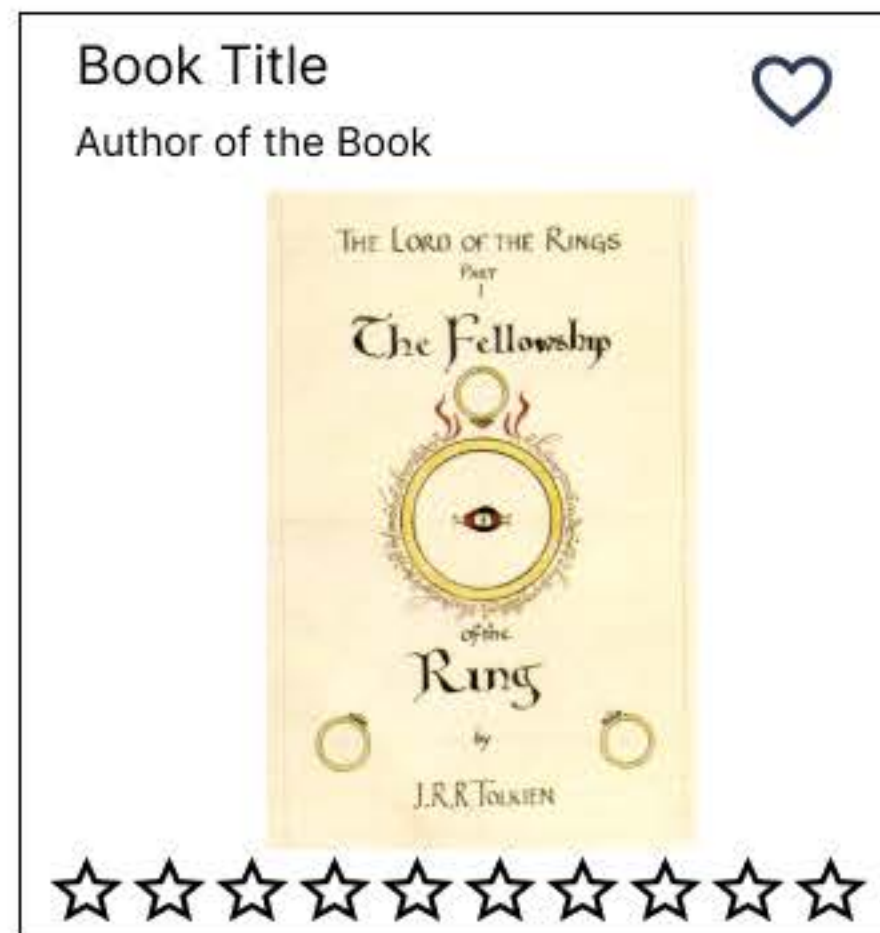
Continue

Recommended for you

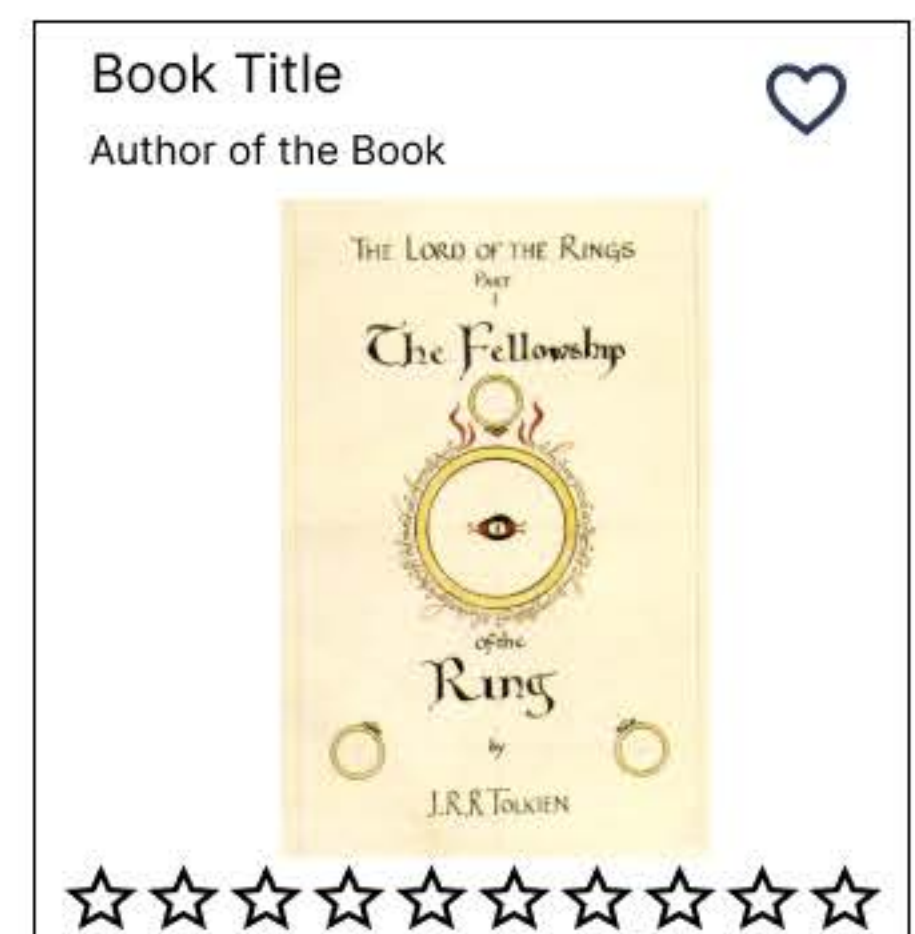
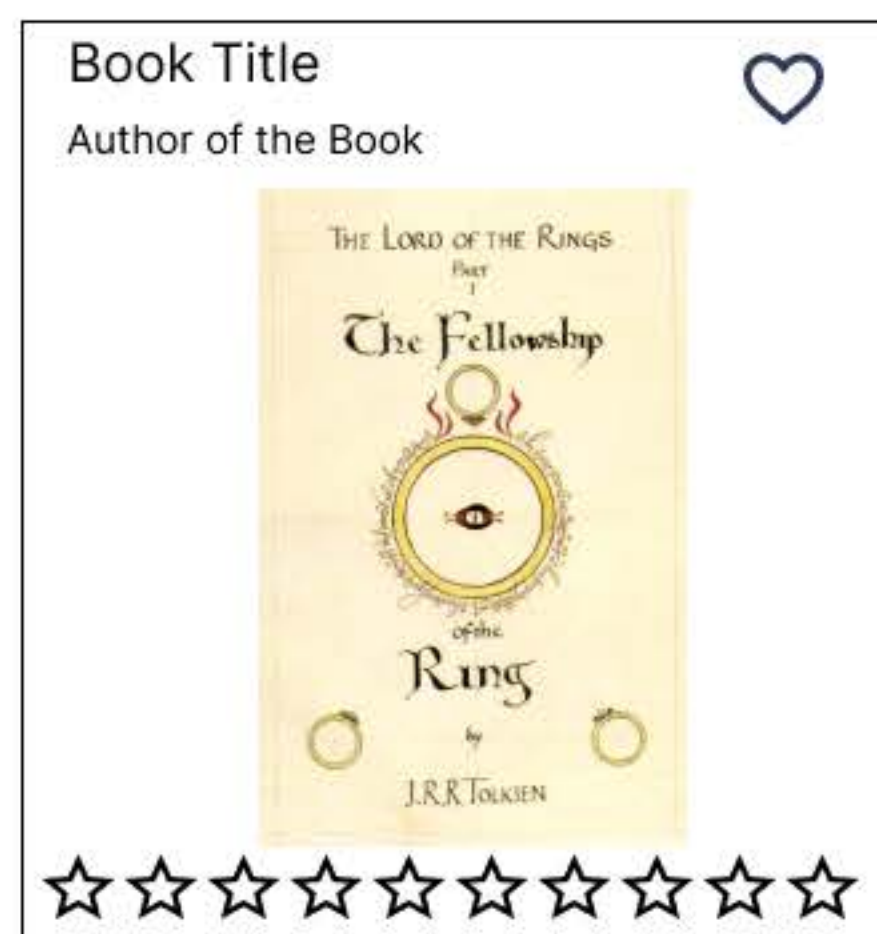
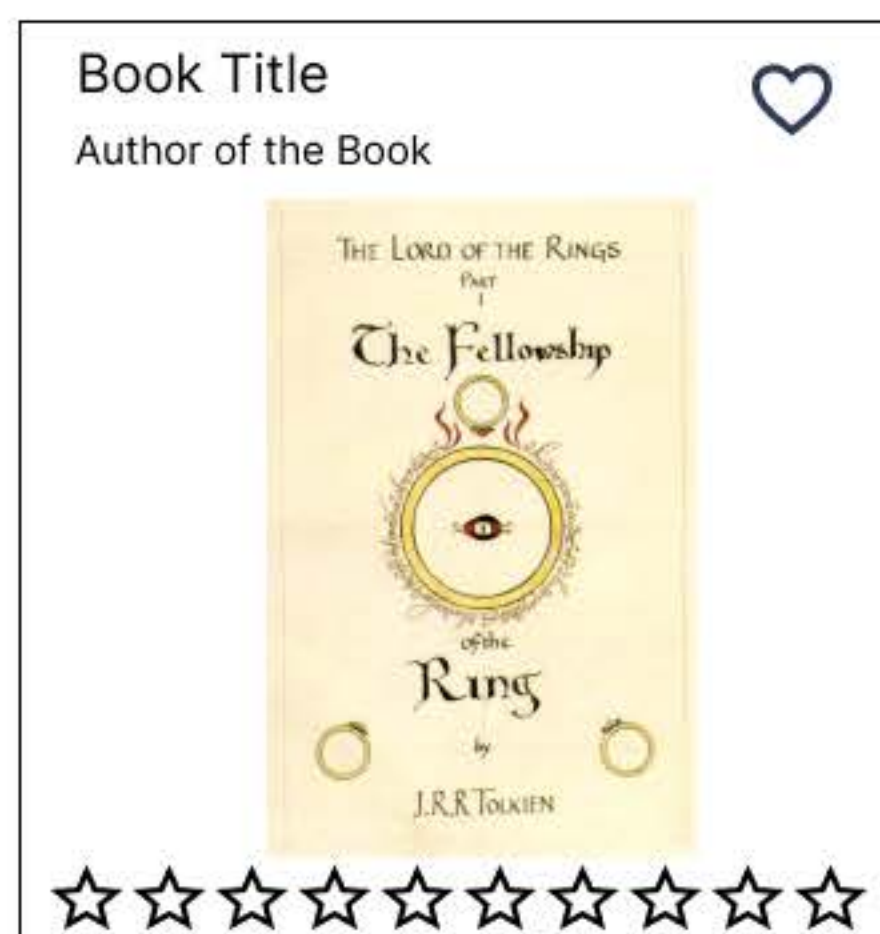
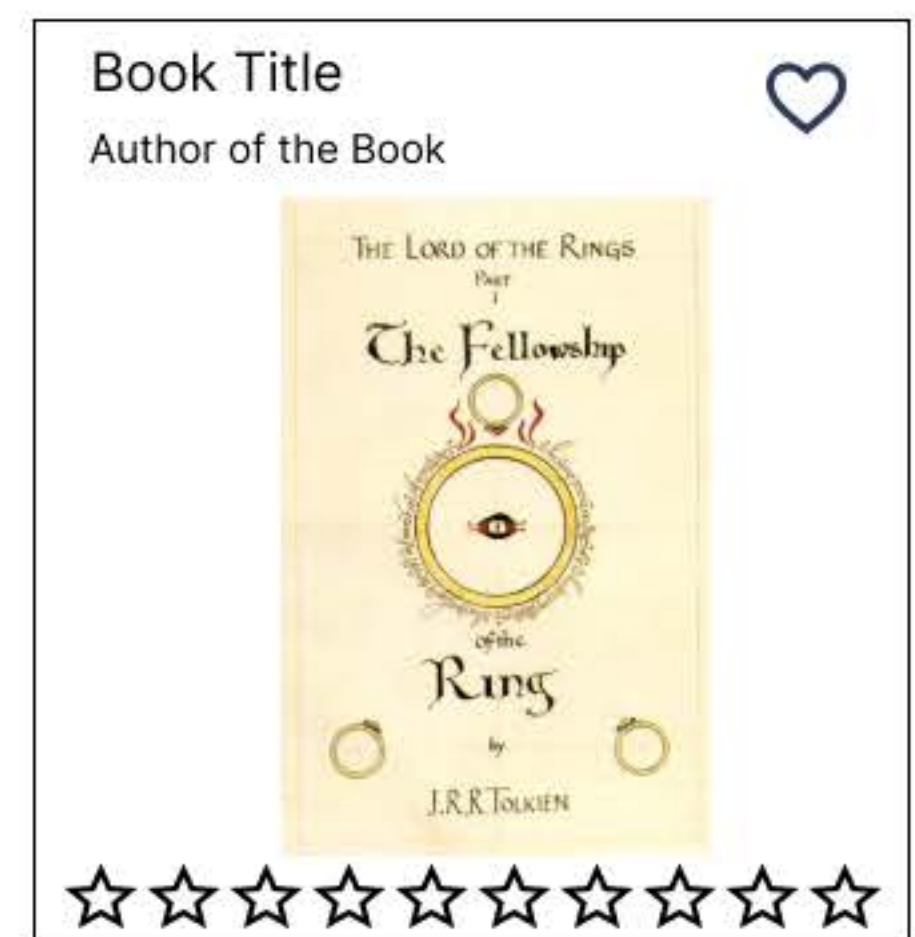
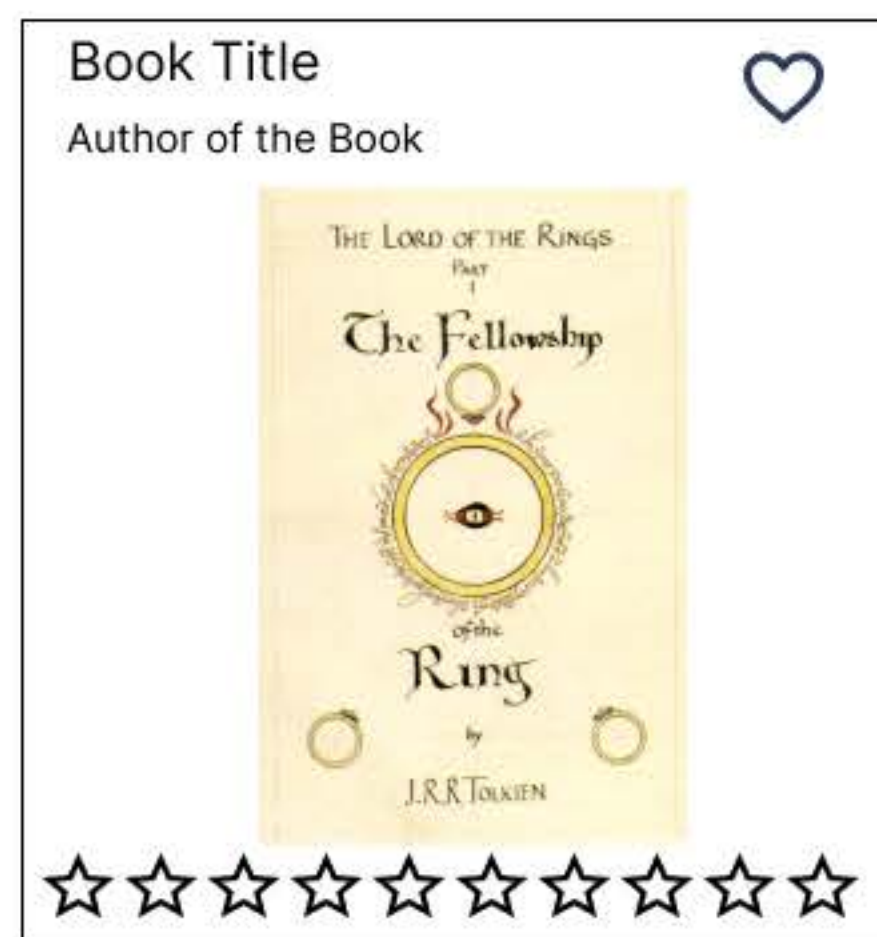
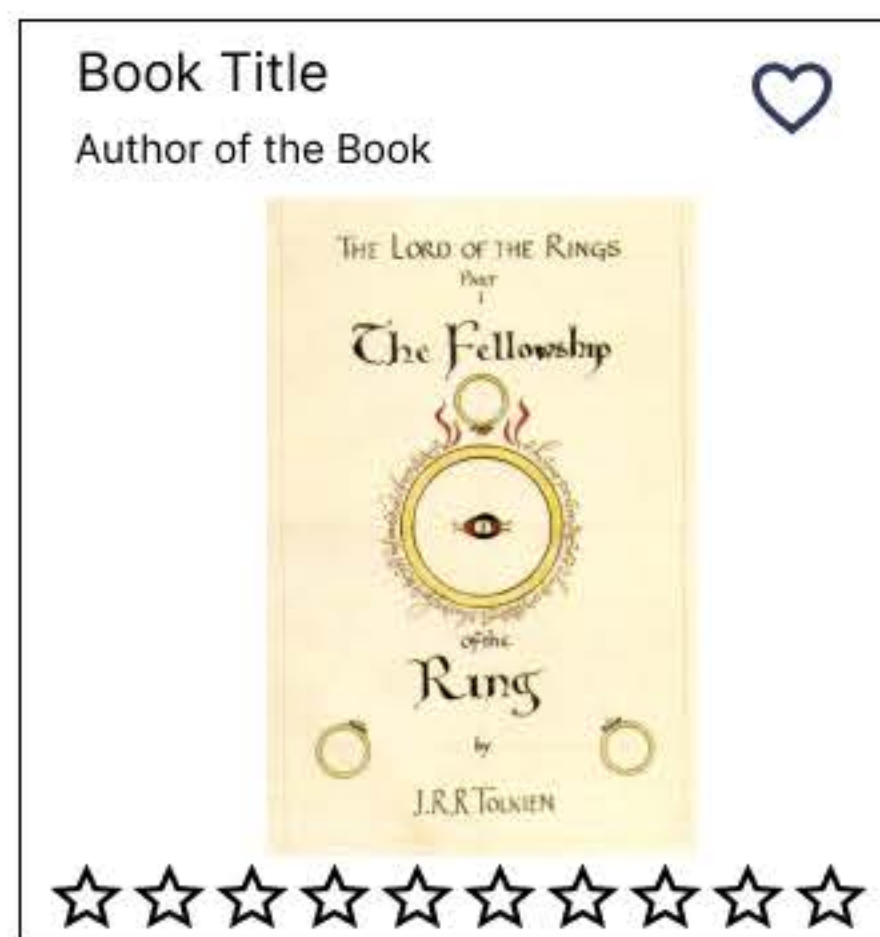




Recommended for you



Search results

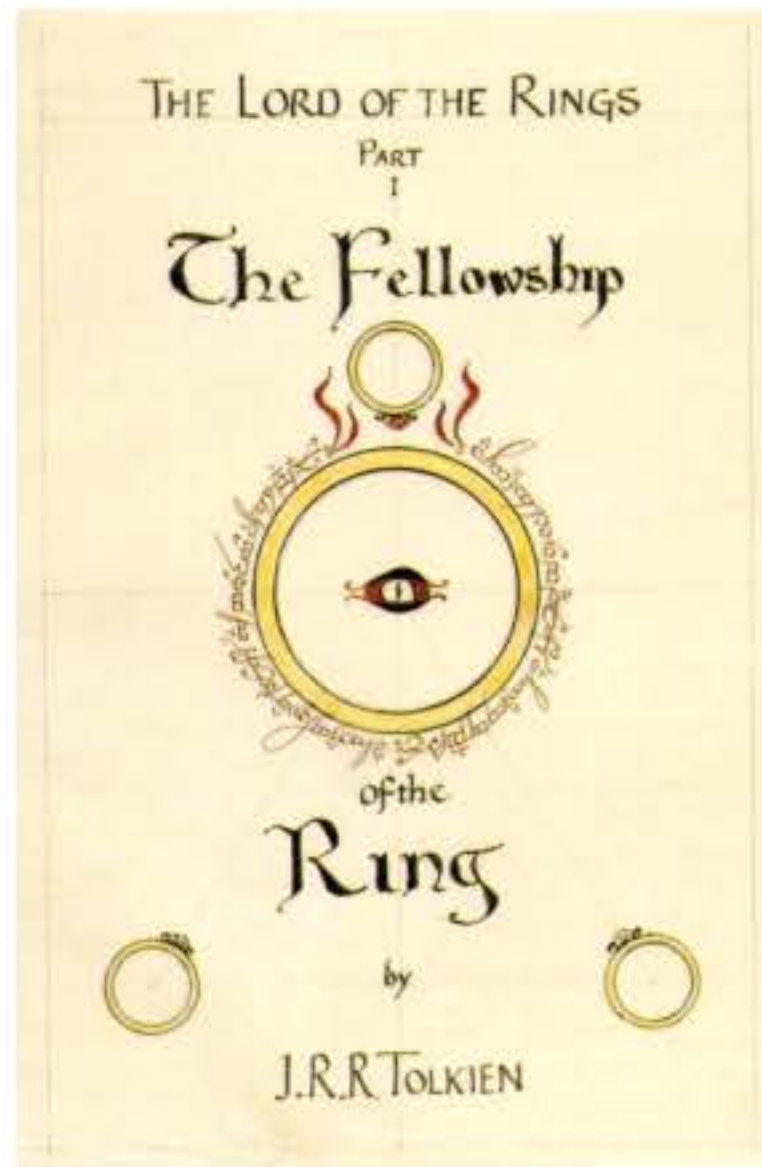




Herr der Ringe: Die Gefährten

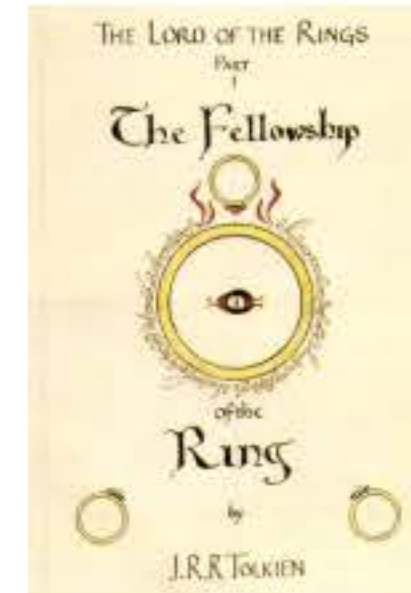
JRR Tolkien - 25.1.1975

☆☆☆☆☆☆☆☆☆☆



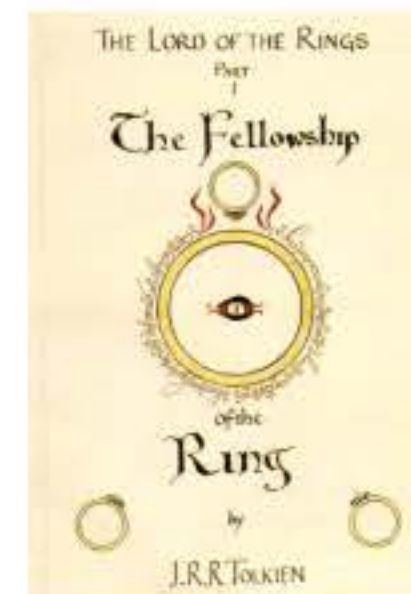
Favorit

Recommended for you



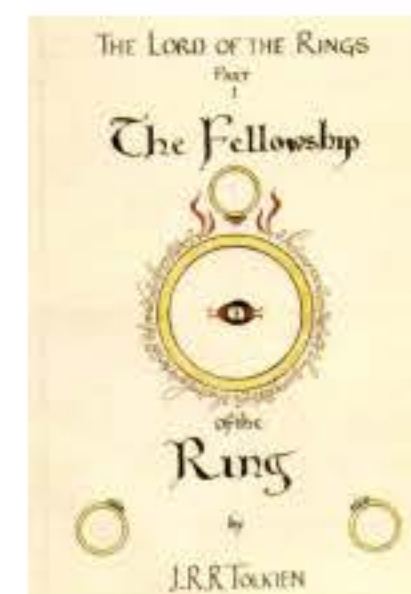
Herr der Ringe: Die Gefährten

JRR Tolkien - 25.1.1975



Herr der Ringe: Die Gefährten

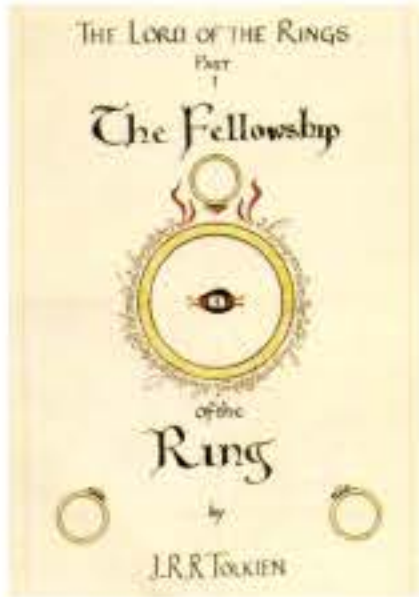
JRR Tolkien - 25.1.1975



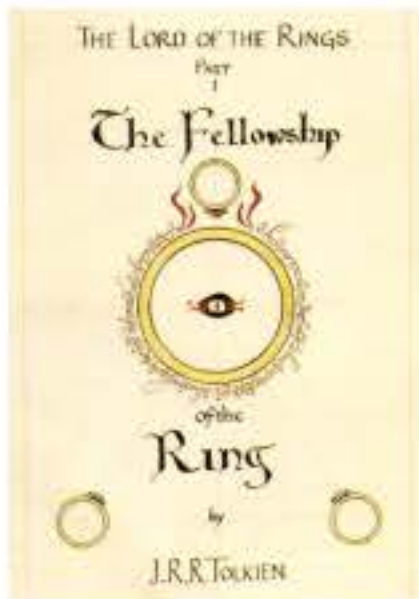
Herr der Ringe: Die Gefährten

JRR Tolkien - 25.1.1975

Your Favorites



Herr der Ringe: Die Gefährten
JRR Tolkien - 25.1.1975



Herr der Ringe: Die Gefährten
JRR Tolkien - 25.1.1975



Herr der Ringe: Die Gefährten
JRR Tolkien - 25.1.1975

