

# 3380

---

Towards the discovery of thematic  
attributes in books with semantic  
similarity clusters



**How many books is it possible  
to read in one lifetime?**



# Goodreads (why it matters)

goodreads Search books

Home My Books Browse Com



**Gone Girl**  
by Gillian Flynn (Goodreads Author)

★★★★★ 4.08 Rating details 2,348,654 ratings 133,370 reviews

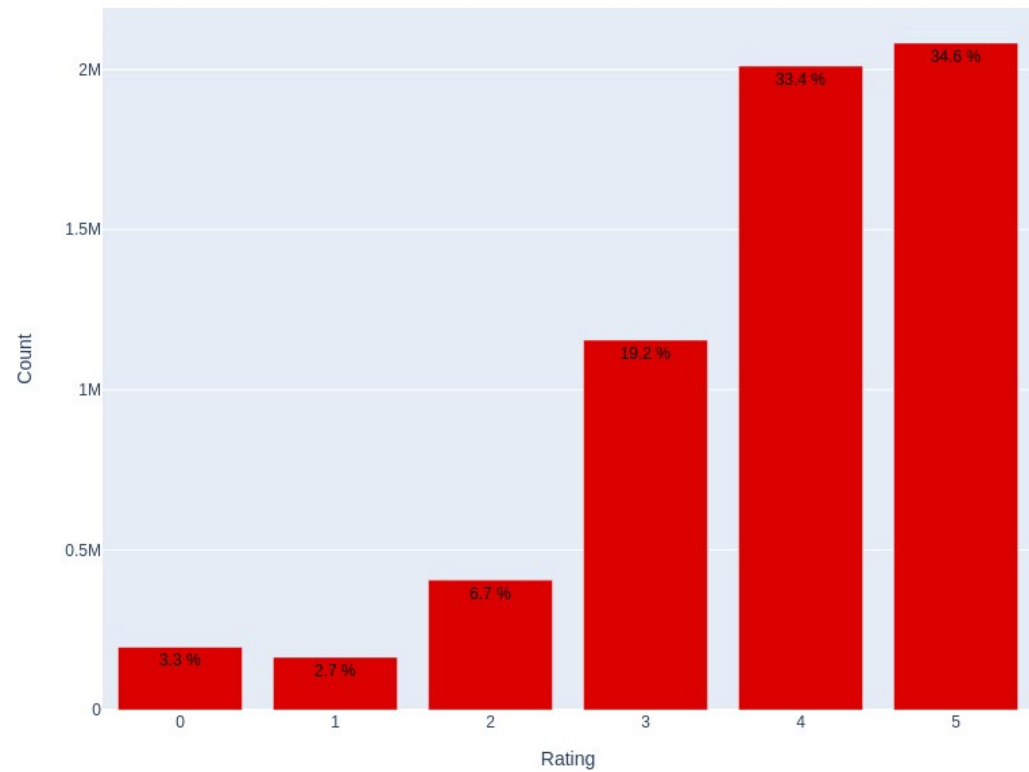
Marriage can be a real killer.

On a warm summer morning in North Carthage, Missouri, it is Nick and Amy Dunne's fifth wedding anniversary. Presents are being wrapped and reservations are being made when Nick's clever and beautiful wife disappears from their rented McMansion on the Mississippi River. Husband-of-the-Year Nick isn't doing himself any favors with cringe-worthy ...more

goodreads choice 2012 WINNER

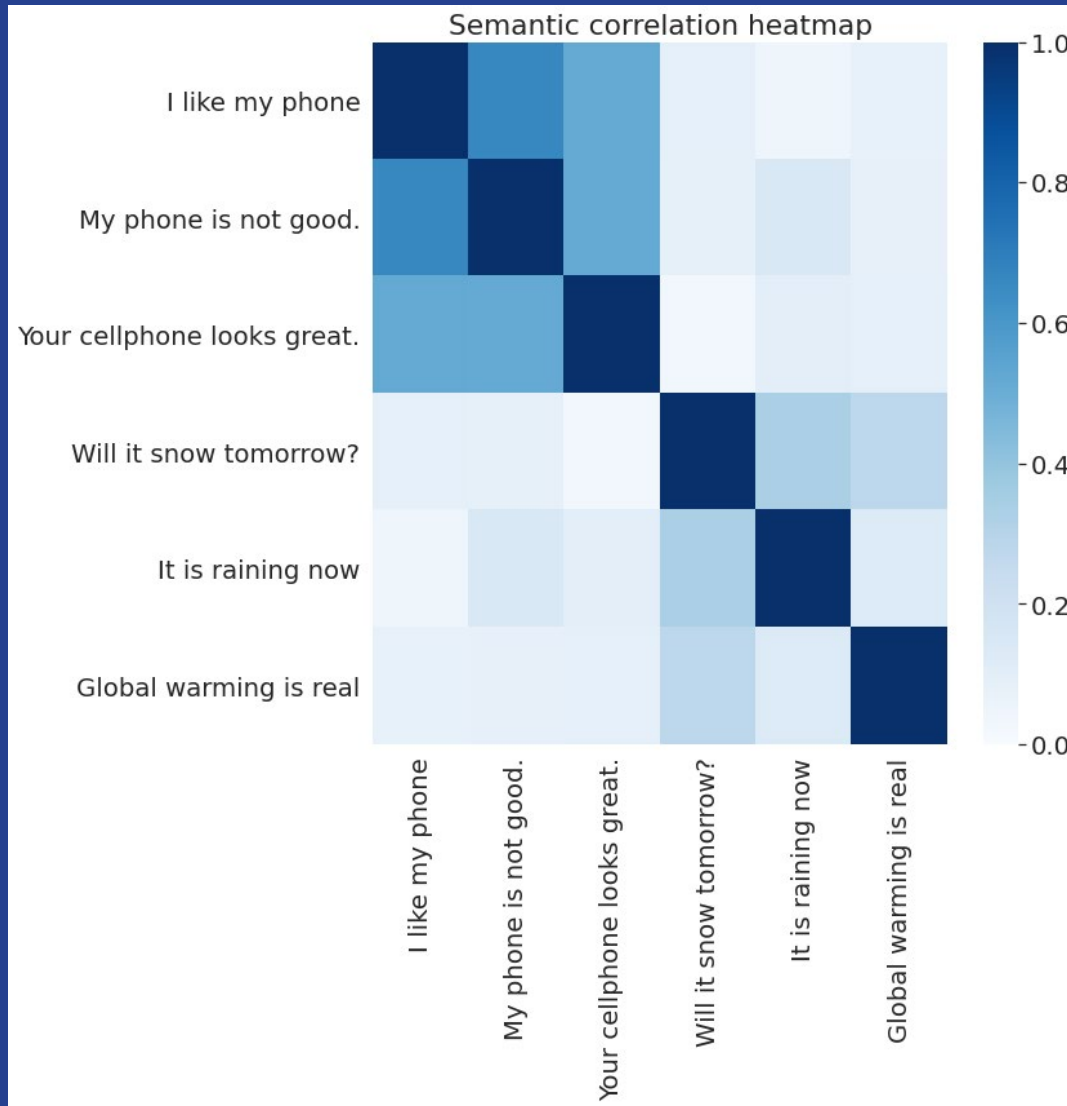
Want to Read Rate this book

Distribution Of 6015347 Goodreads Book Ratings





# What is Semantic Textual Similarity?

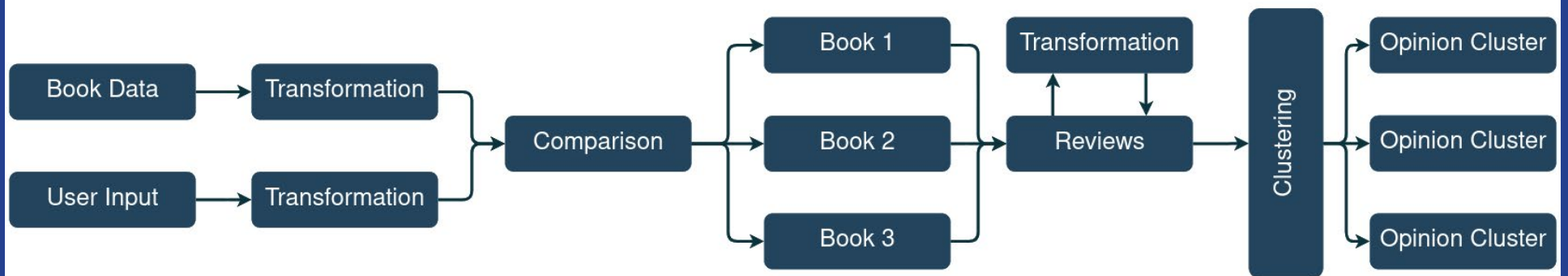


Semantic similarity using  
Tensorflow's Universal Sentence  
Encoder



# Project Pipeline

## 3380 Books





## Demo



# On-Demand Clusters

Only create and save clusters for books at delivery time

# Qualitative Assessment

Human-centric approaches

# Computer Limitations

Built new computer to assist with computation and program development

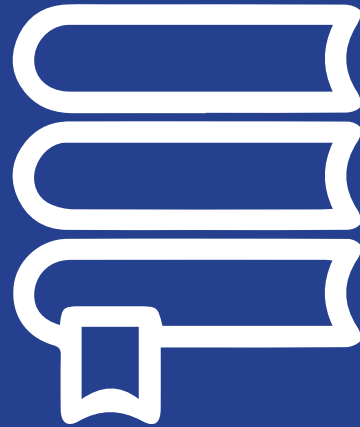


**Thematic recommendations  
based on review data**

**Language detection**

**Larger corpus**





# Thanks!

---

[adravaldes@gmail.com](mailto:adravaldes@gmail.com)  
[github.com/Einhard6176](https://github.com/Einhard6176)  
[linkedin.com/in/adnanvaldes/](https://linkedin.com/in/adnanvaldes/)