

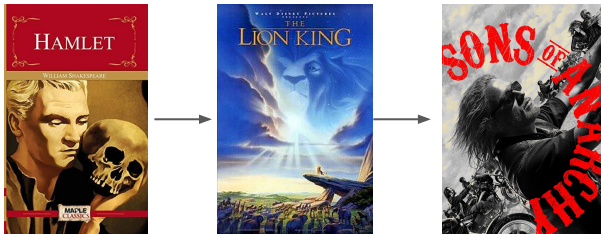
Invisible Adaptations

Project Lead: Grant Glass

Undergraduates: Mikaela Johnson, Nikhil Kaul, Elise Xia

INTRODUCTION

Much of literature is reminiscent of classic storylines. However, William Shakespeare, Mark Twain, Jane Austen, ancient mythology, and common folklore are **seldom acknowledged** as sources of inspiration in modern literature. One of the reasons for this is that **story appropriation**, rather than direct adaptation, is commonplace. By nature, literary analysis is extremely nuanced, and the bounds of adaptation are ambiguous; it is thus difficult and labor-intensive to identify and trace appropriated text.



GOAL

The goal of this project is to introduce a **methodology to identify these appropriated texts** by using Shakespeare's *Hamlet* to look for common signs of adaptation (famous quotes, character names, and settings) in primary sources, as well as themes and public interpretation found in secondary sources.

Primary Texts:

- Look for quotes/settings/characters on a digital library source, HathiTrust

Secondary Texts:

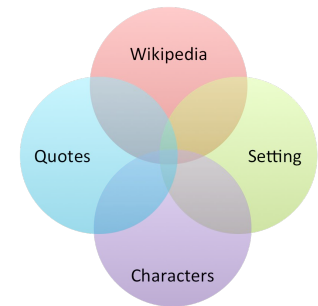
- Break down key topics of novel and use Wikipedia to compare to other literature

*Ultimately, **combine these** to create a threshold for what is an appropriation.

SOURCES



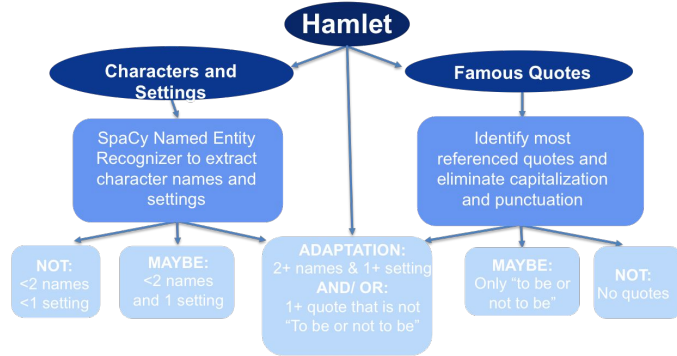
IDEA



The greater the overlap, the greater our confidence it is an adaptation.

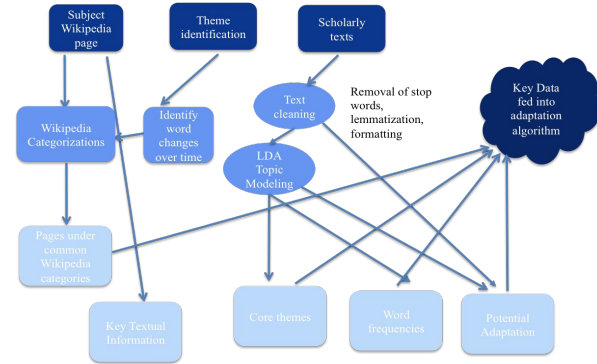
Methodology

Primary Sources



- **SpaCy Named Entity Recognizer (NER)** uses the Natural Language Toolkit to tag names and geographical locations—we then use these names and locations to identify characters and settings.
 - ◆ 2 matching character names and 1 matching setting is the cutoff for adaptation.
- **Quote Matching** was achieved by using the original text’s most-referenced quotes found the JSTOR digital library.
- Every quote is stripped of capitalization and non-letter characters in order to avoid false negatives due to syntax difference.
 - ◆ 1+ matching quotes (2+ if one of the quotes is “to be or not to be”) is the cutoff for adaptation.

Secondary Sources



- We used Latent Dirichlet Allocation (**LDA**), a generative model that assumes **all documents are a mixture of topics**, with each topic then **represented as a distribution over words**.
- By collecting scholarly articles and secondary sources from Wikipedia, we are able to produce key themes and ideas about the subject text.
 - ◆ These are used to analyze potential adaptations using “theme” matching. If an adaptation contains **at least two common themes** or belongs to **at least one common Wikipedia category**, it is a potential adaptation. Combining these elements provides us with levels of confidence in defining a work an “invisible adaptation”

Takeaways



This image displays the **volume ids** that HathiTrust generated for our **quote matching** algorithm

CONCLUSIONS

- Our literature review allowed us to fully understand the **arbitrariness** and **debatable** nature of the Digital Humanities
- Our **Wikipedia API** and **topic modeling** was able to **return titles of works of literature and film** which fit under each category listed.
 - ◆ The more categories a work fits under, the more likely it is an adaptation.
- **HathiTrust research** returns **volume ids** of works corresponding to title names in different levels of adaptations (closer it is, the more similar it is).

FUTURE WORK

- Combine our primary and secondary methodology to create one cohesive algorithm
- Establish a confidence interval
- Transfer Volume ID's to Titles of Works

This is a screenshot of the **themes** we analyzed as important to *Hamlet* and then, based on research for how those concepts have **changed over time**, the corresponding **Wikipedia categories** we found for possible adaptations to fall under

Theme Term	Wikipedia Categories
Power	Political Art, Political Fiction, Political Novels, Political Plays, Political Literature, Monarchies, Works Set in Castles
Familicide	Familicide in Fiction, Fratricide in Fiction, Regicide, Fictional Fratricide, Fictional Regicide, Matricide
Betrayal	Deception, Obfuscation, Works about Divorce
Incest	Adultery in Plays, Incest in Films, Adultery in Fiction, Sexuality in Plays, Adultery in Plays, Incest in Plays
Ghost	Ghost in Written Fiction, Ghost novels, Fictional Ghosts, Ghosts in Film, Ghosts in Popular Culture, Fictional Spiritual Mediums, Ghost Stories, Ghosts in TV
Revenge	Revenge in Fiction, Films about Revenge, Revenge in TV, Video Games about Revenge, Novels About Revenge
Romance	Romantic fiction, Romantic novels, Romantic novels by subgenre, Paranormal romance novels, Romance Films, Romance TV
Madness	Mental illness in fiction, Fictional characters with neurological or psychological disorders, Mental illness in film, Mental disorders in film, Diseases and disorders in film
Inability to Act	Time management, Anxiety, Human Behavior, PTSD in Fiction, Fictional characters with PTSD, Grief
Acting/Performance	Metafiction, Metafictional Works, Metafictional Plays, Metafictional TV, Metafictional Films, Metafictional Techniques
Death	Novels About Death, Death and TV, Death in Fiction, Cemeteries in Fiction, Fiction About Murder, Fictional Skeletons, Suicide in Fiction, Fictional Suicides
Tragedy	Comitragedy, Tragic Hero, Tragic Plays, 9+ Bourgeois Tragedy