

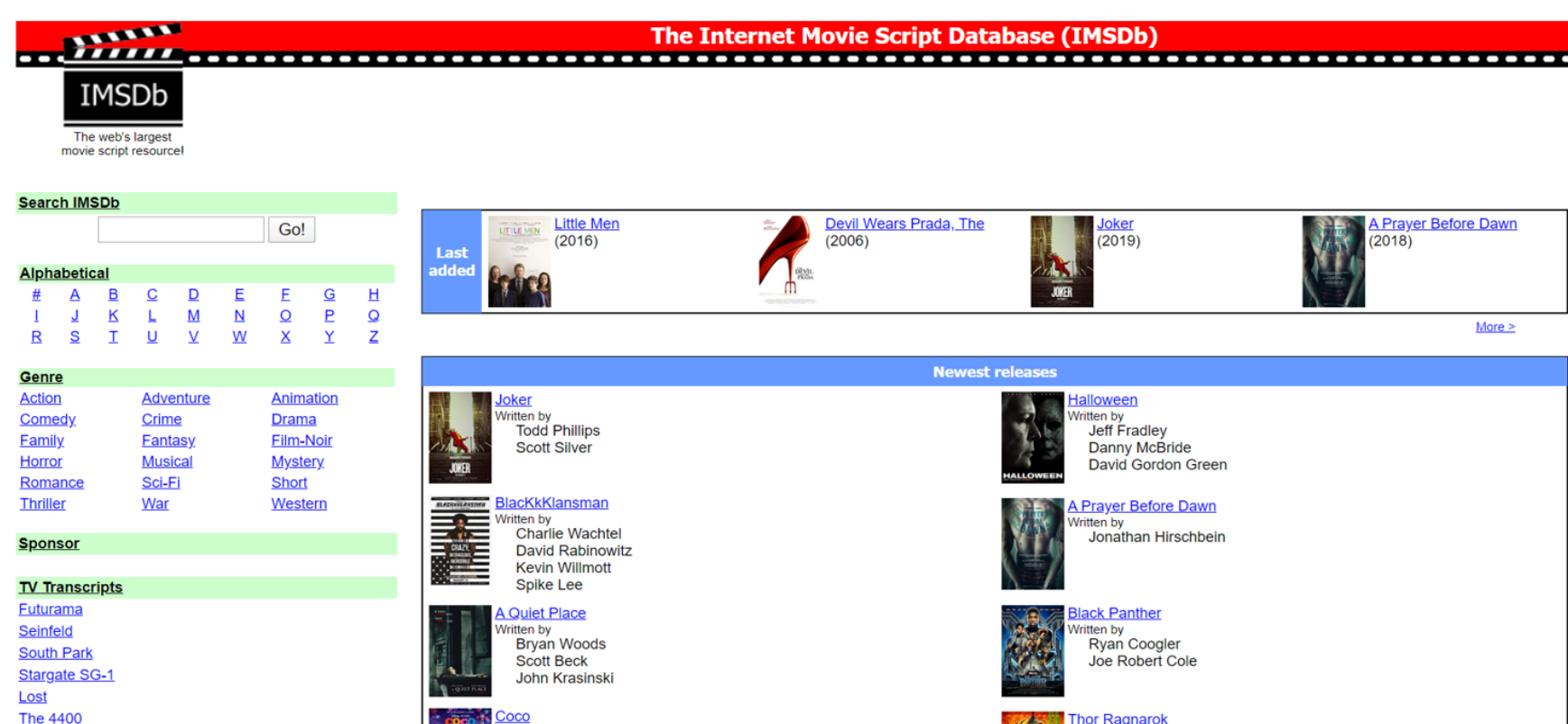
劇本裡的千層套路——從情感分析談電影

旁聽生 曾理陽



劇本蒐集

來源：The Internet Movie Script Database (IMSDb)



文本處理

- 使用strsplit()
- 正規表達式

分析

- 使用Tidyttext
- 句子分解為單詞
- 利用“nrc”情緒字典集進行情感分析

視覺化

- 利用ggplot2做出直方圖
- 利用gganimate做出GIF圖

Goals

你們好奇同類型電影存在某種淺規則嗎？

什麼氣氛下會有床戲、親吻，甚至什麼人物要領便當、鬼會什麼時候從轉角蹦出來嚇人。

或許電影劇情就藏在情緒中，讓我們做個情緒分析，用不暴雷的方式暴雷吧。

Methods

劇本蒐集 → 文本預處理 → 分析處理 → 視覺化、動態化

使用到的package：rvest, dplyr, ggplot2, gganimate, stringr, tidyttext

本專案方法參考 Shuyi Wang: [如何用Python和R對故事情節做情緒分析？](#)

Results

驚悚片的千層套路

- 何時前方高能？

影片進度的10%、30%、70%、90%處最有可能



戰爭片的千層套路

- 爆炸效果很刺激？看看導演們是怎麼燒錢的。

開頭就是大把大把鈔票在燃燒的味道



愛情片的千層套路

- 橫刀奪愛才是真愛？
- 真金不怕火煉，真愛不怕唇槍舌劍？
- 與世不爭使你憑實力單身？

愛情片設計多少衝突，讓人心裡小鹿亂撞？

