# Significance at the Movies

CLARA BOOTHBY

Cross-evaluation of metrics to estimate the significance of creative works (2015)

- **Max Wasserman**
  - Ph.D. student, Amaral Lab @ Northwestern
  - Applied math and Statistical analysis of sports

- **Xiao Han T. Zeng**
  - Ph.D. Student, Amaral Lab @ Northwestern
  - Studies scientific publication and citation data and gender disparities in STEM

- **Luis Amaral**
  - Professor, Amaral Lab @Northwestern
  - Complex systems, Biology
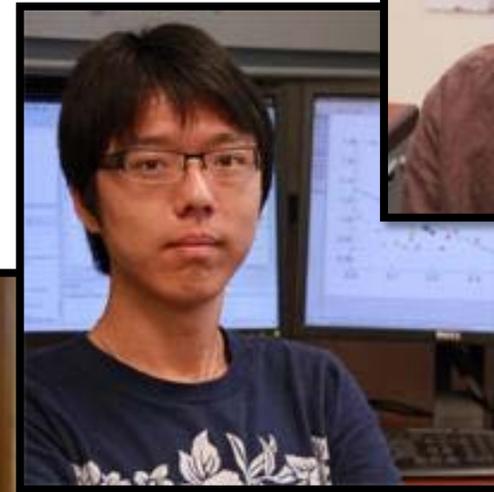
Why I selected this paper

- Politics of Canonization

- Machine Learning

Why I selected this paper

- Politics of Canonization

- Machine Learning

However!

Why I selected this paper

- Politics of Canonization

- Machine Learning

However!

- Focus on mostly on their premises

- Criteria for significance

Significance!

- "Lasting importance of a creative work"

- Latent Quality

Significance!

- "Lasting importance of a creative work"

- Latent Quality

- Induction into the National Film Registry

LIBRARY OF CONGRESS

Variables!
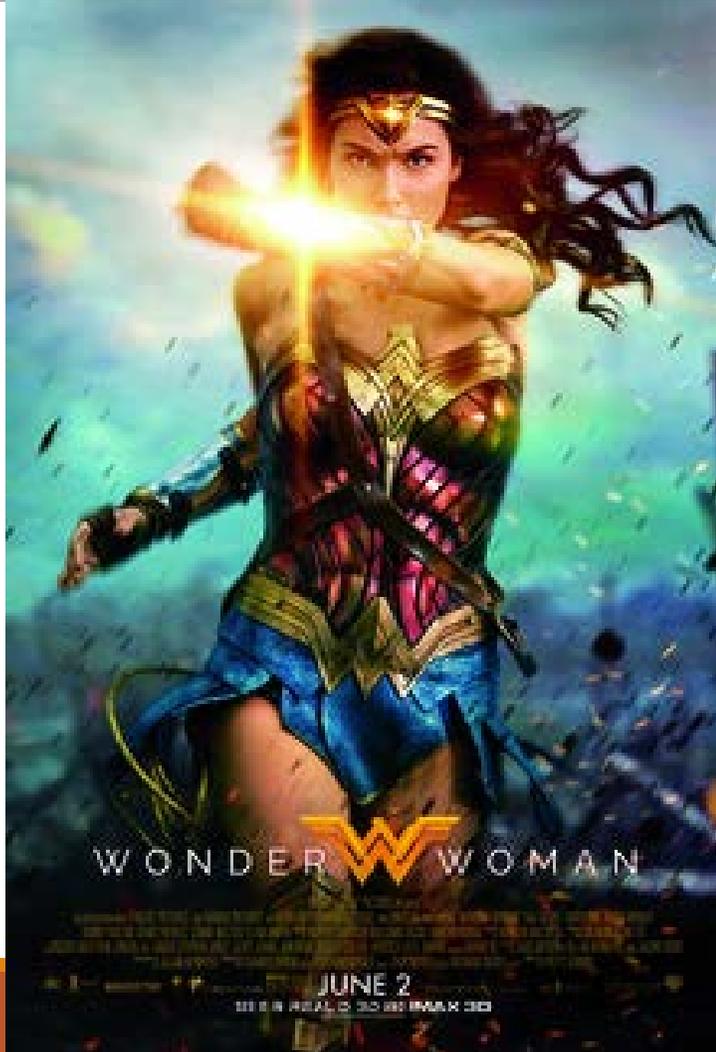
- Expert Reviews

- Amateur Reviews

- Automated Data

Variables!

- Expert Reviews

- Amateur Reviews

- Automated Data

- CITATION DATA!!

Methods

- Single Variable Analysis

- Multi-variable Analysis
  - Random-forest

To select features that could predict significance (i.e. inclusion in the NFR)

Example: Willy Wonka

June 27, 2014: Paper received

- Willy Wonka (1971) is not in the NFR

- BUT it has a high long-gap citation rate

Example: Willy Wonka

June 27, 2014: Paper received

- Willy Wonka (1971) is not in the NFR

- BUT it has a high long-gap citation rate

December 1, 2014: Paper approved

Example: Willy Wonka

June 27, 2014: Paper received

- Willy Wonka (1971) is not in the NFR

- BUT it has a high long-gap citation rate

December 1, 2014: Paper approved

December ?, 2014: Willy Wonka inducted into the NFR

The National Film Registry

- 700 films total; 25 new films each year; Films must be >10 years old to be eligible

- Public nominates films; Librarian of Congress and National Film Preservation Review Board make final selections

- 40 films were directed by women. No accessible data about racial demographics of the directors

Questions

- What new insights or information does this method provide to us that we could not have been learned otherwise?

- Considering what we know of the method, do you feel that this method truly measures significance? Why or why not?

- What responsibilities do scholars have when working on projects that intersect with another field?