Natalie Perez

Professor Daniel Anderson

ENGL480 Digital Humanities History and Methods

30 January 2020

Investigating Classic Horror Cinema via Data Mining, or the “Frankenstein Complex”

**Why would anyone investigate classic horror cinema and monsters?**

When Universal Pictures revealed that the *Universal Classic Monsters: Complete 30-Film Collection* would be released on Blu-Ray, Tianyu Jiang, a Master’s student enrolled within the Department of Media Studies at Stockholm University had a *bloodcurdling* epiphany. Being a horror cinephile, or “someone very fond of horror motion pictures,” Jiang was fascinated by the promise of behind-the-scenes documentaries and archival footage of some of the most notable horror films of the 1930s/1940s, ranging from *Dracula* (1931) to *The Wolf Man* (1941).These classic monsters from almost one hundred years ago are still so notable in the late 2010s and early 2020s, and Jiang started to question *why.* Jiang states that “the mysterious names of the monsters are the representatives of a cultural phenomenon associating with monster and horror, and the meanings of these names have exceeded beyond their cinematic figures” (5). She believes that these classic monstrous figures are notable within pop culture today, surpassing beyond their origins and introductory films … but, why were these classic horror films so notable in the 30s and 40s? Why is classic horror cinema so significant, and how was this branch of cinema impacted by history? Hmm…

In an effort to utilize both digitalization and humanities research practices, Jiang focuses on using digital archives and cinema studies/film history research to answer various questions. While investigating theoretical works on classic horror cinema and film history, Jiang proceeds to perform a meta-inquiry of data mining classic horror cinema data via the Media History Digital Library (MHDL) in order to investigate film spectatorship through the lens of sexual identity and the subsequent cultural significance of these films. Now, let’s explore the meaning and significance behind these practices, Jiang’s results, and what they mean for the future of the relationship between digital tools and the humanities.

**The “Frankenstein Complex” and how it defines digital humanities**

Jiang references Isaac Asimov’s term, the “Frankenstein Complex,” defined as the human fear of “mechanical men,” or enhanced technology made at the hands of humans. She relates this concept to digital humanities, denoting “a tension between the humanistic research tradition and the emerging digital practices in the field” (12). There is a sense of a dichotomy between digitalization and the humanities, but Jiang’s investigation aims to create a bridge between the two disciplines, demonstrating the relationship they can have within the digital humanities.

This investigation combines theoretical research on classic horror cinema and film history with digital tools to analyze data on classic horror cinema within the archives available on the Media History Digital Library (MHDL). Through this project, Jiang aims to investigate film spectatorship of classic horror cinema, hoping to expand the existing model of queer spectatorship while investigating film theorists’ works regarding the relationship between monsters and the marginalized.

**What is the theorized relationship between classic horror cinema and queer spectatorship?**

Upon discussing early gay literature, James Jenkins of *Valancourt Books* states that LGBT characters and themes within horror fiction can be traced back to the late 1700s (Healey). According to Jenkins, “the traditional explanation for the gay/horror connection is that it was impossible for [LGBT authors] to write openly about gay themes back then (or even perhaps express them, since words like 'gay' and 'homosexual' didn't exist), so they sublimated them and expressed them in more acceptable forms, using the medium of a transgressive genre like horror fiction,” resulting in various queer-coded characters and monsters within the horror genre (Healey).

How does this impact classic horror cinema from the 1930s/1940s? Well, in 1943, the “Hays Code” was implemented, which included a strict set of rules regarding what was appropriate for filmmakers to show on screen. According to Inés Mendoza-Pérez of *Control Forever*, despite LGBT issues not being listed or denounced explicitly, there was a code calling for “correct standards of life” and “the sanctity of the institution of marriage,” which implicitly stated that any explicit LGBT characters and themes would not be allowed within American film (“Queer-Coding and Horror Films”). Despite not being allowed onscreen, LGBT characters and themes were ever present in classic horror cinema through implicit themes and subtext. In “Queerness in the Horror Genre, Offensive or Progressive?”, Hannah Stevens points out that queer characters were “forced to the background of film, relocated to roles that consisted of mostly villains … in order to skirt under the rules, [filmmakers] began creating queer characters who were monsters, and thus fictional enough to get away with breaking the rules.”

Despite classic horror cinema having queer-coded characters and LGBT themes, it helped to popularize harmful stereotypes involving sexual perversion and queerness (vampires seducing the same sex for blood, werewolves turning into violent animals at night, etc.) that impacted the way LGBT people were viewed, and ultimately, these stereotypes still impact LGBT representation in media today. Whether it was hyper-femininity or hyper-masculinity, these stereotypes indirectly harmed the LGBT people the films were attempting to include within cinema.

Jiang aims to investigate the relationship between classic horror cinema and queer spectatorship by analyzing data and trends involving monsters and queer people in cinema, before and after the implementation of the Hays Code, within digitalized journals via the Media History Digital Library (MHDL).

**What is the Media History Digital Library?**

Before diving into Jiang’s investigation, it’s important to look into the history of the Media History Digital Library (MHDL) and understand its digital capabilities for seeking information and extracting cinematic data.

*Origins of the Media History Digital Library (MDHL)—*

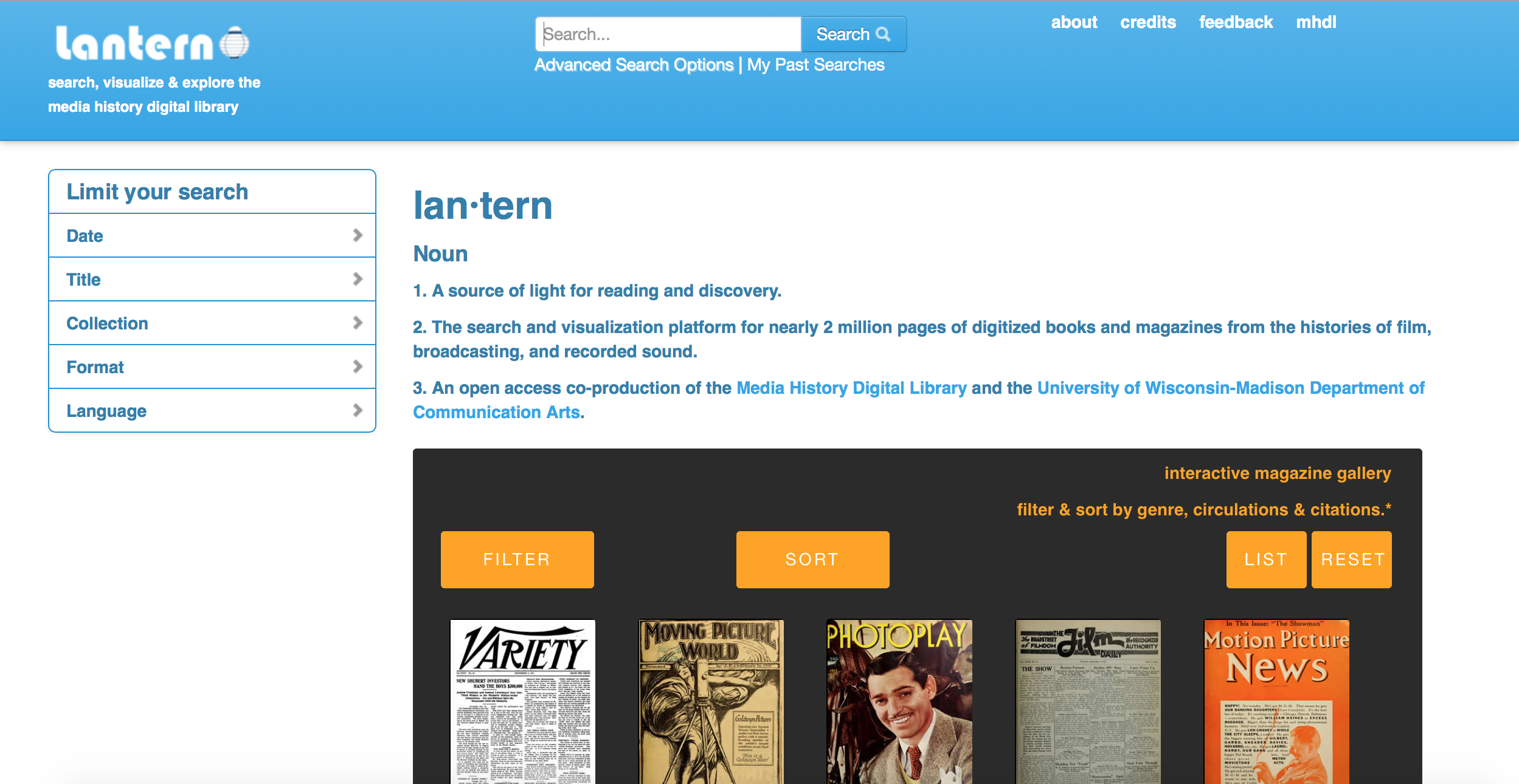
David Pierce, a film archivist and historian, founded the MHDL in 2009, in an effort to improve the way in which historians could find materials related to their research by digitizing classic film and providing public access to media-related journals, trade papers, and fan magazines (Jiang 17). Joining the project in 2011, Eric Hoyt, a digital humanities scholar at the University of Wisconsin-Madison, began building the MHDL website (which, you can find right here: <http://www.mediahistoryproject.org>) and digitalizing several materials. Throughout this process, Hoyt and his team launched a search engine called “Lantern” for the website in 2013, and then later, added “Project Arclight” in 2015, in an effort to, according to Jiang, “make the MHDL’s large-scale collection more user-friendly and analytically” (17).

**

Screenshot of MHDL’s main webpage. *Media History Digital Library*, Media History Digital Library, www.mediahistoryproject.org/.

*Lantern—*

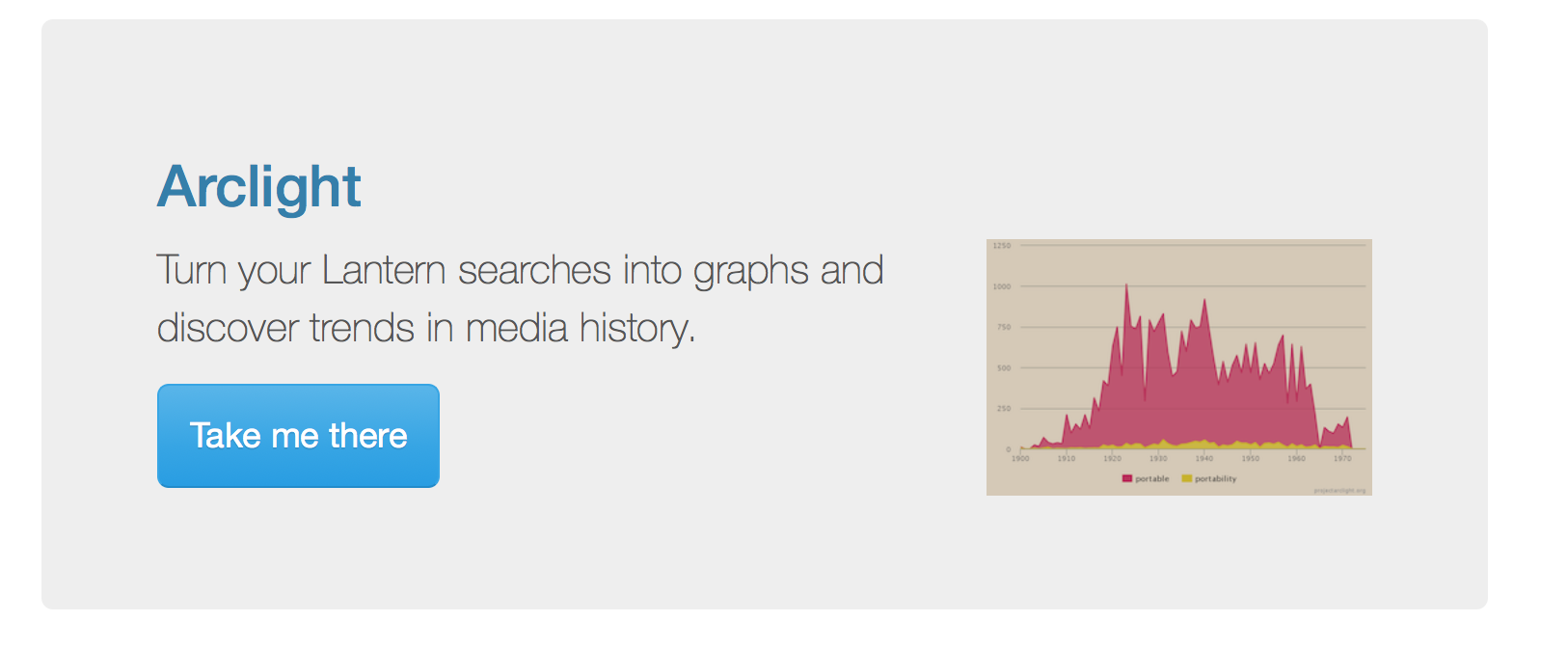
As stated on the search engine’s webpage, Lantern allows for “the search and visualization of nearly two million pages of digitalized books and magazines from the histories of film, broadcasting, and recorded sound,” which creates a great tool for research whether one is a historian or undergraduate student (“Media History Digital Library”). In his essay, “Lenses for Lantern: Digital Mining, Visualization, and Excavating Film History’s Neglected Sources,” Hoyt points out his main objective is for Lantern to “create a user experience that offers fast, powerful search and a sense of context about the underlying sources” (147). Emphasis is placed upon the Lantern’s ability to perform a full-text search quickly, especially in high volume traffic, so Hoyt built Lantern using open-source materials of Solr, a search platform that provides access to other notable archives, such as the Internet Archive, The Usenet Archive, and the Homeland Security Digital Library (HSDL) (Jiang 18).



Screenshot of Lantern’s webpage. *Media History Digital Library*, Media History Digital Library, www.mediahistoryproject.org/.

*Project Arclight—*

Amazingly, Project Arclight allows users to create graphs based on their Lantern searches in order to analyze trends in media history, which ultimately will provide Jiang with a digital means to investigate the relationship between classic horror cinema and historical events involving cinema (“Media History Digital Library”).

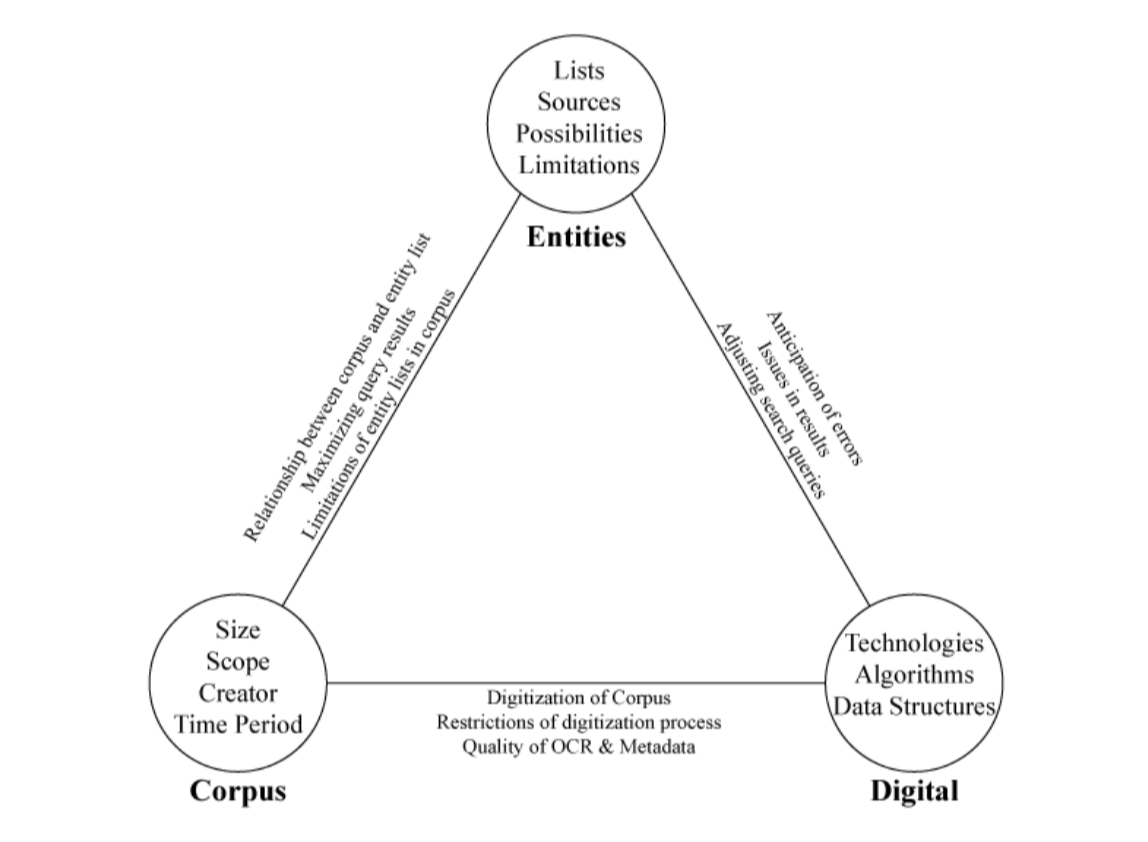
**Screenshot of embedded hyperlink to Project Arclight. *Media History Digital Library*, Media History Digital Library, www.mediahistoryproject.org/*.*

**What methods are utilized to navigate the Media History Digital Library?**

In order to perform a meta-inquiry, or a search of information regarding data and metadata, Jiang aims to investigate two methods to mobilize the MHDL’s collection, and she describes both of them: “[an] analytical and interpretive method named Scaled Entity Search (SES), and [a] pedagogical method named Prediction-Observation-Explanation (POE) strategy” (21). Sounds complex, huh? Let’s try to understand these *scary* methods!

*Scaled Entity Search (SES)—*

Data discovery methods should help with modeling a topic or visualizing a network, and a keyword search in a search engine could be necessary, but how does one interpret and pick from the results, especially if it is a large amount of data? Here is where Scaled Entity Search (SES) *lurks* in! Presented at the 2014 IEEE International Conference on Big Data, SES is described as an “interpretive framework,” working to be a “humanistic method” of big data analysis (Jiang 22). This may sound contradictory, but essentially, SES intends to use someone’s interpretations to find related data, which is useful for one doing research or investigating a question, and it makes sense.

When thinking about a question, it’s important to utilize SES’s “triangulated”interpretive framework involving the entities, the corpus, the digital, and their overlapping relationships. According to Kit Hughes, who has written about SES, one must “balance critical understandings of the chosen entities and corpus with knowledge of how digital technologies shape and frame results” (“Introducing Scaled Entity Search (SES): Interpretive Framework”). This method encourages the user to think about what the results say (the entities), what factors impact the variety of the results (the corpus), and what digital tools are being utilized to display said results (the digital). Utilizing this method when searching through computational data allows the user to display a “humanities”-based approach, which can ultimately encourage better use of the data. 

Jiang, Tianyu. The SES triangle method of interpretation. “Frankenstein Complex” in the Realm of Digital Humanities: Data Mining Classic Horror Cinema via Media History Digital Library (MHDL). 2019, http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-169638.

*Prediction-Observation-Explanation (POE) Strategy—*

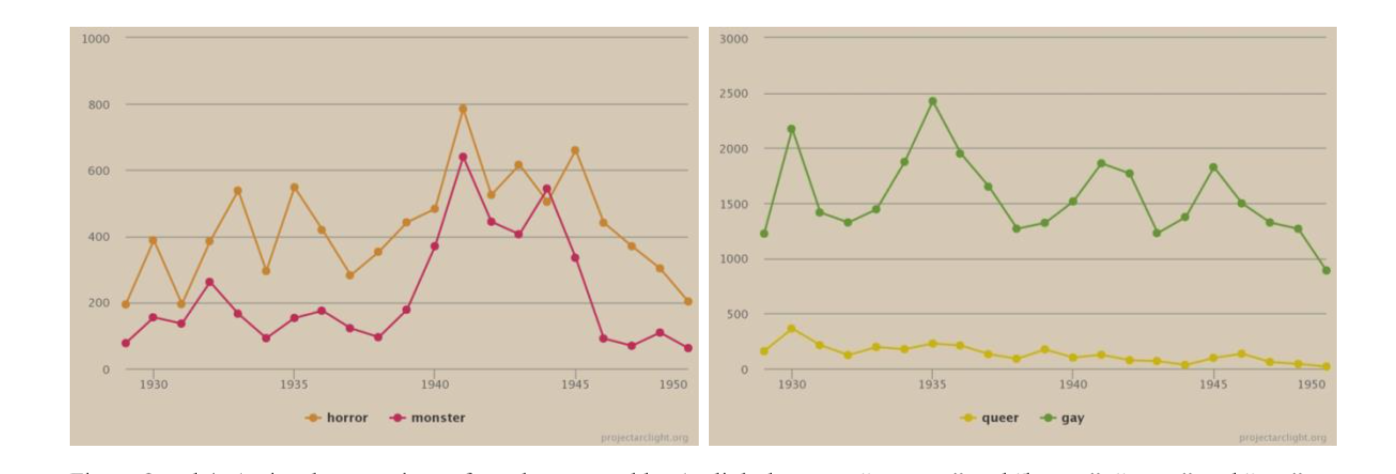
When Hoyt launched Project Arclight in 2015, he introduced the POE strategy into humanities research. Jiang describes it as a strategy that “requires the students to make predictions, observe an experiment, evaluate the outcomes [by] comparing [them] with their initial predictions, and explore the reasons and causes behind the experiment” (26). It can be difficult to use this strategy when dealing with some factual data, such as historical records, but it’s useful for trying to understand trends in data.

**The efforts of data mining classic horror cinema**

Jiang initially attempts to depict a horror movie-going scene during the 1930s/1940s by utilizing the digitalized materials on Media History Digital Library (MHDL) with the methods of Scaled Entity Search (SES) and distant reading (39).

*Framing the search entities—*

Jiang describes the process of determining her searches for Lantern: “I asked myself ‘why do *these people* like *this horror* in *this place* at *this particular time*?’ and accordingly designed the search entities draft into four categories: *these people*–names from the Universal Studio productions and the film column writers; *this horror–*including the title *Frankenstein*, its star Boris Karloff, and the director James Whale; *this place*–names of movie theatres selected from earlier readings including Rialto, Majestic Texas and RKO Kansas City; *this particular time*–the search narrowed down to a rough period” (40). By utilizing SES, Jiang specifies the information most crucial to the investigation. In addition, she also decides on the keyword “monster” instead of “horror,” and “queer” instead of “gay,” to improve the precision and readability of the visual patterns supplied by Arclight (41).

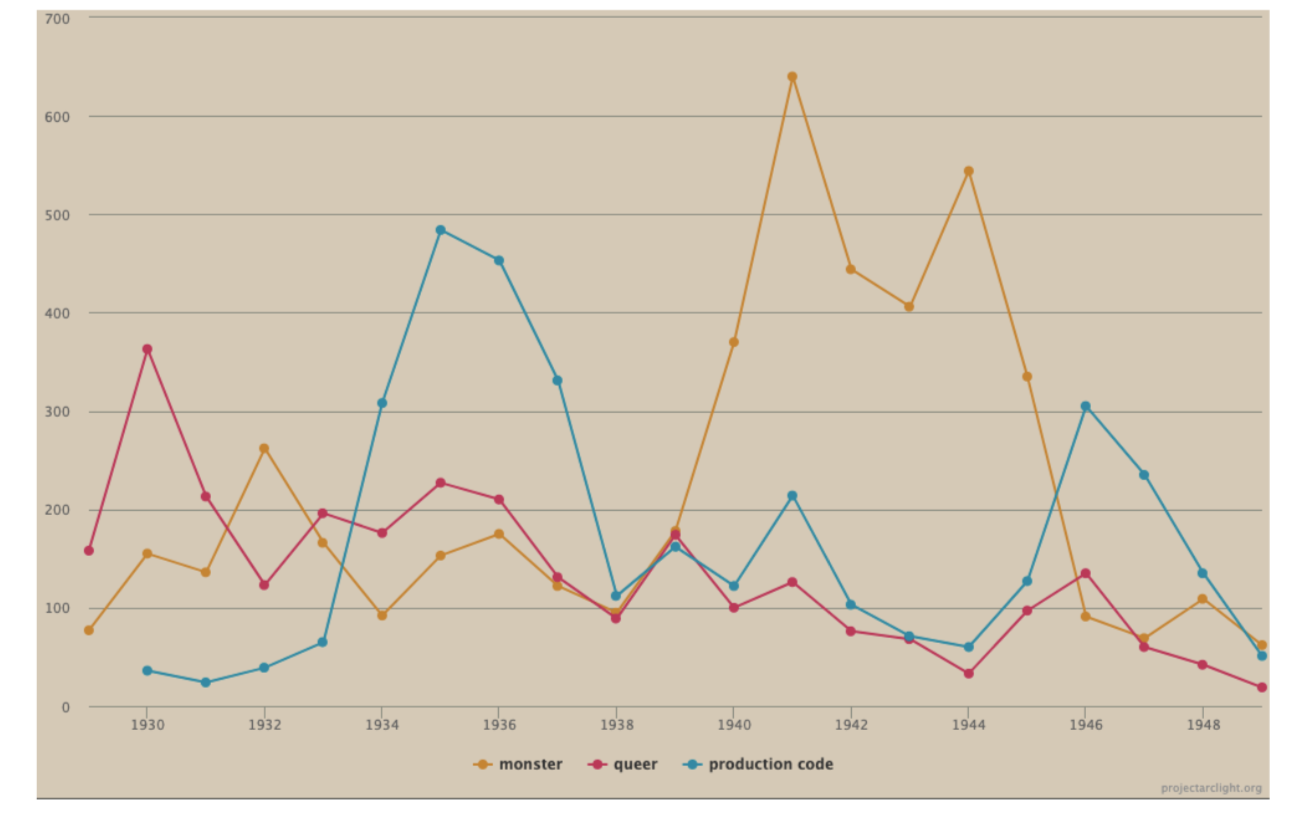


Jiang, Tianyu. A visual comparison of results returned by Arclight between “monster” and “horror”, “queer” and “gay”. “Frankenstein Complex” in the Realm of Digital Humanities: Data Mining Classic Horror Cinema via Media History Digital Library (MHDL). 2019, http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-169638.

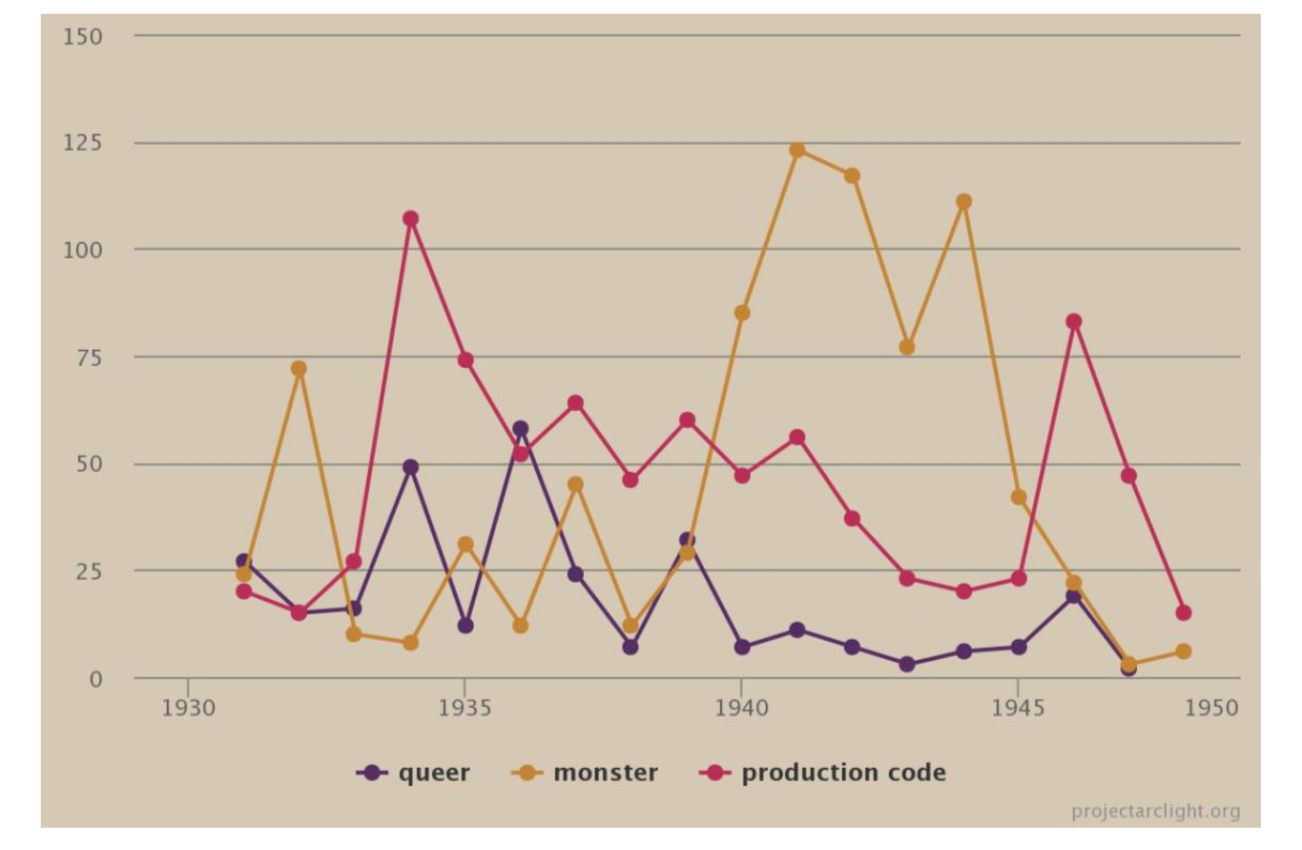
*Distant reading—*

By activating Project Arclight, Jiang’s Lantern searches of “monster,” “queer,” and “production code” are visualized, and she aims to use distant reading, “an attempt at utilizing big data analytics for the purposes of literary scholarship,” to investigate her questions (45).

Jiang’s first search utilizes MHDL’s entire collection of digitalized journals, and she later decides to narrow down the results during the second search by filtering for journals only from the *Motion Picture Herald* (46-7). In her second search, she states that there seems to be more correlation between the nodes involving “monster,” “queer,” and “production code,” once the data was narrowed down (47). However, in order to attain more accuracy, she states that the patterns should be “cross-examined by filtering the journals, adding complementary entities, and a closer survey of particular nodes that might lead to useful primary materials if a further explanation of the observed pattern is in question” (49).



Jiang, Tianyu. Search results of the designed entities in the entire MHDL’s digitalized journals. “Frankenstein Complex” in the Realm of Digital Humanities: Data Mining Classic Horror Cinema via Media History Digital Library (MHDL). 2019, http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-169638.



Jiang, Tianyu. Search results of the designed entities in *Motion Picture Herald*. “Frankenstein Complex” in the Realm of Digital Humanities: Data Mining Classic Horror Cinema via Media History Digital Library (MHDL). 2019, http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-169638.

**The future of the “Frankenstein Complex” and digital humanities**

Jiang outwardly acknowledges the intrigue behind the digital humanities, stating that there is a “puzzling yet strong complexity binding the humanities and digitalization” (60). After conducting meta-inquiry on classic horror cinema via Media History Digital Library (MHDL) by utilizing Scaled Entity Search (SES) and distant reading, she retrieved some information relating monster cinema and queer spectatorship to each other, demonstrating the significance of Project Arclight and MHDL, as long as the search entities are well designed. Upon the question of whether or not digital methods can be utilized in the future to learn more about classic horror cinema, and just cinema in general, Jiang states, “I believe the meta-inquiry has provided a positive answer” (61). Ultimately, discourse on humanities subjects can be further explored through usage of the digital humanities, leading to a future for the “Frankenstein Complex” and *shocking* revelations.

Works Cited

Healey, Trebor. “Early Gay Literature Rediscovered.” *Huffington Post*, Huffington Post, 7 Dec.

2017, [www.huffpost.com/entry/early-gay-literature-redi\_b\_5373869](http://www.huffpost.com/entry/early-gay-literature-redi_b_5373869).

Hughes, Kit. “Introducing Scaled Entity Search (SES): Interpretive Framework.” *Project*

*Arclight*, Project Arclight, 15 Oct. 2014, projectarclight.org/arguments/introducing-scaled-entity-search-ses-interpretive-framework/.

Hoyt, Eric. "Lenses for Lantern: Data Mining, Visualization, and Excavating Film History’s

Neglected Sources." *Film History: An International Journal*, vol. 26 no. 2, 2014, p. 146-168. *Project MUSE* [muse.jhu.edu/article/548806](https://muse.jhu.edu/article/548806).

Jiang, Tianyu. “Frankenstein Complex” in the Realm of Digital Humanities: Data Mining

Classic Horror Cinema via Media History Digital Library (MHDL). 2019, http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-169638.

“Media History Digital Library.” *Media History Digital Library*, Media History Digital Library,

www.mediahistoryproject.org/.

Mendoza-Pérez, Inés. “Queer-Coding and Horror Films.” *Control Forever*, 26 Oct. 2018,

controlforever.com/read/queercoding-and-horror-films/.

Stevens, Hannah. “Queerness in the Horror Genre, Offensive or Progressive? Fembot

Magazine.” *Fembot Magazine*, 28 Oct. 2019, fembotmag.com/2019/10/28/queerness-in-the-horror-genre-offensive-or-progressive/.