When Your Darlings Bite Back:

A Century of Horror Film about Making and Unmaking

Tonya Howe

MVIS 5703: Capstone Development and Presentation

Spring ‘21 Cohort

Apr 30, 2022

When Your Darlings Bite Back: A Century of Horror Film about Making and Unmaking

> "Kill your darlings, kill your darlings, even when it breaks your egocentric little scribbler's heart, kill your darlings."
>
> — Stephen King, *On Writing: A Memoir of the Craft*

Summary:

This project visualizes the shape of horror film, particularly horror about art subjects, as described in IMDb from 1900-2020. It takes the form of an **interactive network visualization,** as well as **three 2D visualizations** imagined as vinyl wraps in arthouse cinemas. Horror film is deeply influenced by the work of the camera and the act and power dynamics of looking [1]; it is also a genre that is frequently marginalized as unwholesome, pressing the limits of what can/should be seen (Ochoa, 2011). Carol Clover (1992) writes that "Horror privileges the eyes…because it is *about* eyes…watching horror" (p. 167). I extend this observation to its logical conclusion: horror is also, then, about the act of making, creating aesthetic objects or experiences, often through the rubric of unmaking. Does horror about making have a privileged place in the genre? What might this tell us about the way horror film "thinks" of itself? Where are these films produced? How does the sub-genre shift over time? The visualizations each showcase aspects of this subset of horror film, concluding with an invitation to the audience.

Rationale:

When I was growing up, a tomboy amongst boys, I was early introduced to the glory of *Monty Python and the Holy Grail*. To this day, I know almost all the lines, and have no shame when it comes to reciting them, loudly, at any given point in time. The scenes that stuck with me? The Black Knight and the Killer Bunny scenes. Why? They were both *gory* and *aware of their fakery.* It's like when you see the boom mic come down in the corner of the screen and just…hover there. You become aware that this is a film, and once you notice that, every aspect becomes self-aware. I think that horror is, overall, an exercise in camp–to varying extents. One of the most unique features of the genre is the way it is watched in part in order to stimulate laughter on the part of the audience (Wood, 1979; or the feeling of being "grossed out," as Linda Williams notes). Therefore, the self-awareness of the genre is important; as Clover goes on to say, "A strong prima facie case could be made for horror's being…the most self-reflexive of cinematic

[1] See Laura Mulvey's originary statement on the role of looking in film, "Visual Pleasure and Narrative Cinema" (1975), and especially Chapter 4 of Carol Clover's *Men, Women, and Chainsaws* (1992) which addresses the act of looking in horror. For a discussion of the impact of film on the body of the spectator, see Linda Williams' extraordinary essay, "Film Bodies: Gender, Genre, and Excess" (1991).

genres" (p. 168). Does data support this? Ultimately, I plan to write a book aimed at a broad audience about the subject, and this is a first foray.

Data Preparation:

This project uses datasets downloaded as-is from IMDb, as well as text scraped from IMDb. The daily data packages directly available contain quite a bit of information; however, they do not contain summaries, synopses, or country of origin. To collect this textual data, I developed a Python script to scrape the IMDb records by unique identifier. Because I needed to gather comparative data to test my hypothesis about horror as a genre, I gathered this textual data from a variety of other genres, as well–specifically, drama, comedy, thriller, mystery, fantasy, sci-fi, action, crime, animation, and romance. I chose only these genres, because they represent a good cross-section of films in IMDb. This resulted in a dataset of 330,086 rows. To simplify matters, I used only feature-length films, though I have another dataset including shorts.

IMDb genre data by direct download is composed of an array with up to three different associated genres; so too with scraped country of origin data. This meant I needed to aggregate each unique identifier with all relevant combinations. I did this in both R and Tableau Prep.[2]

The keyword tokens took time to develop, and I recreated this list several times. It is still not perfect–and it will never be. I created a script that would look through all the scraped textual data, remove punctuation, tokenize, and associate each film with one or more keyword groups. I used both Python and Tableau Prep to generate this list of keywords, after identifying initial directions:

1.) Art or artist ('art', 'artist', 'artists', 'artista', 'artistthe', 'artes')
2.) Create ('create', 'created', 'creates', 'creating', 'creation', 'creations', 'recreate', 'creator')
3.) Photography ('camera', 'camerawielding', 'photo', 'photograph', 'photographer', 'photographers', 'photographs', 'photography')
4.) Recording or filmmaking ('filmmaker', 'filmmakers', 'videotape', 'recording', 'recordings', 'videotaped', 'videotapes', 'videotapethe')
5.) Painting ('graffiti', 'graffitistyle', 'killerpaint', 'lifeportrait', 'paint', 'paintbrush', 'painted', 'painter', 'painters', 'painting', 'paintings', 'portrait', 'portraits', 'selfportrait', 'wetpaint')

[2] One rabbit hole I went down included connecting countries of origin to World Bank data like GDP. I did not end up using the GDP data, as film industry size, which was generated by counting the number of total films produced by country, was more relevant for the project. But this did reveal problems in the IMDb country of origin data, particularly in terms of current and past country names.

6.) Sculpture or the plastic arts ('miniature', 'sculpt', 'sculpted', 'sculpter', 'sculptor', 'sculptors', 'sculptress', 'sculpture', 'sculptures', 'sculpturesmeanwhile', 'statue', 'statues', 'statuette', 'taxidermist', 'taxidermy')
7.) Dance ('dance', 'dancer','dancers', 'dances',)
8.) Museum, gallery, or exhibition ('museum', 'museuma', 'museumafter', 'museumand', 'museumas', 'museums', 'museumthe', 'museumwith', 'stationmuseum', 'waxmuseum')
9.) Theater or acting ('acting', 'actor', 'actors', 'playacting', 'theater', 'theaterat', 'theaters', 'theatre', 'theatregoers', 'theatres', 'theatrically')
10.) Writing ('author', 'writer', 'Writing', 'writing')

I populated the node list for the network diagram with all existing data as well as generated fields for decade and node type descriptor. The node list contains 4154 rows and 10 columns. The edge list, containing 18,870 rows, connects genre, country, and token to each film to create a directed graph.

Data Analysis and Interpretation:

I was surprised by how many horror films are actually about art in some way. However, this finding was not discovered without some trauma. About three weeks from the due date, I realized I had made a human error in my data analysis while looking just at genre distribution and themes of making–what I thought was 10% was actually 1%. I went back to the data and thought about other ways to show the relations of things. I calculated percentages of all horror films (films with 'horror' in the three-genre array) with *any* subjects of making, by year and decade as well as overall, and then percentages of all not-horror films (films without 'horror' in the three-genre array), in the same way. I used only films whose textual record contained 2 or more references to keyword, to indicate a substantial focus, as earlier efforts resulted in too high a number of films referring to a painting or a statue in a very offhand way.

This proved illuminating. While the raw numbers are quite small, proportionally, there is a vivid difference between horror film about art subjects and not-horror about art subjects. From there, I was able to identify years and decades with high points, proportionally. This became the basis for the several of the 2D visualizations. In 1953, 55% of horror produced was about art. Non-horror films presented only 9% on art subjects in the same year. Overall, the percentage of horror films about art to non-horror films about art is 60/40. While I ended up not exploring genre fully in this prototype, my analysis did reveal relevant patterns–in the 1950s, for instance, when sci-fi and movie musicals seemed to be everywhere, the keyword groups "create" and "dance" were the highest in proportion to all horror films about art subjects. After the 1980s, the disparity between horror films about art subjects and not-horror films about art subjects notably decreased–perhaps due to the vast increase in film production

overall, but perhaps also due to less experimentation in feature films (a factor related to commercial viability).

I wanted to explore the relationship between horror film making and specific countries, which is the subject of the second 2D visualization. This revealed that even amongst countries with the tiniest of film industries, horror films are produced, and even horror films about subjects of making. I did some initial analysis of when horror films appeared in each country's timeline by decade, and only one such visualization (Ireland) made its way into the final prototype. I included an arc diagram to show creativity tokens associated with horror films produced in each country.

The third 2D visualization highlights films that are highly correlated to themes of making by decade, stacked in a stream area graph and divided out. I highlighted a handful of films and noted where they fell in the timeline, including a brief storyline for each. At this point it became clear that the language of the summaries and synopses was notably biased. IMDb, like Wikipedia, is user-generated; we know that Wikipedia editors are predominantly male, white, and middle-aged[3], and it makes sense that this trend would be similar for IMDb contributors. Therefore, the films that are summarized or include a synopsis will include descriptions that are colored by the demographics of the authors–this is particularly notable in the treatment of gender, where women are typically referred to as girls, the subject of sexual assault is treated cavalierly, and so on. When I pulled out the specific films for focus, I edited the summary text both for length and for content, responding to the gender bias visible in many of the existing summaries. I also included a QR code here to the network diagram, with an invitation to the reader-as-fan to contribute their own summaries to IMDb.

The network diagram is geared toward the fan who enjoys discovering new films and unexpected connections. I enabled the user to search or browse by title, keyword, country of origin, and genre. The nodes are color-coded, as well. Each node, when clicked, reveals an information pane that carries more details, like all the countries and genres associated with that film, film summary, date, decade, and so on, as well as a list of all the connected nodes, themselves clickable. I used Gephi and Sigma JS, a plug-in for Gephi that enables the graphs to be hosted live via GitHub. Formatting was done by hand in HTML and CSS.

Audience:

Horror has long been both among the most popular of genres and the most derided of genres (Wood, 1979). My audience therefore is twofold: 1.) genre fans, and 2.) resistant film scholars. I want these audiences each to take something away–validation, and a sense of significance that often comes with history. The network visualization will allow genre fans to find new films,

[3] See Mandiberg, Meyer, and "Gender bias in Wikipedia."

perhaps from other countries and eras. The 2D visualizations help film scholars see that these films have a history. I sought to speak to these different audiences with the overall aesthetic, which tries to walk a line between grotesque and refined. I envision the final prototype existing as vinyl wraps in Landmark E Street Cinema in Washington, DC. These venues tend to attract visitors who identify both as fans and as "serious film" aficionados.

Design:

The design elements of my prototype use muted reds, moody blues or greens, milky off-whites, and not-quite-blacks, a palette drawn from 17th century Dutch still life paintings. I wanted to use colors typically associated with horror in unexpected, more refined ways. The thematic connections between Dutch still lives and horror about unmaking are I think well-suited to this topic, describing as they do the workings of death amidst life and creativity in a very aesthetic manner which is highly contrived, self-aware, and conventional. I've chosen the Bely display typeface for headers, and TT Norms Ultra Light for general text. Bely has something gothic about it, and it contrasts well with the thin and elegant but also slightly off-putting TT Norms.

Application and User Experience:

The three static data visualizations went through numerous iterations, tending toward simplicity. I created coxcomb charts of genres and tokens by decade; radials by country and token and decade; I created dot plots and lollipop charts and interactive maps. I ultimately discarded these because they did not work with the overarching thematic goal, which was primarily to validate this subgenre and make an initial foray into describing its history. I focused on a minimum number of shapes in my final static pieces to create a cohesive aesthetic. I got feedback from cohort members and friends knowledgeable about horror film both to assess iterations of the visual design and the content. This feedback was especially helpful in identifying films to pull out, and in refining the streamgraphs for readability, especially in developing the "read me" callouts. Because the final proposed venue is the long wall space in Landmark E Street Cinemas, I reversed the timeline reading orientation throughout each piece to correspond to how visitors physically encounter the space, walking from right to left.

I worked between Python, R, Tableau, Tableau Prep, and Illustrator, and I also used Flourish, RawGraphs.io, and other online visualization tools. The network analysis uses Gephi, an open-source platform. This is one of the most frequently-used network visualization softwares out there; there is a learning curve, but once the basic structure of a network is understood, the platform is pretty intuitive and very customizable. The learning

curve for all of these platforms, to be honest, was pretty high–if I had to do this over, I would probably just use Tableau to create an interactive, as I ended up drafting many of the vizzes in Tableau and then saving them as PDF to import into Illustrator (which I *also* do not know well).

In terms of design, I initially wanted to use more overt references to 17th century Dutch still life painting, but I ended up refining this into more of a gesture that I hope is still evident to those who are familiar with such paintings.

Future Directions:

This prototype resulted in a mass of data, of which I used only a part. Ideally, this data will inform a book project. Some future avenues for development include using natural language processing to identify keywords, and taking a close look at how arts and artists are depicted. This last will entail a self-generated dataset that looks at key tropes in horror films about making (like blood as paint, bodies as scuplture, and so on) and assesses the role of the artist, the sex of the artist, the sex of the killer, and the significance of the topic to the film as a whole. I have begun work on this dataset, and it was my initial hope that this could be incorporated into this prototype, but alas, time was against me. In the short term, I will develop a Tableau version of the prototype for public display.

References

Clover, Carol. (1992). *Men, women, and chainsaws: Gender in the modern horror film*. Princeton UP.

Ochoa, George. (2011). *Deformed and destructive beings: The purpose of horror films*. McFarland Press.

Mandiberg, Michael. (2020). Where Wikipedia's editors are, where they aren't, and why. *The Atlantic,* 23 February 2020. https://www.theatlantic.com/technology/archive/2020/02/where-wikipedias-editors-are-where-they-arent-and-why/605023/

Meyer, Robinson. (2013). 90% of Wikipedia's editors are male—here's what they're doing about it. *The Atlantic*, 25 Oct 2013. https://www.theatlantic.com/technology/archive/2013/10/90-of-wikipedias-editors-are-male-heres-what-theyre-doing-about-it/280882/

Wikipedia. Gender bias in Wikipedia. *Wikipedia,* 1 March 2022. https://en.wikipedia.org/wiki/Gender_bias_on_Wikipedia

Williams, Linda. (1991). Film bodies: Gender, genre, and excess. *Film Quarterly*, 44(4), 2-13. http://www.jstor.org/stable/1212758

Woods, Robin. (1986). *Hollywood from Vietnam to Reagan*. Columbia UP.