## 1 The Birth of a Field

In 2016 Adrienne LaFrance, writing for *The Atlantic*, described "The Six Main Types of Storytelling, As Identified by an AI." A few months later the *Washington Post* announced that "Researchers Have Quantified What Makes Us Love Harry Potter." A group led by Andrew J. Reagan had published research suggesting that "The Emotional Arc of Stories Is Dominated by Six Basic Shapes," along with an example from the Harry Potter series as shown in Figure 1.

These findings offered proof for Kurt Vonnegut's thesis that all stories have simple shapes. Vonnegut wrote a master's thesis in anthropology for the University of Chicago on these simple shapes of stories. It was rejected but he later reprised it in a lecture that is now available on YouTube. He shows us shapes that look like curved waves that undulate above and below a horizontal axis. The horizontal x-axis plots the unfolding time of the story and the y-axis maps the rise and fall of misery and good fortune. As just one example, Vonnegut draws the shape of the Cinderella fairy tale. As shown in Figure 2, the story starts with misery since Cinderella can't go to the ball like her stepsisters. Then the shape climbs with Cinderella's good fortune since the fairy godmother helps her to go after all. The arc falls again with Cinderella's departure from the ball at midnight before rising one last time with the happy ending when Cinderella is reunited with her prince.

Anthropology as a field has long explored the shapes of stories, whether the folktales of Vladimir Propp and Claude Levi-Strauss – on which more in a moment – or the fairy tales invoked by Vonnegut. Apparently even anthropologists have their limits, however; Vonnegut's thesis was rejected. In the video Vonnegut gives a deadpan performance as we hear the audience chuckle. The simple shape of the story is both fun and, when he points it out, all too simple. Too fun and too simple is Vonnegut's own surmise of why it was rejected, as he writes in *Palm Sunday*.

Vonnegut muses in his talk that there's no reason why the "beautiful shapes" of stories can't be fed into a computer. He was right, although today we are more likely to see these beautiful shapes emerge as output rather than input. Early work modeling the emotional arc of *Romeo and Juliet* was first undertaken by David Bamman, who hired human labelers on Mechanical Turk to analyze the play scene by scene. Ted Underwood then compared their findings with a computational analysis and published the results on his blog, *The Stone and the Shell*. The computational graph that Underwood created, shown in Figure 3, relied on a recent tool developed by Matthew Jockers: Syuzhet. Underwood found that his model and Bamman's comported quite well.

Cambridge University Press & Assessment
978-1-009-27039-7 — The Shapes of Stories
Katherine Elkins
Excerpt
More Information

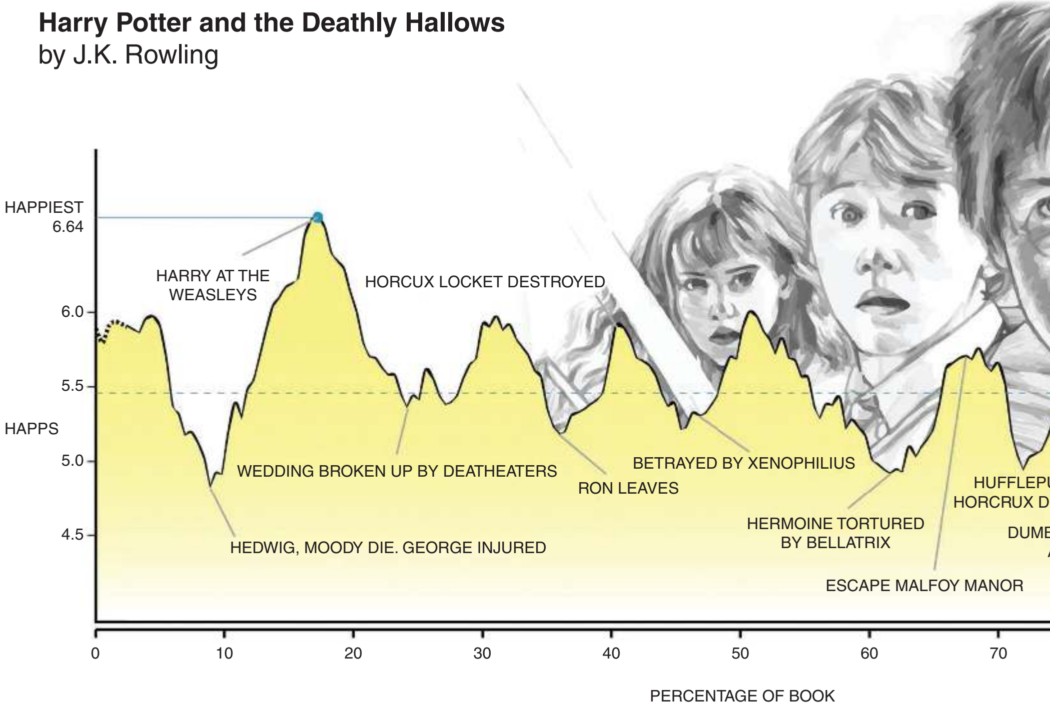


**Figure 1** Sentiment analysis of Rowling's *Harry Potter and the Deathly Hal*

 

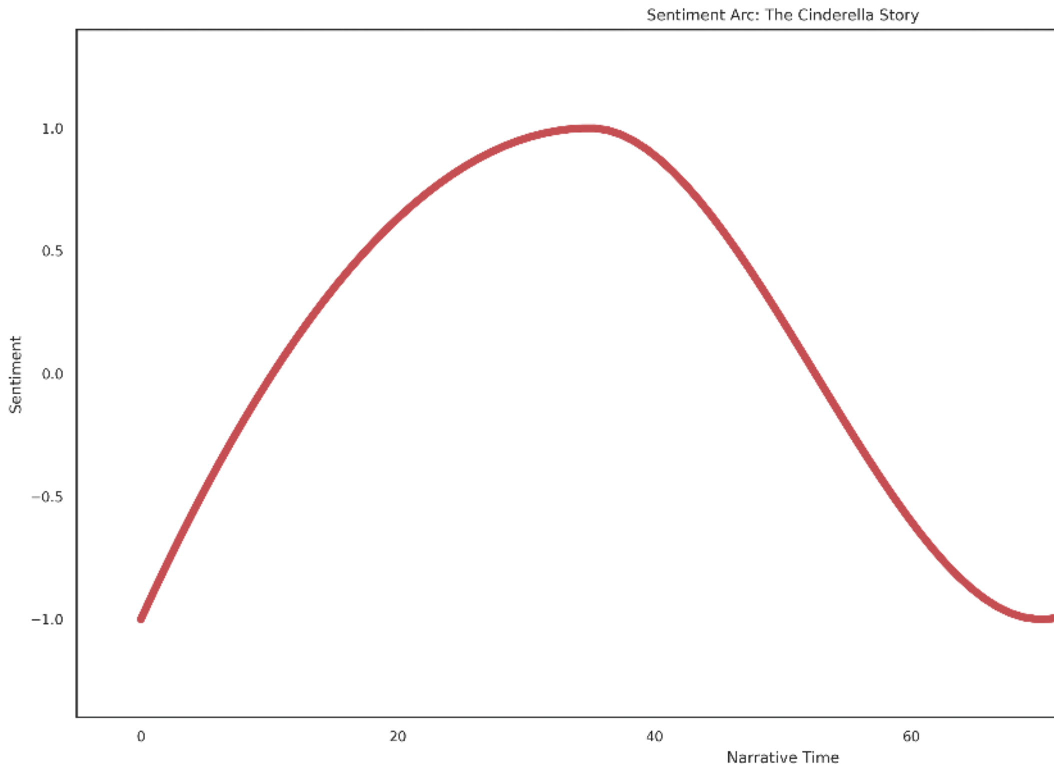


**Figure 2** Story shape: Double person in the hole

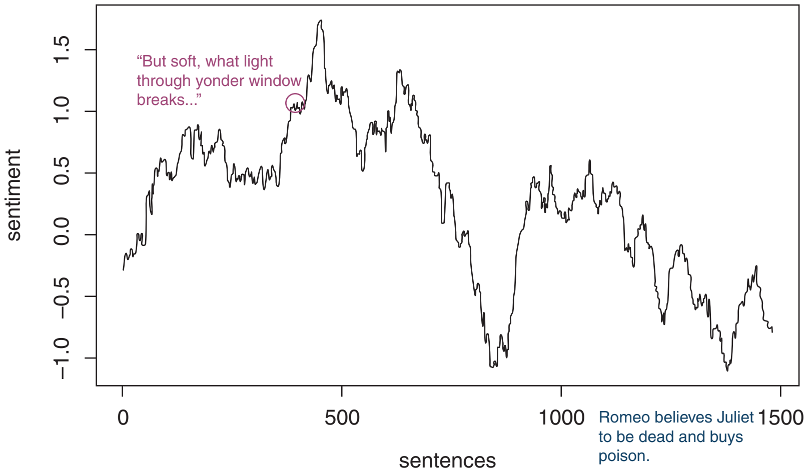


**Figure 3** Sentiment analysis of Shakespeare's *Romeo and Juliet* (Underwood, 2015)

Jockers, who had developed the tool, was also comparing it against hand-coded novels. He too found that the results comport fairly well with human judgment in James Joyce's *Portrait of the Artist As a Young Man* (2015). Tragedy proved easy to graph, but it's a bit more surprising that *Portrait of the Artist As a Young Man* evinced such a clear emotional arc. *Portrait* combines elements of life writing alongside representations of consciousness, aesthetic theory, and even church sermons – hardly a story of the kind Vonnegut describes.

The year after these developments hit the mainstream press Jon Chun and I decided to try out another modernist novel. Would a novel known for being relatively plotless – Virginia Woolf's *To the Lighthouse* – still evince an underlying structure? We found the answer to be yes. This finding suggests that the typical periodization of the modernist novel as highly experimental in form may require more nuance since we have found that modernist novels can in fact surface classic story shapes.

Underwood poses another periodization question. *Romeo and Juliet* may be quite easy to graph, he suggests, given that tragedy conforms to a typical structure. But what about the realist novel, which the literary historian Ian Watt describes as moving away from plots that are so predictable? Could realist novels be harder to plot than many modernist novels *or* tragedy? In fact, we have found in our lab that this approach can reveal latent structure in a wide variety of narrative genres well beyond the realist novel, from screenplays to *Shark Tank* episodes, political speeches to poems, judicial opinions to financial news reporting.

These findings make evident the fact that this method for surfacing arcs does not expose plot directly. Instead, it surfaces an underlying sentiment structure that occurs even when very little happens plot-wise. Reagan and his group call it an emotional arc and, if you try out their Story Lab tool, the "Hedonometer," you can explore the emotional arc within a no-code web browser environment. Their tool offers keywords that contribute to the valence of the sliding window of text. Yet, looking at the list of words and the emotional arc, it's hard to make sense of it from a scholarly point of view. What is happening in each section? How well does it comport with a critic's understanding of the narrative? It is impossible to tell, although one can certainly guess if one knows the novel well.

Jockers's work moves toward an approach more familiar to literary scholars by assessing how well the points of inflection – the peaks and valleys of the sentiment arc – comport with the passages often selected for close reading. In the case of his own investigations, he finds peak detection of the minima and maxima comport surprisingly well to the key passages or "cruxes" often chosen by scholars. Our exploration of more than one hundred novels confirms that the tool often – though not always – surfaces what many scholars might consider key crux points.

Within the digital humanities, sentiment analysis once necessitated the use of the specialized statistical programming language R. Only recently have many of us been working to make it available to those who work in the much more popular general programming language Python. Jon Chun has also worked to make available a much wider array of tools and approaches using easily accessible low-code Jupyter notebooks.[1] SentimentArcs offers a wide range of sentiment models from simpler to the more complex. We have spent several years exploring these many models, many narratives, and many ways of "smoothing" the arc to produce the shapes that Reagan and his group describe. Section 2 on methods and the case studies that follow in Section 3 rely on the foundational SentimentArcs corpus that we created.[2]

---

[1] Infinite thanks are due to Jon Chun, without whom this Element could not have been written.

[2] Chun explains:

> SentimentArcs' corpora consists of 25 narratives selected to create a diverse set of well recognized novels that can serve as a benchmark for future studies. The composition of the corpora was limited by the effect of copyright laws as well as historical imbalances. Most works were obtained from US and Australian Gutenberg Projects (Gutenberg, 2021) (Project Gutenberg, 2021a) . . . Several dimensions of diversity were considered for inclusion including popularity, period, genre, topic, style and author diversity. The first version of our corpus includes only English, although Proust and Homer are included in translation. SentimentArcs has processed a larger set of novels, including some in foreign languages. The initial reference corpus is in English since performance across all ensemble models was uneven in less resourced languages (Dashtipour et al.,

What we have found is that narrative sentiment arc is much more complicated than the Hedonometer might lead one to believe. For one thing, there are many different tools with different methods of surfacing sentiment, and sometimes the thirty-five different models we've worked with can seem like thirty-five different "readers" in the room. We are far from having a single model that is empirically "best" for all narratives and sometimes even the state-of-the-art models can struggle with specific texts. Different tools, we have found, work better for different texts, and the model and text must be jointly optimized for best results. Finding the best fit between a model and a particular narrative can be challenging. What follows in these pages is the description of an ensemble method that relies on considering a variety of models to arrive at an optimal fit.

Narratives are "fuzzy," and in computational terms we need to think of probabilistic confidence intervals rather than simple deterministic point values. There is no way to derive a model that gives simple black-and-white answers with 100 percent certainty. We are also assuredly not in the realm of the hypothesis-driven scientific method that aims to test if our data reveal underlying phenomena distinct from random chance by using tests of statistical significance. Nor, unfortunately, can we use what is termed supervised learning, which relies on a labeled data set to train models that can evaluate new unseen data sets. We cannot label the emotional arc of Virginia Woolf's *To the Lighthouse* and expect it to accurately model Joseph Conrad's *Heart of Darkness*, for example. Each narrative is unique in its own way, and there is no universal ground truth we can use to compare, evaluate, and select the "best" sentiment model.

Instead we are in the realm of probabilistic models and what is termed exploratory data analysis. Optimal model fit will depend on the unique linguistic nature of each narrative, and determination of that fit relies on what is called a human-in-the-loop (Yeruva et al., 2020) to evaluate that fit and arbitrate between the differing results of competing models. A human-in-the-loop or, in the case of narrative analysis, a critic-in-the-loop, presupposes that all models are wrong and that the ultimate judgment as to which model is useful can be determined only by a human reader. This approach stresses the importance of the human expert and relies on computational models as an aid rather than the final arbiter of any ground truth. For that reason, what follows will describe a method that can be used to

2016). SentimentArcs' corpora spans approximately 2300 years from Homer's *Odyssey* to the 2019 *Machines like Me* by award-winning author, Ian McEwan . . . In sum, the corpora includes (1) the two most popular novels on Gutenberg.org (Project Gutenberg, 2021b), (2) eight of the fifteen most assigned novels at top US universities (EAB, 2021), and (3) three works that have sold over 20 million copies (Books, 2021). There are eight works by women, two by African-Americans and five works by two LGBTQ authors. Britain leads with 15 authors followed by 6 Americans and one each from France, Russia, North Africa and Ancient Greece.

assist the reader but never replace the reader. Chun's SentimentArcs makes it easier for scholars to quickly visualize all of this nuance for any narrative, and his tool offers efficient methods to quickly assess divergences that still need to be evaluated by the critic. The pace of innovation in the field is rapid and tools will continue to evolve and change. Nonetheless, methods for evaluation of the models will remain the same and we already have enough tools that work well to begin leveraging them to yield insights in the field.

Until now, there have been a few key reasons sentiment analysis has not been widely adopted. First, it has been difficult to determine the optimal model beyond comparing a few models on specific texts and engaging in guesswork. Second, there has been an absence of clear methodology to get the most out of the approach. Watching both professional scholars and students in the classroom struggle with how to evaluate and validate models has convinced us that best practices – firmly grounded in an understanding of the approach – are needed. Finally, scholars have until now failed to demonstrate how sentiment analysis can yield insights into narrative that are compelling for most scholars. As a scholar who has published widely using more traditional methods, my goal here is to leverage this new approach to yield interpretative results in ways that align with more traditional approaches.

In the following pages I start with the history of the shapes of stories and of sentiment analysis for narrative before turning to the ways in which this approach dovetails with recent trends in literary scholarship. Then I detail methodological aspects of using the tool: close reading of cruxes for the validation of a single arc, evaluation of an ensemble of models for best fit, and assessment of questions surrounding what is called "smoothing." Finally I turn to interpreting the models as a method of yielding insights into individual novels as well as surfacing larger questions surrounding narrative more generally. From an investigation of plotless and postmodern novels to life writing, explorations of race, gender, and colonialism to issues of translation, sentiment analysis can assist the reader in surfacing interpretive insights. Moreover, it has much to show us about the role of emotion in narrative, as well as both the singularity and the shared structure of narratives – the shape and the shapes – of stories.

Even with such a useful approach, the real work begins where the model ends. In spite of the headline, sentiment analysis doesn't actually offer quantitative reasons for why we love Harry Potter, nor does it confirm that all stories share the same shape, though many do once smoothed. For all the newness of the approach, the traditional methods of the literary critic are still needed: intimate knowledge of the narrative and a close reading of the language of sentiment that forms the peaks and valleys of the arc. The scholarly questions raised by the approach are thus more complex than we first expected, and in a good way. For now the aim of what follows is to suggest the many different ways and contexts in which sentiment analysis can be leveraged to explore new questions that may be of interest to literary scholars. Because much of what we analyze in literature are "edge cases" – because so much canonical literature strays from predictable use of language – sentiment analysis is often stretched, sometimes to its limit, by the challenge narratives present.

It is therefore not surprising that sentiment analysis has been widely adopted in other fields before it has been in narrative studies. Mäntylä, Graziotin, and Kuutila (2018) summarize the evolution of the approach, first to mine customer opinion and, more recently, to explore social media. Meanwhile the number of papers published in the field continues to grow exponentially. Our hope is that literary studies will finally adopt it as one approach – among many – to explore the shapes of stories.

## 1.1 A Brief History

Do stories have shapes? Aristotle, in *Poetics*, was the first to suggest that they do. In tragedy, he wrote, all events are interconnected and demonstrate the change in the protagonist's fortune. The shape is determined by these events, which give rise to the protagonist's happiness and misery. Plot and the protagonist's emotions are thus tightly interwoven, and together they form the shape of tragedy. Aristotle was also the first to suggest the importance of emotion in the experience of tragedy. Spectators experience a *catharsis* when they are able to react to the tragedy with feelings of pity and fear. The shape of the story lies in the events that trace the happiness and misery of the protagonist, but this shape of tragedy is intimately connected to our own emotional reaction.

In the nineteenth century novelist and playwright Gustav Freytag outlined a pyramid-like structure for a dramatic plot in which he described a rising and falling structure of actions that forms a triangular shape. In the mid-twentieth century the literary critic Northrop Frye, building on Freytag's pyramid, suggested a U and an inverted U structure depending on whether it was a comedy (U) or a tragedy (inverted U).

In addition to these simple shapes, others have tried to determine the underlying shapes of stories by breaking stories into atomic units. Vladimir Propp focused on the linear unfolding of events, classifying each action as a discrete unit and demonstrating that, in Russian folktales, these essential actions always occur in the same order. Claude Levi-Strauss also broke stories into atomic units, called mythemes, that can be explained, he suggested, by thought processes common to all cultures. Both the formalism of Propp and the structuralism of Levi-Strauss, then, were attempts to determine the underlying shape of stories through an analysis of its most essential elements, whether plot or theme.

Levi-Strauss was probably influenced by the work of Carl Jung, whose theories of unconscious thought formed the basis for an explanation for basic archetypes. Instead of plot events or themes, a Jungian approach focuses on the most basic elements of a "hero" or protagonist, thereby defining archetypes, the most discrete elements of a character, as they are found across stories. Building on these many ways of suggesting that all stories are crafted from the same elements, Joseph Campbell popularized the notion of a "monomyth" and the "hero's journey," combining fundamental plot elements with character archetypes to suggest similar patterns in all stories. More recently, Christopher Booker categorized stories in his *Seven Basic Plots*. He employs the same kind of elemental breakdown of action, giving examples of seven basic plots that include "Overcoming the Monster" and "The Quest."

In addition to all these simple shapes, atomic events, and character types as shown in the timeline of Figure 4, there are recent applications to film. George Lucas credits Campbell's work with helping him bring his *Star Wars* draft to completion, and the final version relies on the fundamentals laid down in the hero's journey. Dan Harmon has further popularized this approach by emphasizing the "story circle" underlying many well-known movies. Indeed this story circle is a common method used in creative writing workshops and classes today. If you are wondering whether screenplays evince the same underlying shape found in novels, our preliminary investigations suggest they often do.

These shapes of stories look a bit different from the triangles and U's described by the first theorists in the field. One question remains, however. You may be wondering, "Are all stories as simple as Cinderella?" and you would have good reason. As the highly esteemed narratologist Jim Phelan once asked me, "If all stories have the same basic shapes, why should we care?" A very good question indeed.

## 1.2 Aren't All Stories Unique?

Shapes fed into (or spewed out of) computers don't interest most literary scholars, and this may be one reason that modeling the simple shapes of stories has failed to gain traction in the field. With our attention to the particular

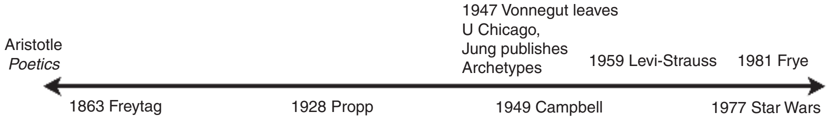


**Figure 4** Theories of story shape

passage and a concern with the unique nature of each story, simple shapes may even seem to work against our attempt to explore the individual aspects of each story. A good way to think about this is to imagine oneself standing at a certain distance from the story. From a distance, one can see the simple shapes that all stories share. From a closer perspective, each story has a pattern as unique as a fingerprint.

While I will touch on questions of the simple shapes of stories, therefore, the primary focus of what follows will be on exploring the unique fingerprint of stories, the *many shapes* of stories rather than the *simple shape* of all stories. In the following sections I'll show how sentiment arcs can help us understand particular texts as well as the differences between the works of a single author. Moreover, as more of us employ these approaches and, collectively, we graph a larger number of stories, we should be able to explore even more deeply the ways in which stories differ from each other. Collectively, we can begin to create a new body of research that investigates differences in sentiment arcs across time periods and genres, across different cultures, and for different readers.

Exploring sentiment arcs, perhaps much like the experience of narrative, can be quite emotional: thrilling and puzzling, inspiring and confusing. Those inclined to think of computational approaches as binary and definitive – all zeros and ones – will be surprised to find that the process of analyzing a sentiment arc feels very much like what we already do. It offers an exploratory method at its very best, often inspiring more questions than answers and requiring close reading and analytical skills to yield insight.

The risk of all computational approaches, of course, is that some will view the computer's output and assume the computer has done the work for them. In fact, knowing what to do with the graph is not obvious at all, and it's taken hundreds of novels and several years of exploration to develop a reliable method. Those who haven't read the story may not be able to glean much beyond the simple shape. Those who understand the story but don't understand the approach will similarly be at a loss. One has to understand how the model works in order to make sense of what it shows.

For that reason this Element will spend quite a bit of time moving between describing the approach and giving examples of how to leverage this understanding by incorporating literary methods. The focus here will be on the kinds of questions that literary scholars already ask using more traditional methods. All too often work in digital humanities is too technical to be of interest to many literary scholars. Oftentimes digital humanists also ask questions that are more in line with the fields of data analytics or the broad sweep of literary history than with what many of us do in examining the unique aspects of a particular story.