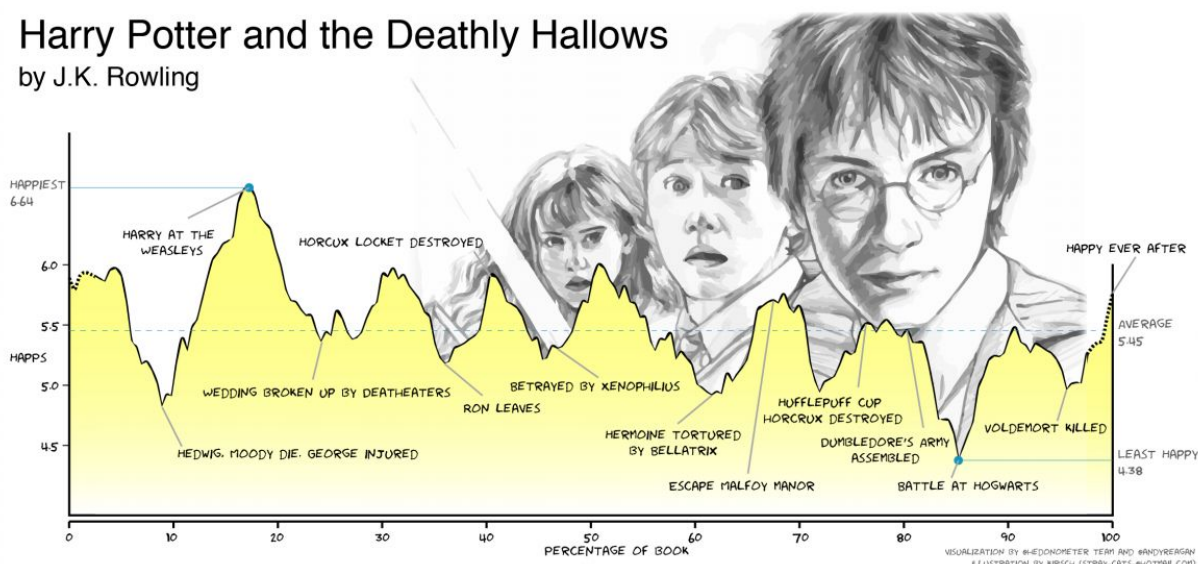


Data mining analyses suggests there are just six main story arcs in Western literature

July 7 2016, by Bob Yirka

Harry Potter and the Deathly Hallows

by J.K. Rowling



Annotated emotional arc of Harry Potter and the Deathly Hallows, by J.K. Rowling, inspired by the illustration made by Medaris for The Why Files. The entire seven book series can be classified as a “Rags to riches” and “Kill the monster” story, while the many sub plots and connections between them complicate the emotional arc of each individual book. The emotional arc shown here, captures the major highs and lows of the story, and should be familiar to any reader well acquainted with Harry Potter. Credit: arXiv:1606.07772 [cs.CL]

(TechXplore)—A team of researchers working at the Computational Story Laboratory (in Vermont) has conducted [data mining](#) and text analysis of approximately 1,700 stories from books available on the

Project Gutenberg website and in so doing have concluded that there are just six main story arcs among them. They have written a paper detailing their study and have uploaded it to the *arXiv* preprint server.

Casual readers and scholars alike have debated the number of story arcs that appear in conventional Western literature, some have become so commonplace that they are considered cliché (boy and girl meet and fall in love, something tears them apart, they are happily reunited), while others are less so. Sadly, despite all the study and debate, no real consensus has been reached. In this new effort, the researchers took a [logical approach](#) to the problem by downloading a lot of [books](#) and then using searching techniques to sniff out story arcs. Their electronic analysis was based on looking for and categorizing emotional polarity in text, using what they describe as 'word windows' which they slid all the way through a story piece by piece.

Once it had been done, a line drawing type of chart could be drawn showing the emotional peaks and valleys as the story unfolded. After that, it was just a matter of running the same program on a lot of books, in this case books that have passed into the public domain—and including only those that were popular enough (by noting number of downloads) to warrant inclusion—and then comparing them. In looking at the averages, the program was able to show that there were just six main story arcs among all the books studied, which the team gave the self-explanatory names: Icarus, Oedipus, riches to rags, Cinderella (which has become the basis of modern romance stories), man in a hole and rags to riches.

The researchers note that there were exceptions to the rules, of course, with some following completely unique paths—the six arcs the computer found were merely the most strongly represented. They also acknowledge that the sample size was small and didn't include more modern works. They plan to continue the work, hoping to expand the

study to other languages to see if they might have more or less arcs.

More information: — The team provides interactive visualizations of all Project Gutenberg books at hedonometer.org/books/v3/1/ and a selection of classic and popular books at hedonometer.org/books/v1/ .

— The emotional arcs of stories are dominated by six basic shapes, arXiv:1606.07772 [cs.CL] arxiv.org/abs/1606.07772

Abstract

Advances in computing power, natural language processing, and digitization of text now make it possible to study our a culture's evolution through its texts using a "big data" lens. Our ability to communicate relies in part upon a shared emotional experience, with stories often following distinct emotional trajectories, forming patterns that are meaningful to us. Here, by classifying the emotional arcs for a filtered subset of 1,737 stories from Project Gutenberg's fiction collection, we find a set of six core trajectories which form the building blocks of complex narratives. We strengthen our findings by separately applying optimization, linear decomposition, supervised learning, and unsupervised learning. For each of these six core emotional arcs, we examine the closest characteristic stories in publication today and find that particular emotional arcs enjoy greater success, as measured by downloads.

© 2016 TechXplore

Citation: Data mining analyses suggests there are just six main story arcs in Western literature (2016, July 7) retrieved 4 May 2024 from <https://techxplore.com/news/2016-07-analyses-main-story-arcs-western.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private

study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.