Young Adult (YA) literature influences readers' understanding of identity and society by acting as "mirrors" of their own experiences and "windows" into others'. The power-danger framework scores words based on the meaning they convey across two axes: weak-powerful and safe-dangerous [Figure 1], allowing analysis of the themes, biases, and attitudes involved in literature that can shape readers' perceptions. Through a partnership with the University of Northern Iowa, this team analyzed word usage in 29 YA novels by Black authors and 27 by White authors with a focus on power-danger patterns.

The team developed a pipeline that tokenizes and scores a set of novels, then returns visualizations such as time series and 2-dimensional histograms for each book [Figure 1, 2]. Histograms revealed a bias towards weak-safe language in all novels. Comparison between author groups found African American authors tend to use slightly more powerful-safe language, which suggests an emphasis on themes of resilience and empowerment. White authors tend to use more weak-dangerous language, which suggests an emphasis on themes of external conflict or personal vulnerability. This project bridges data science and literary analysis, offering new insights into representation in YA novels and novel ways to extract meaning from texts.
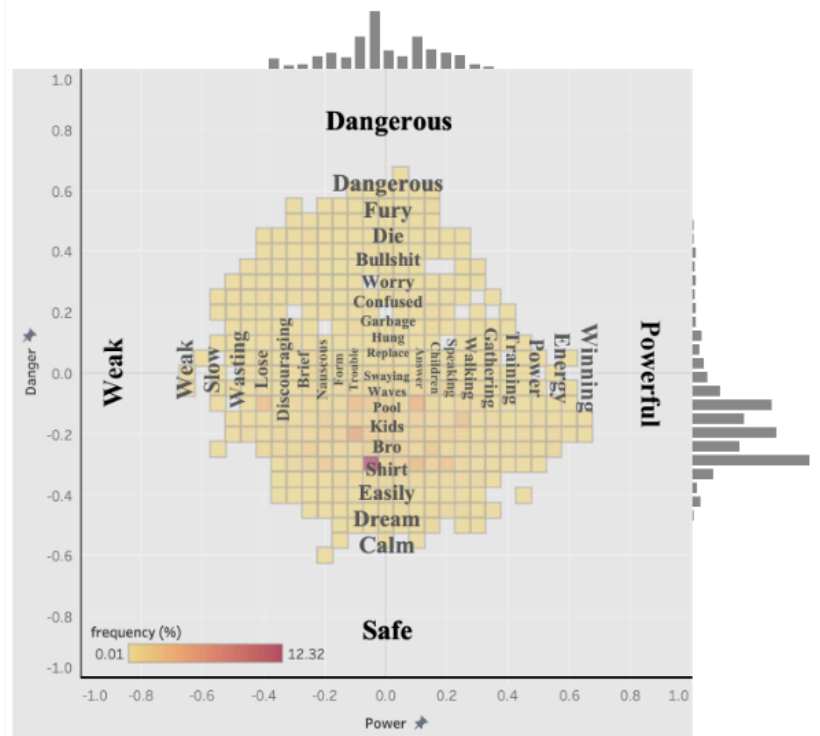

Figure 1. A 2D histogram demonstrating the frequency of power-danger word combinations throughout an entire novel. For the novel "Darius & Twig" by Walter Dean Myers (2014).
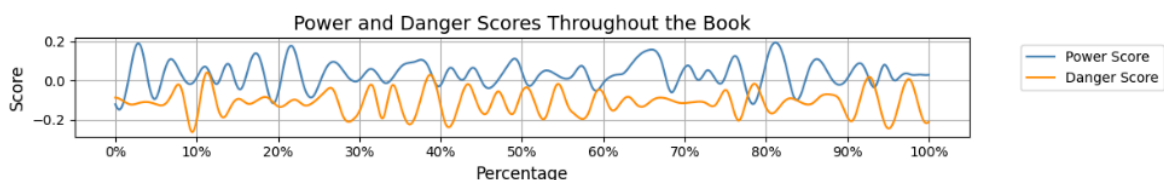

Figure 2. Denoised time series highlighting fluctuations in power and danger scores across the narrative time. For the novel "Darius & Twig" by Walter Dean Myers (2014).