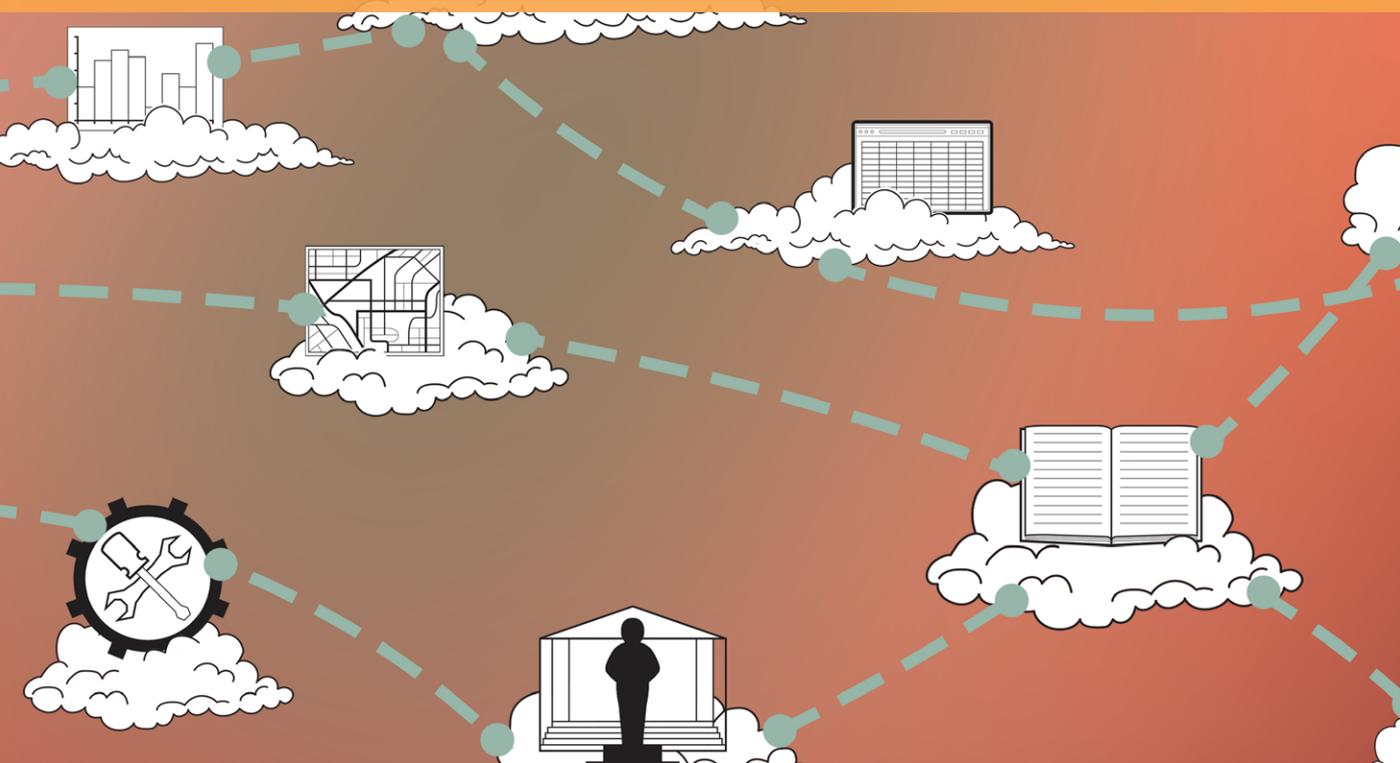


CESTA research anthology 2022



This is an excerpt from the *CESTA Research Anthology 2022*. To read the full *Anthology*, [click here](#). To learn more about CESTA's *Anthology* series and read editions from previous years, [click here](#).

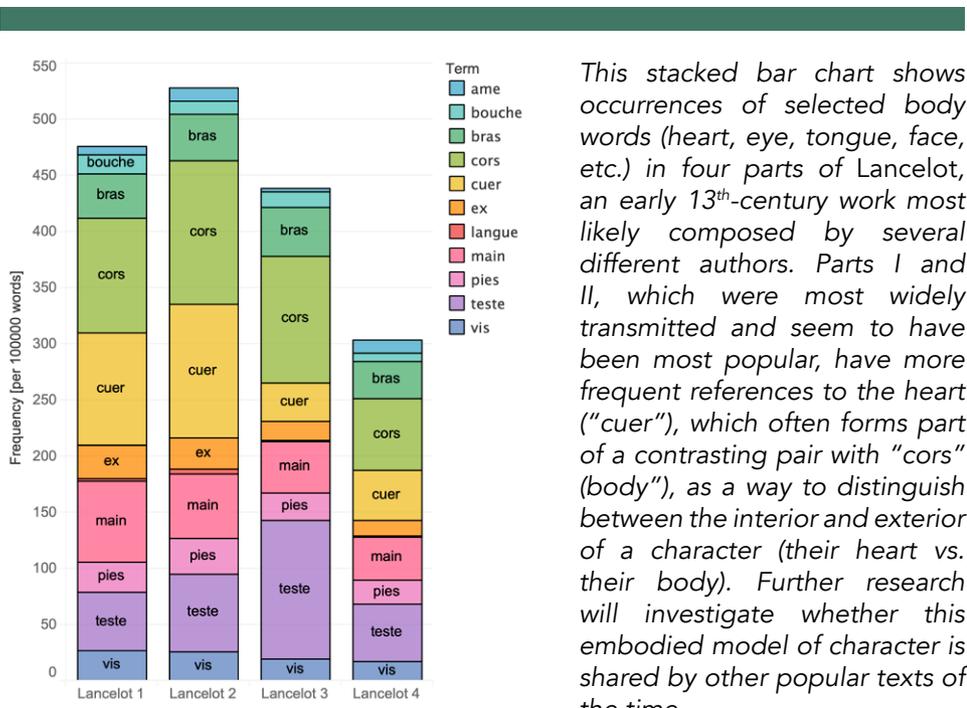


Capturing Formal Innovation in Medieval Fiction by Mae Velloso-Lyons (Comparative Literature)

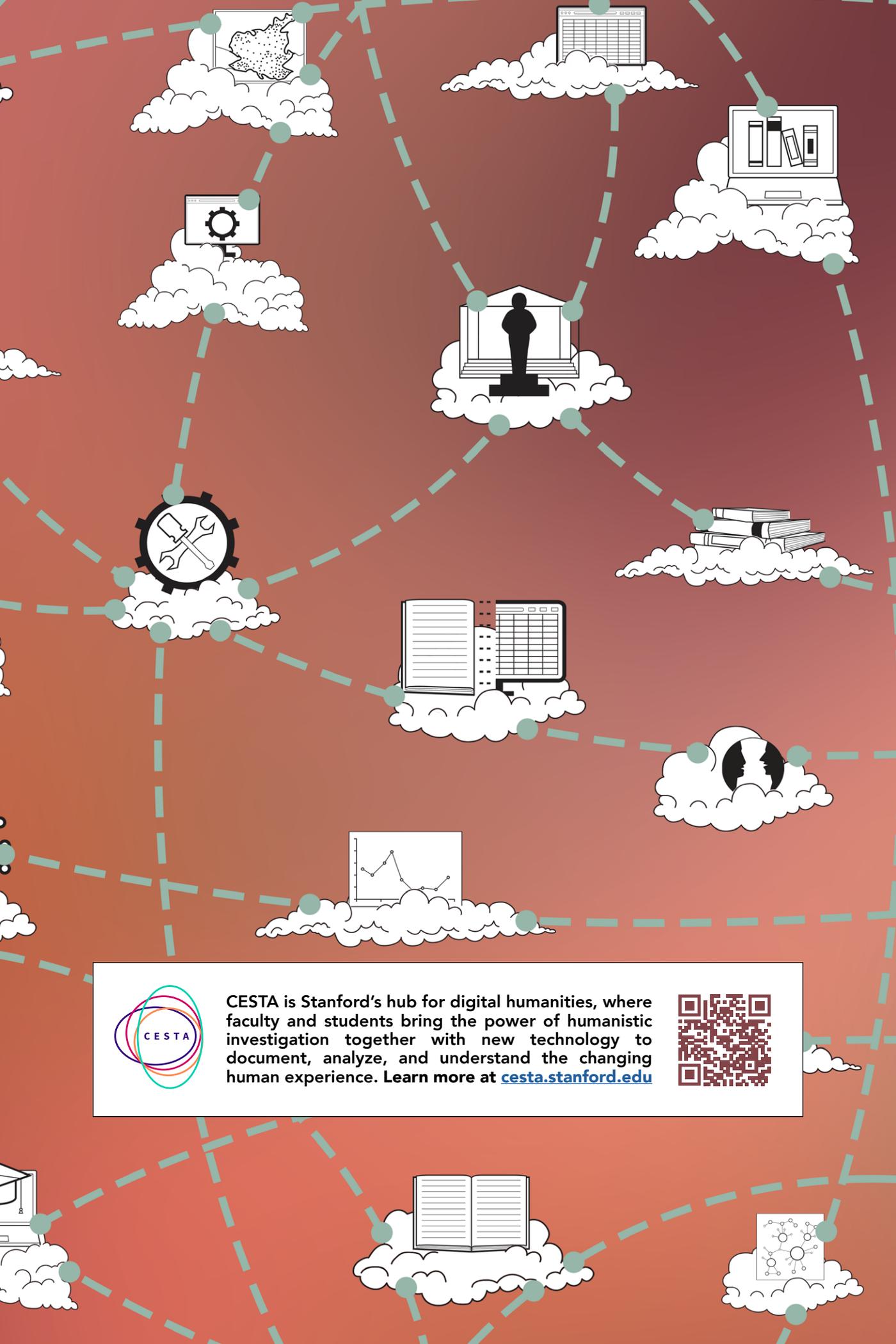
In the first half of the 13th century, several long works of prose fiction were composed in French and gained immediate and widespread popularity, receiving translations and adaptations into a dozen other European languages. These texts were the first novel-like works in Western history since the fall of the Roman Empire and they inaugurated a new age for fiction, with new ways of reading, writing and sharing literature, which have influenced the ways we think about fiction today. Yet these innovative works have received surprisingly little scholarly attention outside of the field of medieval French literature. This project endeavors to build and analyze a digital corpus of medieval prose fiction, using computational methods to identify and quantify formal patterns which would otherwise be difficult to reliably observe and describe.

Building on my earlier research into the *Lancelot-Grail Cycle (LGC)*, one of the best known of these prose fictions, I aimed to create a proof-of-concept by using quantitative methods to test the extent to which the innovative ‘embodiment’ of character in the *LGC* (in particular, the way interiority is expressed through bodily language) is equally present in contemporaneous texts. While my analysis is currently focused on French prose fiction, the goal is to subsequently conduct an analysis across texts in multiple languages, and to make comparisons with fictions in verse.

To build my digital corpus, I used optical character recognition (OCR) software to extract text data from scans of older printed editions of selected works, many of which used non-standard fonts and layouts. My research assistant, Hayn Kim, assisted me with data cleaning by writing custom Python scripts. I now intend to proceed with seeded topic models and the development of a word vector model to assess the consistency of embodied character across my corpus.



This stacked bar chart shows occurrences of selected body words (heart, eye, tongue, face, etc.) in four parts of Lancelot, an early 13th-century work most likely composed by several different authors. Parts I and II, which were most widely transmitted and seem to have been most popular, have more frequent references to the heart (“cuer”), which often forms part of a contrasting pair with “cors” (body), as a way to distinguish between the interior and exterior of a character (their heart vs. their body). Further research will investigate whether this embodied model of character is shared by other popular texts of the time.



CESTA is Stanford's hub for digital humanities, where faculty and students bring the power of humanistic investigation together with new technology to document, analyze, and understand the changing human experience. Learn more at cesta.stanford.edu

