

## **W. Victor H. Yarlott, “AI models for detecting motifs in a text collection”**

**Abstract:** Motifs are distinctive recurring elements found in folklore that have significance as communicative devices in news, literature, press releases, and propaganda. Motifs concisely imply a large constellation of culturally-relevant information – imagine a troll under a bridge, a motif from Scandinavian folklore about which there are many relevant, related ideas: the troll does not own the bridge, yet the troll charges a toll for using the bridge which is enforced with the threat of physical violence. This motif is used effectively when applied to *patent trolls*, who are not the owners of the *idea* they hold a patent for, yet they attempt to extort money through threats of *legal* violence. The broad usage of motifs suggests their cognitive importance as touchstones of cultural knowledge and their cultural relevance hints at their importance for pieces intended to represent an opinion or convince other of an opinion: for example, editorial articles. Thus, we expect that in editorial articles, as compared to non-editorial articles, motifs would occur more frequently. As part of a larger project to develop a system for automatically detecting motifs, we have collected 7,946 news articles containing phrases matching the surface form of an identified motif; we have annotated those phrases as either a true motific usage or not. Of these, 5,109 had either editorial tags or other genre tags; the remaining 2,678 categories did not. Using a high-performance opinion classifier, we re-categorized these articles as either editorial or not, resulting in a total of 117 editorial and 7,829 non-editorial pieces. Calculating the rate of motifs per article, sentence, and token, we found that motifs were roughly three times as frequent (3.75x, 2.95x, and 2.93x, respectively) in editorial articles than in non-editorial articles. We hypothesize this difference is due to editorial articles taking a more casual form of discourse, being crafted to appeal to certain groups, relying on stories for emotional appeal, or arguing for a specific opinion. Further, this strongly suggests the importance of motifs for understanding human communication and the highlights the need for automatic tools for detecting motifs within text. This is part of a research project with scholars at Smart Information Flow Technologies (Laurel Bobrow, Joan Zheng, David McDonald, Chris Miller) and Florida International University (Armando Ochoa, Anurag Acharya, Diego Castro Estrada, Diana Gomez, and Mark A. Finlayson).

**W. Victor H. Yarlott** is a Ph.D. student under the direction of Dr. Mark A. Finlayson at the Knight Foundation School of Computing and Information Sciences at Florida International University (FIU). He received his B.S. and M.Eng. from MIT in 2014 under the supervision of Patrick H. Winston. His research is on narrative and how it relates to culture, emotion, and human intelligence from a computational perspective, focusing on detecting motifs, distinctive recurring cultural “memes,” in narratives.