

Title: Tracing shame between the lines: modeling implicit emotion in Slovenian novels through digital humanities methods

Abstract

Shame has long occupied a central place in literary and cultural analysis, where it is understood not as a simple emotion but as a relational, embodied, and often unspoken experience. In literary texts, shame rarely appears through explicit vocabulary; instead, it emerges indirectly through narrative situations, bodily gestures, silences, and social exposure. This interpretive subtlety poses a challenge for computational approaches to literary emotion, particularly in less-resourced languages.

This paper presents a digital humanities approach to exploring shame in Slovenian novels published between 1800 and 1940. We introduce the first large-scale Slovene literary corpus designed for the study of shame and compare several semi-automatic strategies for identifying shame-related passages, ranging from embedding-based similarity measures and clustering to the use of open, lightweight large language models.

Our results show that while lexical and geometric approaches tend to privilege explicit or prototypical expressions, large language models are better able to capture the implicit, contextual, and culturally embedded manifestations of shame in narrative prose.

Rather than treating these methods as replacements for close reading, we position them as heuristic tools that surface interpretively rich passages at scale: it enables scholars to navigate large corpora by surfacing passages where shame is likely to be at stake, thereby supporting interpretive exploration rather than replacing it.

Keywords

digital humanities; literary emotion; shame; Slovenian literature; implicit affect; narrative analysis; less-resourced languages; computational interpretation

Authors: Boshko Koloski*, Teo Radetić*, Ivana Zajc, Katja Mihurko, Matthew Purver, Senja Pollak

* These authors contributed equally.

Short bios: The authors are a multidisciplinary team affiliated with the Institut “Jožef Stefan” (IJS), the University of Nova Gorica, and Optimal Theta, bringing together academic

researchers of Slovene literature and language technologies, PhD researchers, and industry practitioners. Their work focuses on Slovenian literary history and the creation of corpora for the study of affect and narrative using digital humanities and computational approaches, particularly in the context of less-resourced languages.