

# Using word embeddings to explore the Aboriginality discourse in a corpus of Australian Aboriginal autobiographies

November 1, 2019

## 1 Poster submission - Abstract

The genre of Australian Aboriginal autobiography emerged in the 1960s, with the rise of the Aboriginal rights movement. It is considered a literature of significant socio-political importance, with authors sharing an alternative history as a truth different to the one previously asserted by the European settlers where Indigenous peoples and cultures were either misrepresented or ignored. [1, 3]

With this poster submission, we present intermediate results of the work-in-progress PhD project within the scope of which we are focusing on exploring a newly created text corpus with the word embeddings technique [2]. We have analysed discursive spaces of a corpus of 92 Australian Aboriginal autobiographies published in 1970s-2010s. Here, the particular focus has been on the discourses of race and aboriginality as some of the dominant themes characterizing the genre. The books have been digitised and converted into machine-readable texts, whereas the words have been mapped to real-valued vectors (word embeddings) believed to reflect various linguistic properties and enabling calculating semantic similarity between them.

Word embeddings modelling has recently been used in Digital Humanities as an exploratory technique to complement and guide traditional close reading approaches, which is justified by their potential to identify word use patterns in a collection of texts. Word embeddings capture a lot of cultural

information, as language is considered a reflection of the cultural system of a social group using it.

The results of our initial experiments suggest that the trained word embeddings reflect the historical, political, and cultural environment of the authors and some unique experiences and perspectives related to their racial and gender identity.

## 2 Authors

**Ekaterina Kamlovskaya**, PhD student at the Doctoral Training Unit "Digital History and Hermeneutics", Centre for Contemporary and Digital History, University of Luxembourg. Graduated in 2009 with a degree in linguistics and translation studies from the Kamchatka State University (Russia).

**Christoph Schommer**, Associate Professor in Artificial Intelligence, Computer Science and Communications Research Unit, Interdisciplinary Lab for Intelligent and Adaptive Systems (ILIAS), University of Luxembourg  
`{ekaterina.kamlovskaya, christoph.schommer}@uni.lu`

## References

- [1] Anita Heiss. *Dhuuluu-Yala = to Talk Straight : Publishing Indigenous Literature*. Aboriginal Studies Press, Canberra, 2003.
- [2] Tomas Mikolov, Greg Corrado, Kai Chen, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space. pages 1–12, 2013.
- [3] Adam Shoemaker. *Black words white page: Aboriginal literature 1929–1988*. ANU E Press, Canberra, 3rd ed. edition, 2004.