

# On the unreasonable shallowness of natural language

AKA Adding up word embeddings works far too well

**Lyndon White**

School of Electrical, Electronic and Computer Engineering  
The University of Western Australia

We like to think language is very complicated

This is what we do, and complicated models  
publish well.

and sometimes it is

but sometimes it isn't

and often it looks like it is complicated but  
isn't

SkipGram etc. are basically iterative algorithms for weighted collocation matrix factorization.