

Anonymiser

“Using stylometry to make your text anonymous”

The Team

- Leo Kavanagh
- Federico Cerrato
- Barry Scott
- Conor Moloney
- Sean O'Connor
- Ioannis
- Sebastian Ruder



Problem



In modern society,
retaining anonymity is
getting more
important

In recent years we
have seen what
happens when leaks
are exposed



Currently no easy
way to anonymise
your text

Solution



- Stylometry: Find features of a person's writing style that makes a text traceable to them.
- Use this to inform our predictive model



- Replace these distinguishable features, creating a more “bland” version of the text



- Ensure final output maintains same sentiment and meaning, but is not identifiable

How hard was it?

- VERY!

Did we solve it?

- Kind of

Applied AI

Machine Learning



- Supervised learning to classify input text by author.
- Output distinguishable phrases for authors

Natural Language Processing



- Part of Speech tagging using NLTK
- Find synonyms using Word2Vec / TensorFlow

Sentiment Analysis



- Check accuracy
- Ensure sentiment remains close

DEMO

Impact



- Upward feedback within companies



- “Whistleblower” type situations



- Anonymous Reviews