

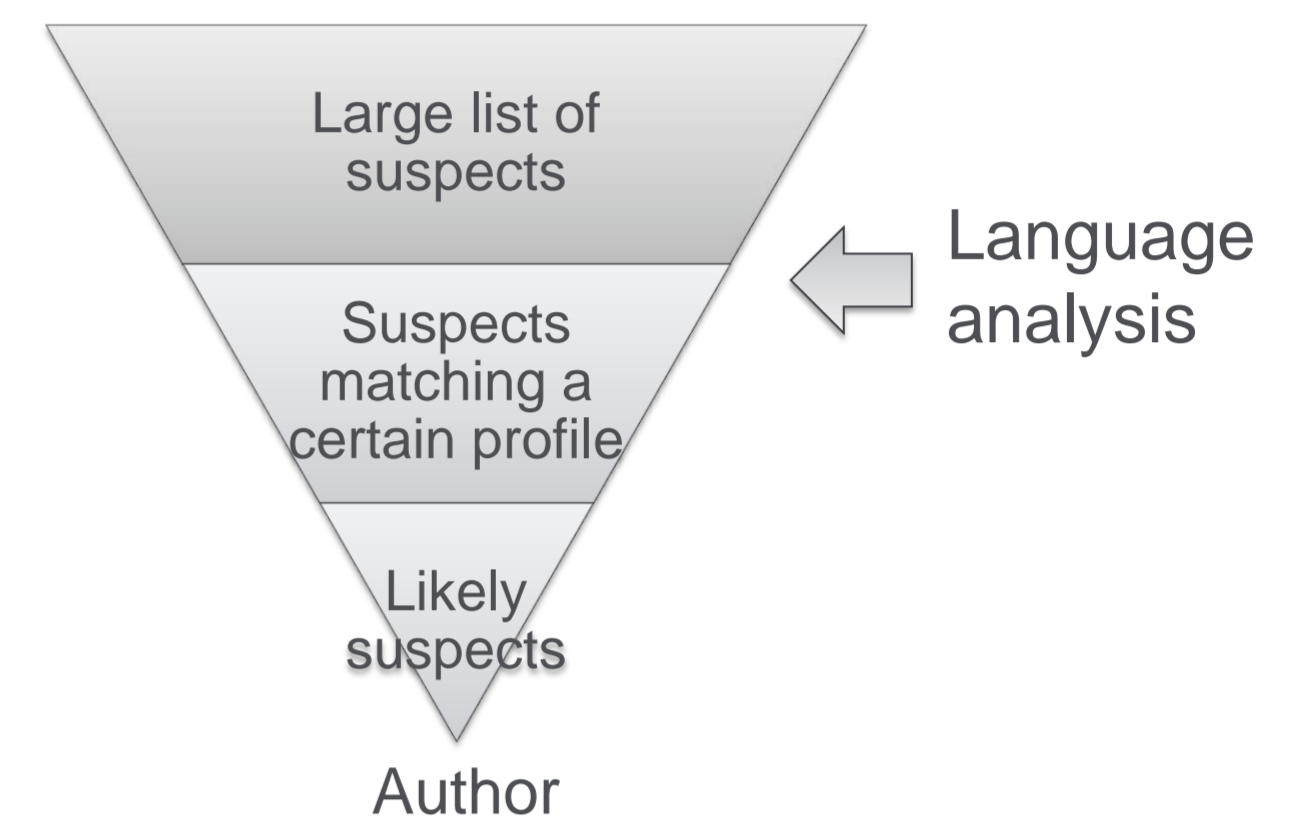
Authorship Profiling in a Forensic Context



Andrea Nini graduated from Aston with an MA in Applied Linguistics with Forensic Linguistics. His research interest lies in the development of a framework for **authorship profiling using quantitative methods**. He is currently completing his Ph.D. on the research of standardized protocols for authorship analysis.

Context

In cases where the anonymous authors of abusive or threatening letters have to be identified, the police or other law enforcement units call in the expertise of linguists to draw up a profile of the most likely suspect. However, a reliable standard model to approach this question is still needed.



Corpus of authentic texts



Texts that are threatening, abusive or spreading malicious information

About 40,000 words
About 130 texts

- Internet searches (e.g. "threatening letter")
- Publicly available data (e.g. Zodiac killer)

Corpus of fabricated texts



95 subjects

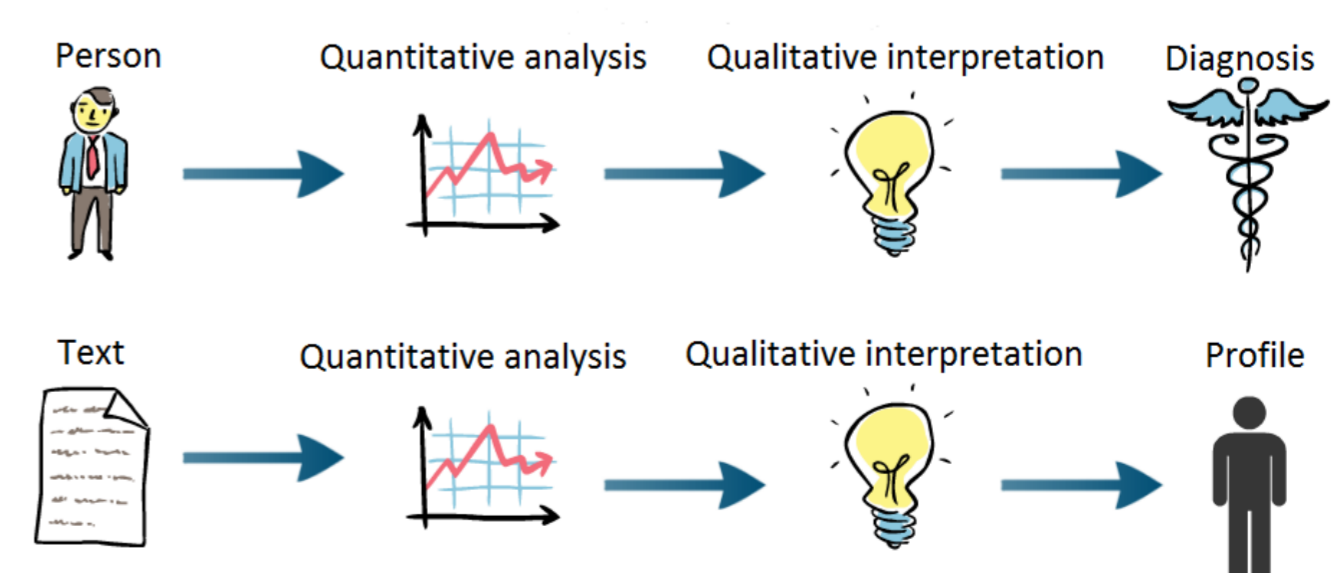
- Gender
- Age
- Level of education

Three 300 word samples:

- Letter of complaint to travel agency
- Letter to Prime Minister about crisis
- Threatening letter to boss

These two data sets are going to be analysed computationally for 130 linguistic variables (e.g. frequency of nouns, verbs, length of clauses) to understand whether they can be used to predict an author's gender, age, education or social class. Research so far suggests that a standard model for profiling could be based on medical diagnosis

Diagnostic Process



Science-informed qualitative process