Also markup the headings

Include:

Mr

Mrs

Dr

Ms

Sir

Apostrophe s

Jr

Sr

Include names in Show titles, or book names

Do not include:

Title like Prime minister, arch bishop

Ignore roman numbers

Features:

Position from stopwords.

Two words beginning with capital letters

Mr,Ms etc preceding

Mr and Ms should be first word

Apostrophe should be last word

Apostrophe succeeding

Distance from a verb

Csv + capital letters

Titles like Leader, Secretary, Prime Minister, Officer, ArchBishop etc preceding the word

Titles following after an ‘is’ or ‘was’ after name..Could have words like ‘nominated’,’elected’ ‘appointed’ etc in between

Sliding window of size 1,2,3,4 for the extractor.