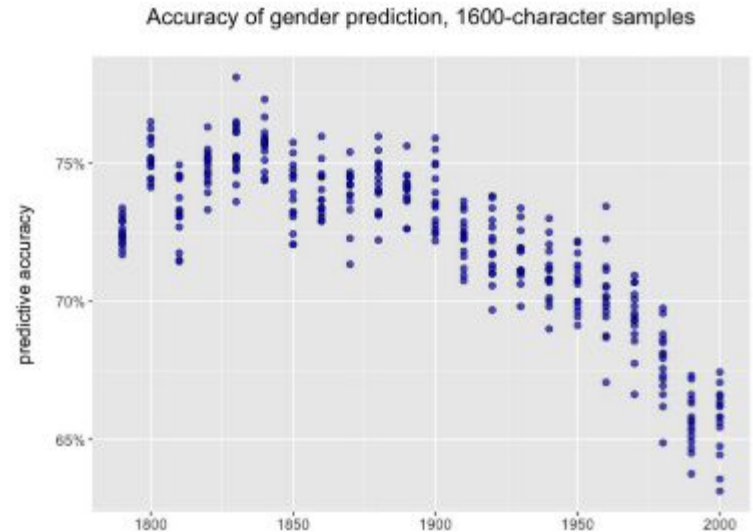# Rolling through the text
Sequential stylometry

ESU 24 @ Cluj-Napoca
Jeremi Ochab, Artjoms Šeļa

Sidenote: "perspectival classification"

- Not all classifications are for improving accuracy

- 'Labels' are cultural conventions

- We can interrogate them

Accuracy of gender prediction, 1600-character samples

Underwood et al. 2019. Transformations of Gender…

Sidenote: "perspectival classification"

- Not all classifications are for improving accuracy

- 'Labels' are cultural conventions

- We can interrogate them

- 'Wrong' classifications is indicative
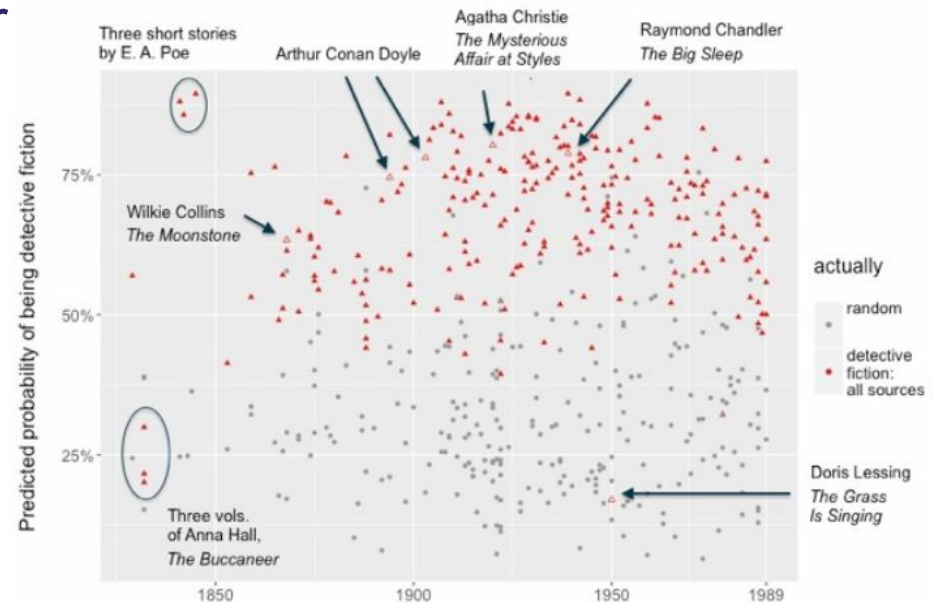
Underwood 2017. Lifecycle of genres
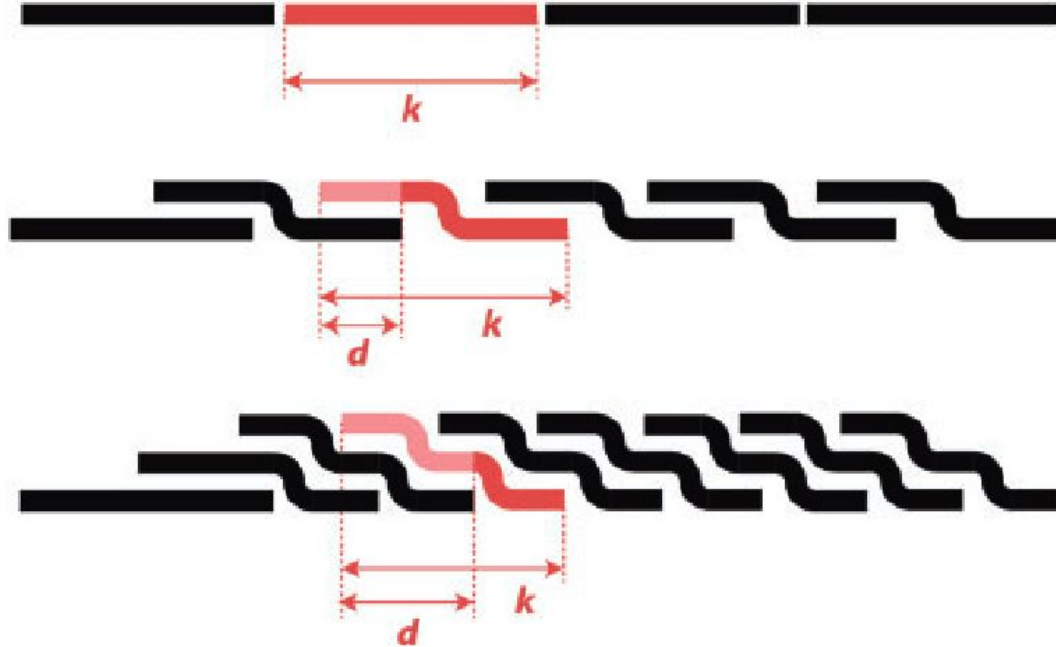
Text-as-sequence

- Heterogeneity

- Multiple authorship

- Editor's signal (?)

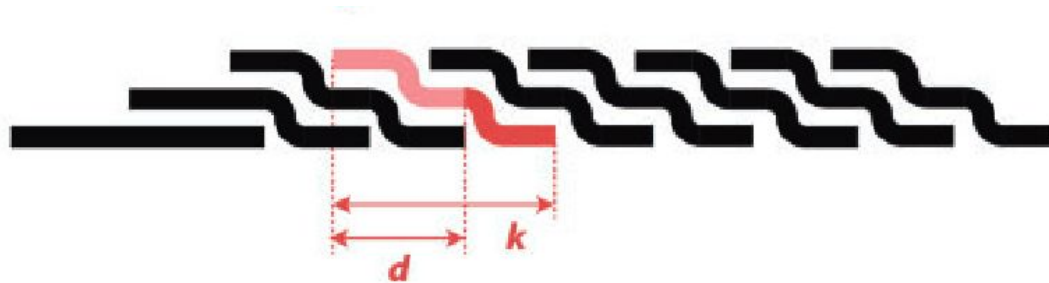- Any kind of textual invasion / injection
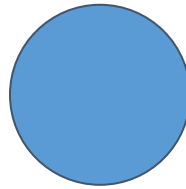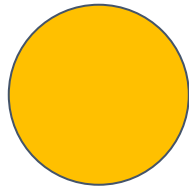
Rolling stylometry

Maciej Eder

Polish Academy of Sciences, Institute of Polish Language, and
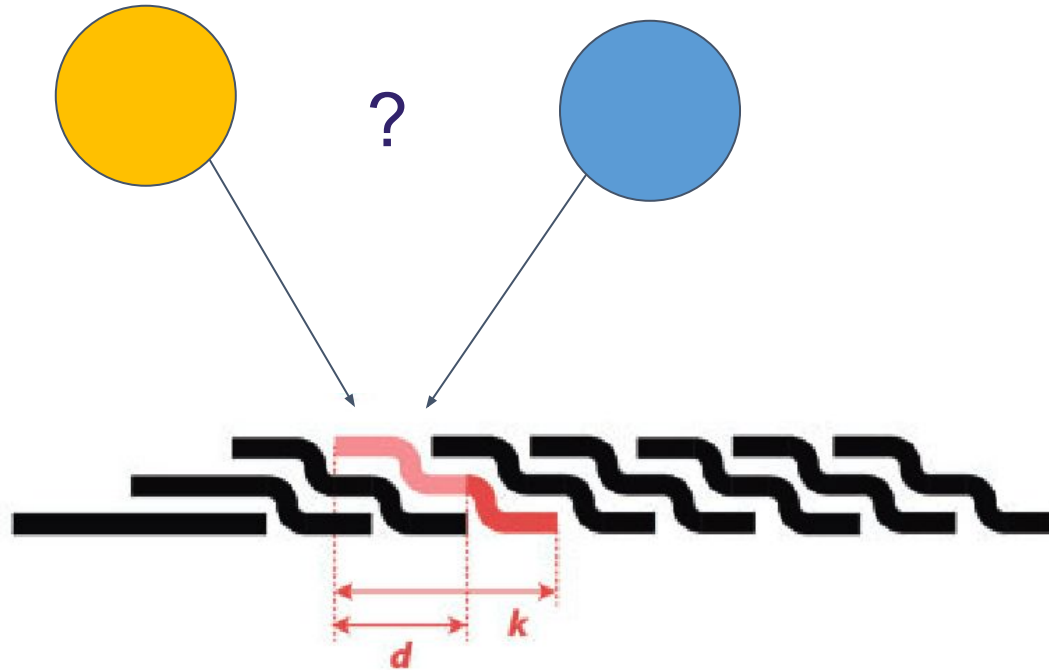Pedagogical University of Kraków, Kraków, Poland
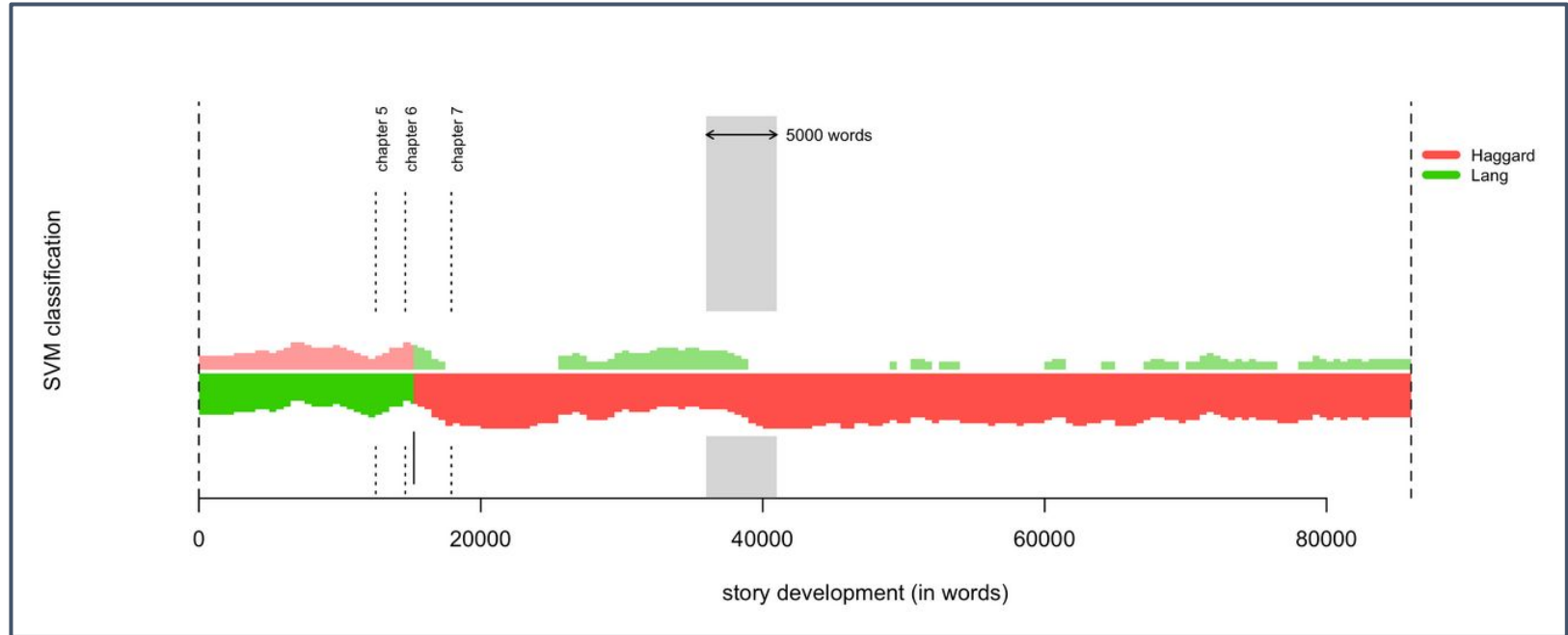
Chunking & overlapping samples

Classification of sequential chunks

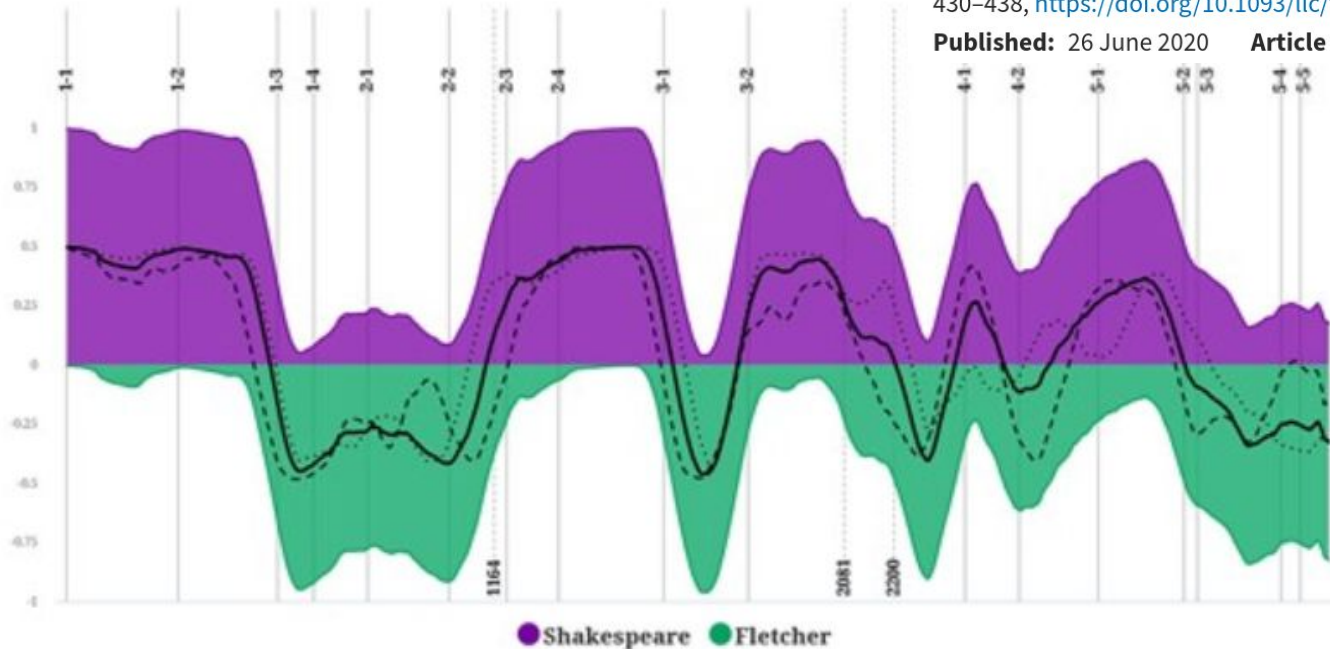# Classification of sequential chunks

Classification of sequential chunks

● Shakespeare ● Fletcher

rolling.classify()

The function doesn't have GUI, so it takes tinkering with arguments to run it:
In this example we have two texts of Ward and two texts of Woolf as **reference set** and *Night and day* novel (Woolf) in **test set**

```
rolling.classify(analyzed.features="w",
                 ngram.size = 1,
                 mfw=500,
                 slice.size = 5000,
                 slice.overlap = 500
                 classification.method="svm")
```
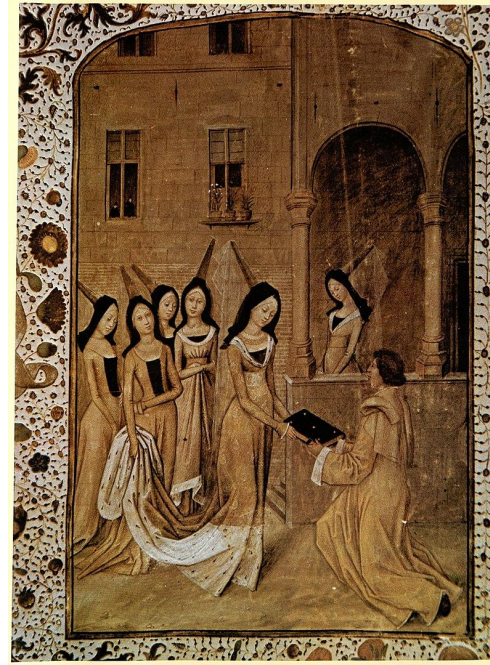
rolling.classify()

The stripe on the bottom is colored according to classification results (red - Ward, green Woolf). A significant portion of text is actually attributed to Ward!
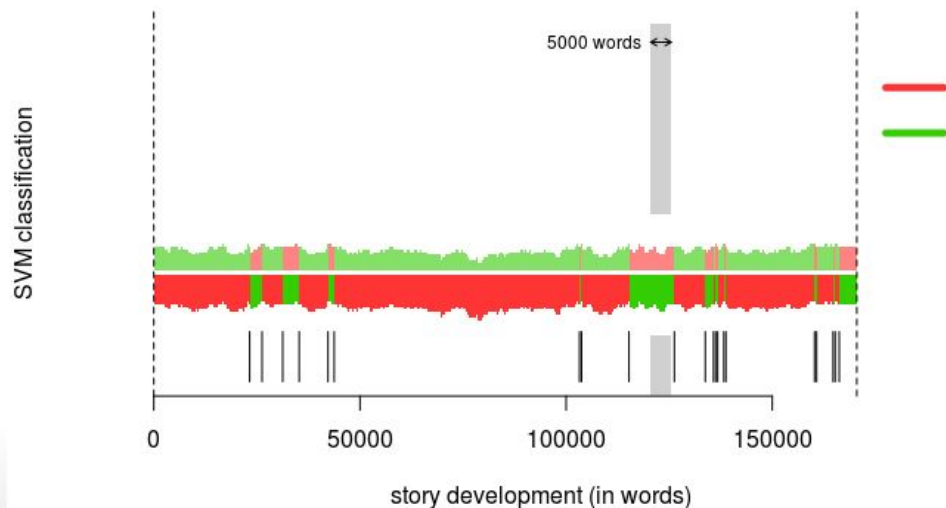
Roman de la Rose (c. 1230-1275)

Guilliame de Lorris

Jean de Meun