

Computational Text Analysis with Stylometry and R – introductions

Joanna Byszuk & Jacek Bąkowski
(Institute of Polish Language of the Polish Academy of
Sciences)

About us

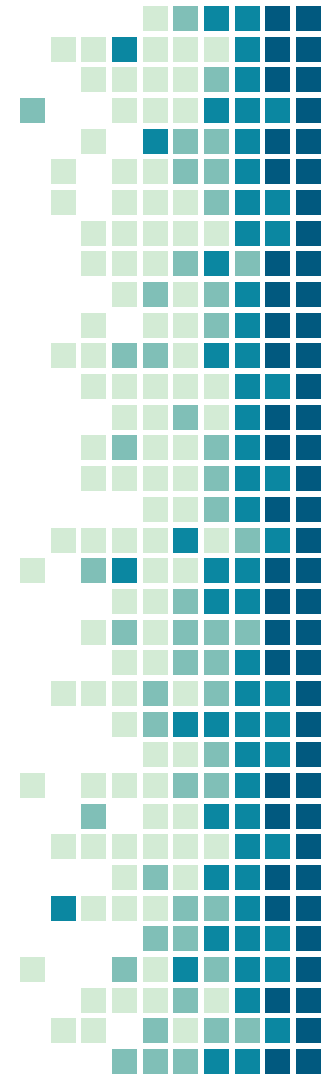
10 Computational 01
01 Stylistics 0101000
11 Group 011010110

PAN **JP** Institute of Polish Language
Polish Academy of Sciences

Computational Stylistics Group in Kraków

Joanna Byszuk – translation studies, linguistics,
electronic information processing

Jacek Bąkowski – mathematics, linguistics



About us

10 Computational 01
01 Stylistics 0101000
11 Group 011010110

PAN **JP** Institute of Polish Language
Polish Academy of Sciences

Research on:

- various aspects of stylometry (methods, languages, problems),
- natural language processing (e.g. HTR, OCR, Direct Speech),
- deep learning in humanities (distributional semantics, multimodal data analysis)

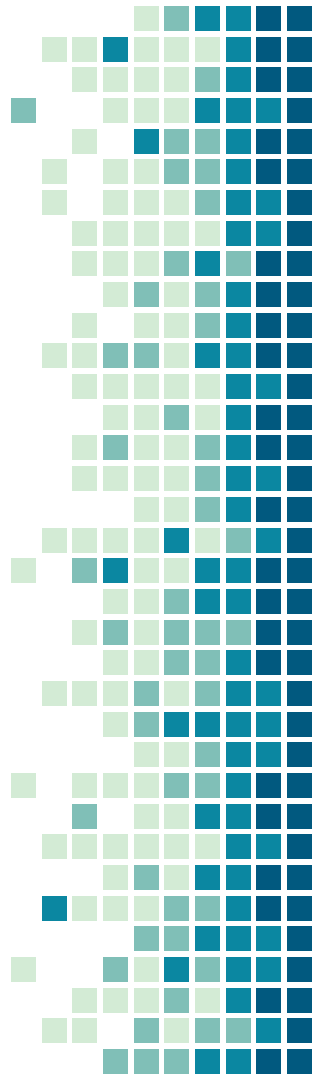


Useful links

<https://computationalstylistics.github.io/>

<https://github.com/computationalstylistics/stylo>

<https://journal.r-project.org/archive/2016/RJ-2016-007/RJ-2016-007.pdf>



Materials

<https://github.com/JoannaBy/DHSI2025-Computational-Text-Analysis> – slides, scripts, anything we'll consider useful

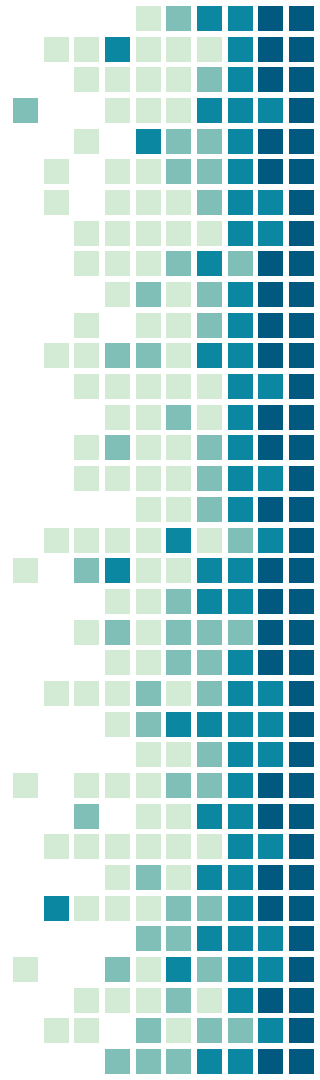


Course plan

Monday 26th May 2025

11.00-12.00 Introductions

13.30-16.00 Introduction to stylometry and textual analysis

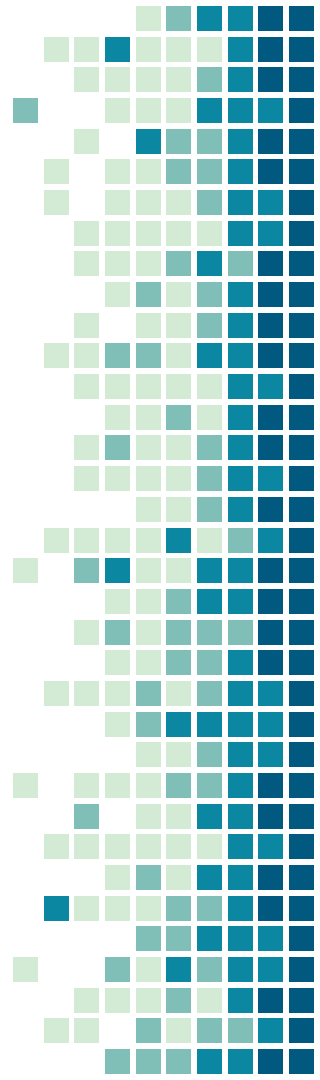


Course plan

Tuesday 27th May 2025

9.00-12.00 Introduction to supervised and unsupervised machine learning, designing a study and first experiments

13.30-16.00 Exploratory analysis – corpus setup and feature engineering, unsupervised methods of text analysis



Course plan

Wednesday 28th May 2025

9.00-12.00 Classification methods and validation of the results

13.30-16.00 Sequential analysis of singular works



Course plan

Thursday 29th May 2025

9.00-12.00 Network analysis

13.30-16.00 Additional useful R functions and tools



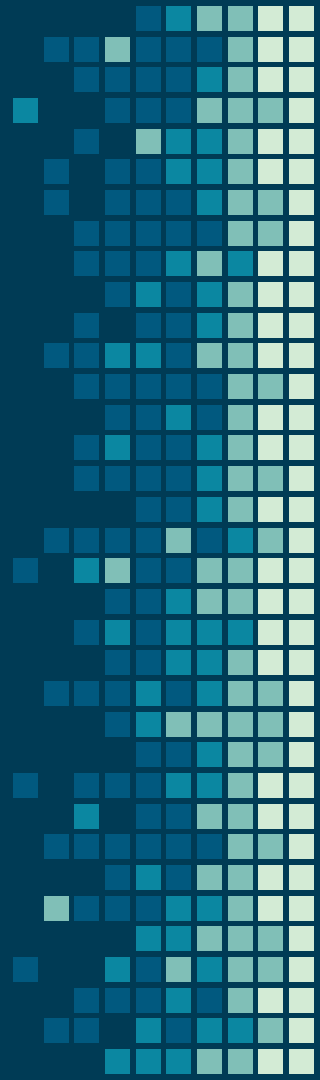
Course plan

Friday 30th May 2025

9.00-12.00 Summary and final works on own projects



And now to you

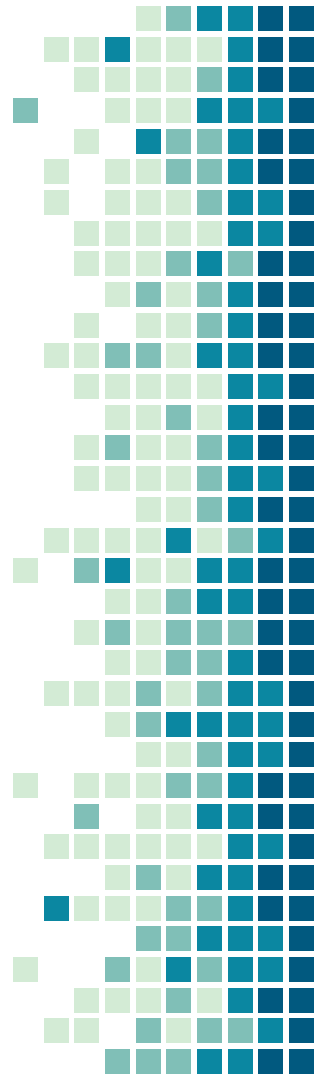


What's your name (and pronouns)?

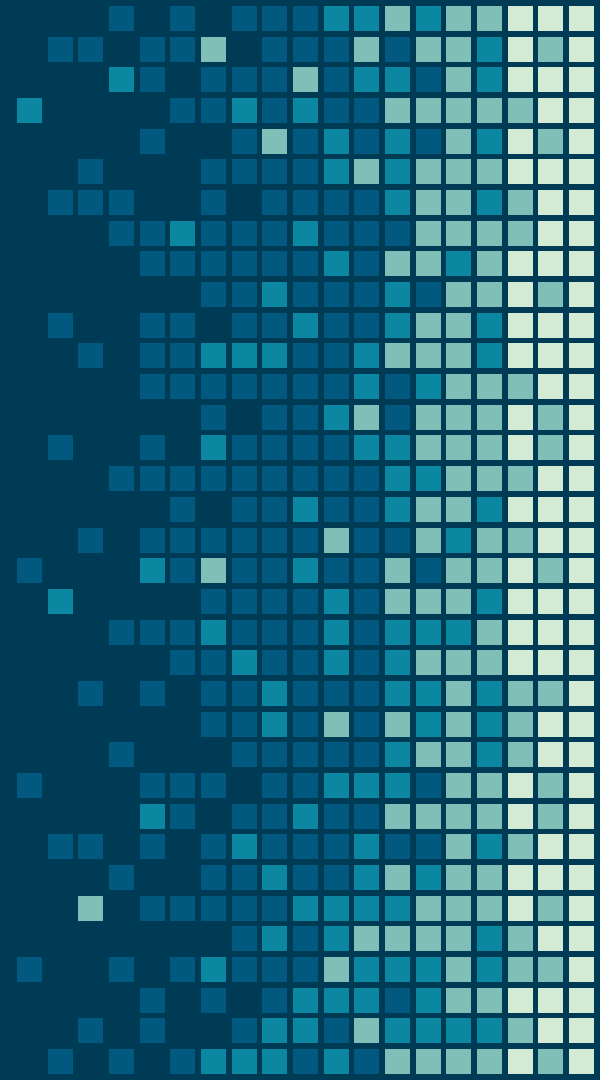
Where are you from?

Why are you interested in stylometry?

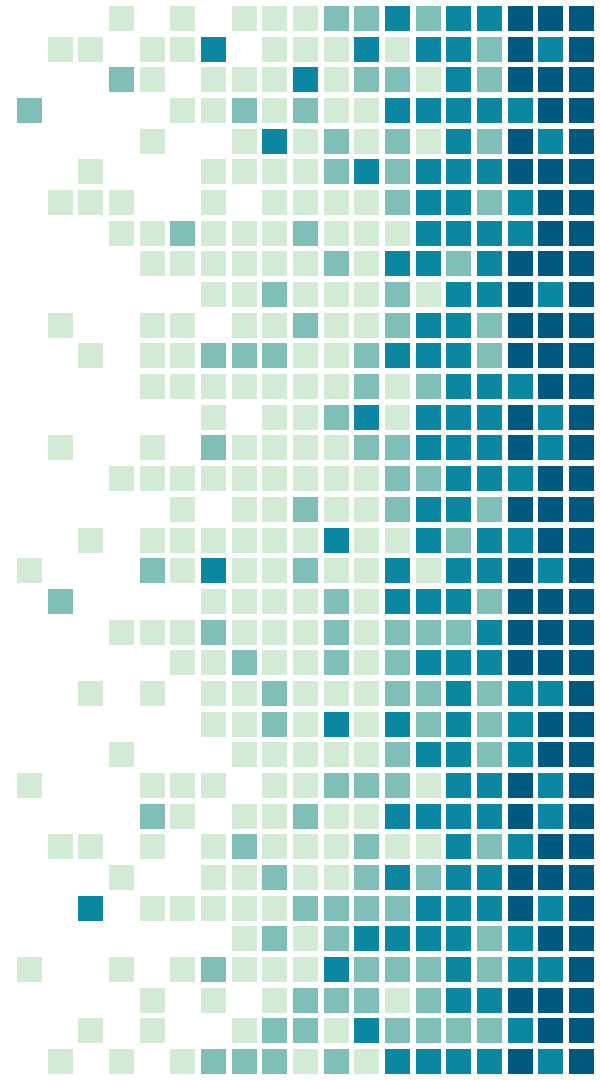
What's the project you'd like to do here?



Testing



Getting started



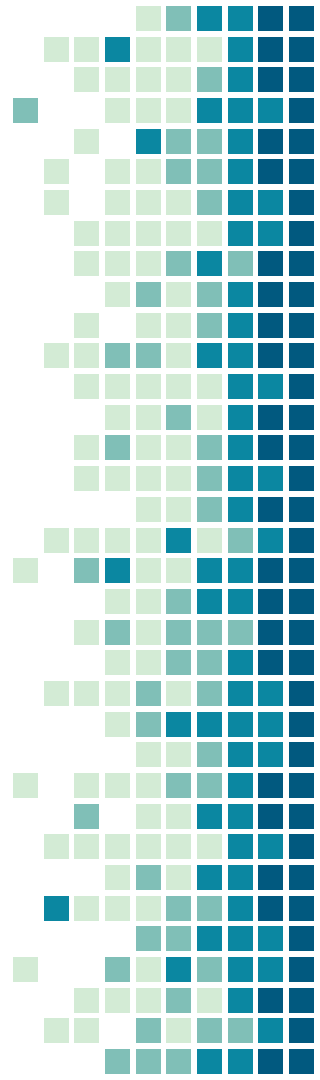
Set working directory:

Command line:

```
setwd("the/path/to/my/favourite/folder")
```

RStudio users: find your directory in the Files panel, then use *Menu > More > Set as Working Directory*

Windows users: use *Menu > File > Change directory*



NEXT:

```
Type: library(stylo)
```

```
And then:
```

```
stylo()
```

