

# Pharmacology research using text analysis in R

Bishwajit Ghose

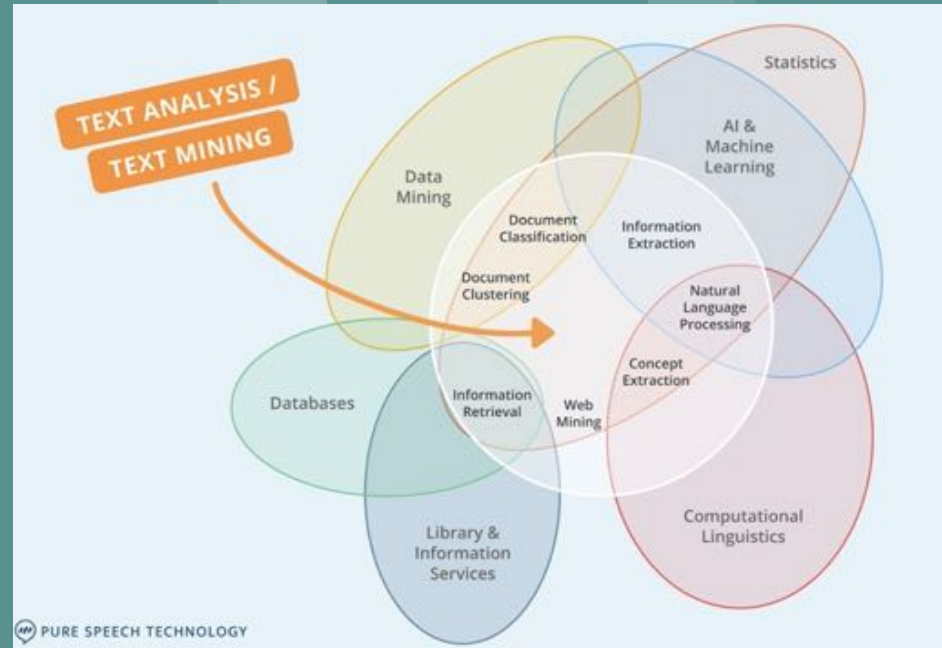


# What is Text Analysis?

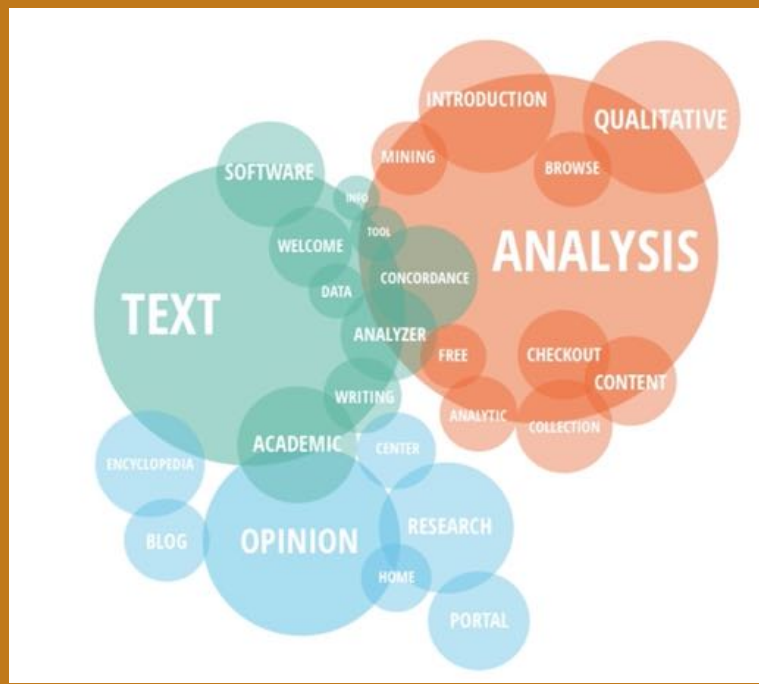
- Text Analysis (TA) aims to extract machine-readable information from unstructured text in order to enable data-driven approaches towards managing content.
- To overcome the ambiguity of human language and achieve high accuracy for a specific domain, TA requires the development of customized text mining pipelines.

# What is the difference between text analysis, text mining and text analytics?

Text analysis, text mining and text analytics are one and the same, and are often used interchangeably.



Text Analysis is the process of analysing unstructured and semi-structured text data for valuable insights, trends and patterns.



It is typically used in instances where there is a need to process large volumes of text-based data for insights, but would otherwise be too resource and time-intensive to be analysed manually by humans.

# So, why do we need this?

- It's precisely because there is so much information that we struggle to truly understand the situation= Information overload/Analysis paralysis.
- We miss crucial insights because there is simply so much. Here's where text analysis comes in.

# Text analysis in R

## Text Analysis in R

Kasper Welbers<sup>1</sup>, Wouter van Atteveldt<sup>2</sup> & Kenneth Benoit<sup>3</sup>

Institute for Media Studies, University of Leuven<sup>1</sup>, Department of Communication Science,  
VU University Amsterdam<sup>2</sup>, Department of Methodology, London School of Economics  
and Political Science<sup>3</sup>

### Abstract

Computational text analysis has become an exciting research field with many applications in communication research. It can be a difficult method to apply, however, because it requires knowledge of various techniques, and the software required to perform most of these techniques is not readily available in common statistical software packages. In this teacher's corner, we address these barriers by providing an overview of general steps and operations in a computational text analysis project, and demonstrate how each step can be performed using the R statistical software. As a popular open-source platform, R has an extensive user community that develops and maintains a wide range of text analysis packages. We show that these packages make it easy to perform advanced text analytics.

Quantitative Methods in the Humanities  
and Social Sciences

Matthew L. Jockers  
Rosamond Thalken

## Text Analysis with R

For Students of Literature

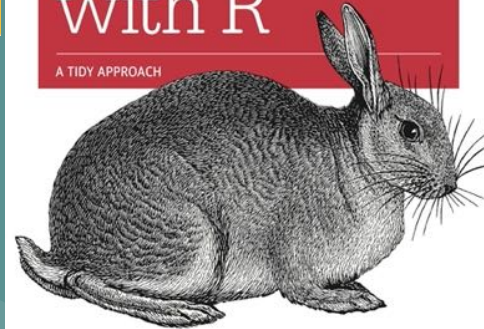
Second Edition

 Springer

O'REILLY

## Text Mining with R

A TIDY APPROACH



Julia Silge & David Robinson

# The 5 Packages We Should Know for Text Analysis with R

1. The All-Encompassing: `Quanteda`  
`install.packages("quanteda")`  
`library(quanteda)`

2. The Transformer: `Text2vec`  
`install.packages("text2vec")`  
`library(text2vec)`

3. The Adapter: `Tidyttext`  
`install.packages("tidyttext")`  
`library(tidyttext)`

4. The Matcher: `Stringr`  
`install.packages("stringr")`  
`library(stringr)`

5. The Show-Off: `Spacyr`  
`install.packages("spacyr")`  
`library(spacyr)`  
`spacy_install()`  
`spacy_initialize()`