

# TEXT MINING

**Fahmi Noor Fiqri**  
By Kodesiana.com



# Text Mining

**Text mining** merupakan proses *data mining* yang berfokus untuk mengolah data berupa teks.

Contoh:

- Sentiment analysis
- Topic modelling
- Chatbot
- Plagiarism detection



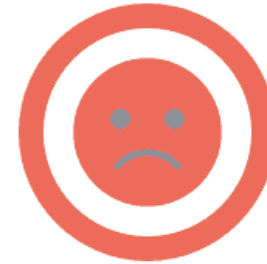
# Sentiment Analysis

**Sentiment analysis** merupakan metode analisis teks untuk mengetahui polarisasi sentimen dari suatu pernyataan.

## Sentiment Analysis



Positive



Negative



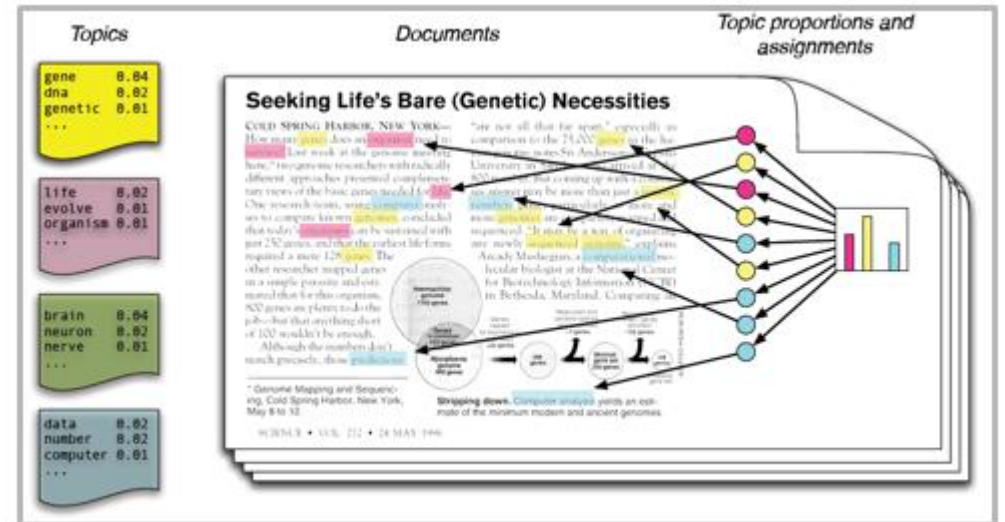
Neutral

# Topic Modelling

**Topci modelling** merupakan metode untuk menentukan kata kunci yang menjadi pembicaraan pada dokumen.

Contoh:

- Latent Dirichlet allocation
- Latent semantic analysis

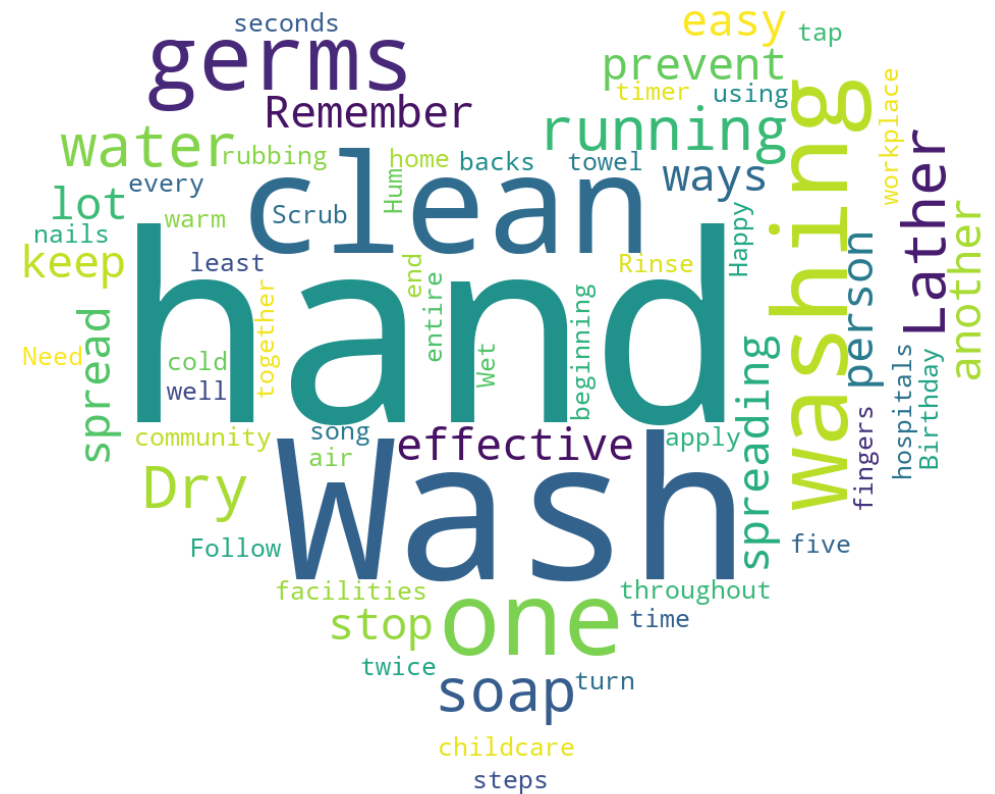


Sumber:

<https://www.analyticsvidhya.com/blog/2016/08/beginners-guide-to-topic-modeling-in-python/>

# Istilah pada Text Mining

- *Corpus*  
Kumpulan teks.
- *Token/word*  
Satu kata.
- *Stopwords*  
Kata yang tidak memiliki bobot.
- *Punctuation*  
Tanda baca.
- *Bag of words*  
Metode representasi token.
- *Word embedding*  
Proses representasi token.
- *Regex*  
Metode pencocokan teks berdasarkan pola tertentu.



# TEXT MINING

**Fahmi Noor Fiqri**  
By Kodesiana.com

