

Natural Language Processing

If any of the following applies to you, you are in the right place:

- As a researcher, do you want to base your empirical findings on linguistic data such as text collections or voice recordings?
- Does your team need help in collecting, managing, processing, or visualising linguistic data?
- Do you want to automate tedious manual tasks?
- Do you intend to make your documents machine-readable?
- Do you want to detect historical trends, political leanings, and personal stance in natural language data?
- Do you have an interesting application in mind which is built on text, audio recordings or eye-tracking data?
- Do you want to use semantic analysis methods like collocations, word embeddings and Neural Networks?
- Do you want to get started implementing language technology solutions yourself?
- Do you have historical documents that you are looking to digitize?

If you need help with the technological side of things, we are here for you!

The [Language Technology](#) group consists of experts who support your text and voice technology needs.

- Text Analytics & Data Mining
- Content Analysis
- Information Retrieval & Extraction
- Part-of-speech tagging, syntactic parsing, semantic tagging
- Sentiment Detection
- Creation and adaption of language models
- Data Classification
- Machine Learning (including Neural Networks & Deep Learning)
- Machine Translation
- Statistical Data Processing
- Audio signal processing (e.g. for human speech or animal vocalizations)
- Speech Recognition and Synthesis
- Data conversion

We offer consulting, coaching, and support in the following scenarios (among others):

- Digitization of printed texts and manuscripts (including OCR)
- Efficient information extraction and analysis of large text collections
- Enrichment of texts with named entities, sentiment analysis, topic modeling, and classification, including multilingual and historical documents
- Advice on tools, software, and best practices
- Help with project applications and common projects

Teasers from our Research

Coaching and Teaching

Example from a Workshop, teaching Content Analysis with R

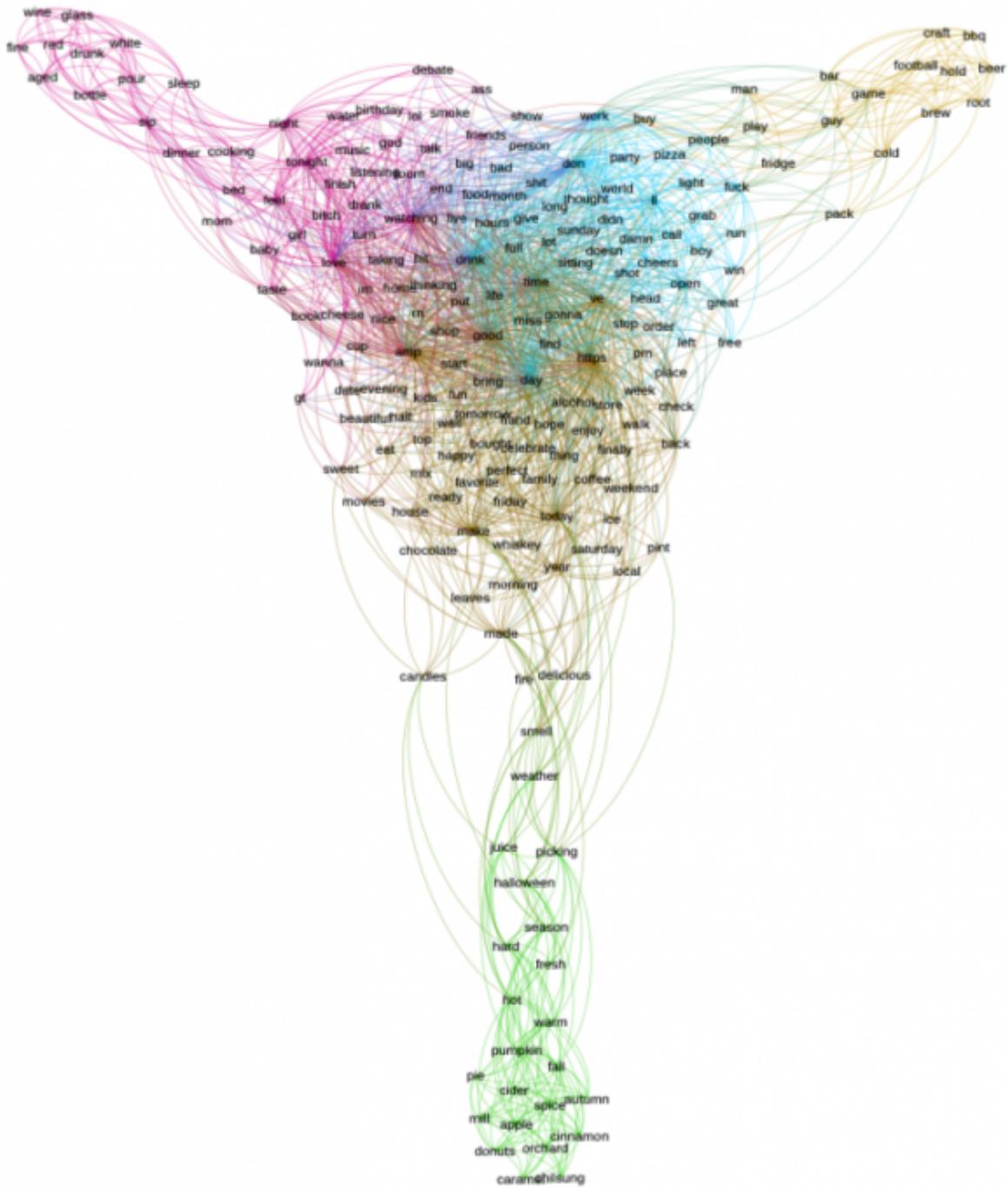
The screenshot shows the RStudio environment with the following components:

- Source Editor:** Contains R code for a function `plot_by_year` that processes a corpus, aggregates word counts by year, and performs a chi-squared test. The code includes comments and function definitions.
- Environment:** Lists global environment variables such as `v`, `w`, `w0`, `w2`, `x`, `y`, and `zipfconstant`.
- Console:** Displays the output of the `plot_by_year` function, including the text "Pearson's Chi-squared test" and a table of data.
- Plots:** A bar chart titled "Relative F of social per 10000 words" showing the relative frequency of the word "social" per 10,000 words across various years from 2000 to 2018.

| Group.1 | x | nc |
|---------|------|----------|
| 1 | 2000 | 0 1061 |
| 2 | 2001 | 1 1142 |
| 3 | 2002 | 12 1809 |
| 4 | 2003 | 21 4558 |
| 5 | 2004 | 14 3789 |
| 6 | 2005 | 13 3835 |
| 7 | 2006 | 31 4671 |
| 8 | 2007 | 43 5325 |
| 9 | 2008 | 49 11126 |

Semantic Analysis

Example of a Result: Conceptual Map of Associations to Wine, Cider and Beer



Stylistics

Characterisation of the Language of Donald Trump

Blick [TV](#) [News](#) [Sport](#) [Meinung](#) [Politik](#) [Wirtschaft](#) [People](#) [Leben](#)

Live auf Blick TV: Das macht heute Schlagzeilen

[News](#) | [Trumps Sprache: Einfach und sprunghaft](#)

26.10.2020, 17:01 Uhr

US-Präsident analysiert

Trumps Sprache: Einfach, sprunghaft und emotional

«Great», «incredible», «deal» und «money»: Diese Wörter benutzt der amtierende US-Präsident Donald Trump besonders häufig. Sein Vokabularreichtum ist im Vergleich mit anderen Präsidenten denn auch mit Abstand am kleinsten.



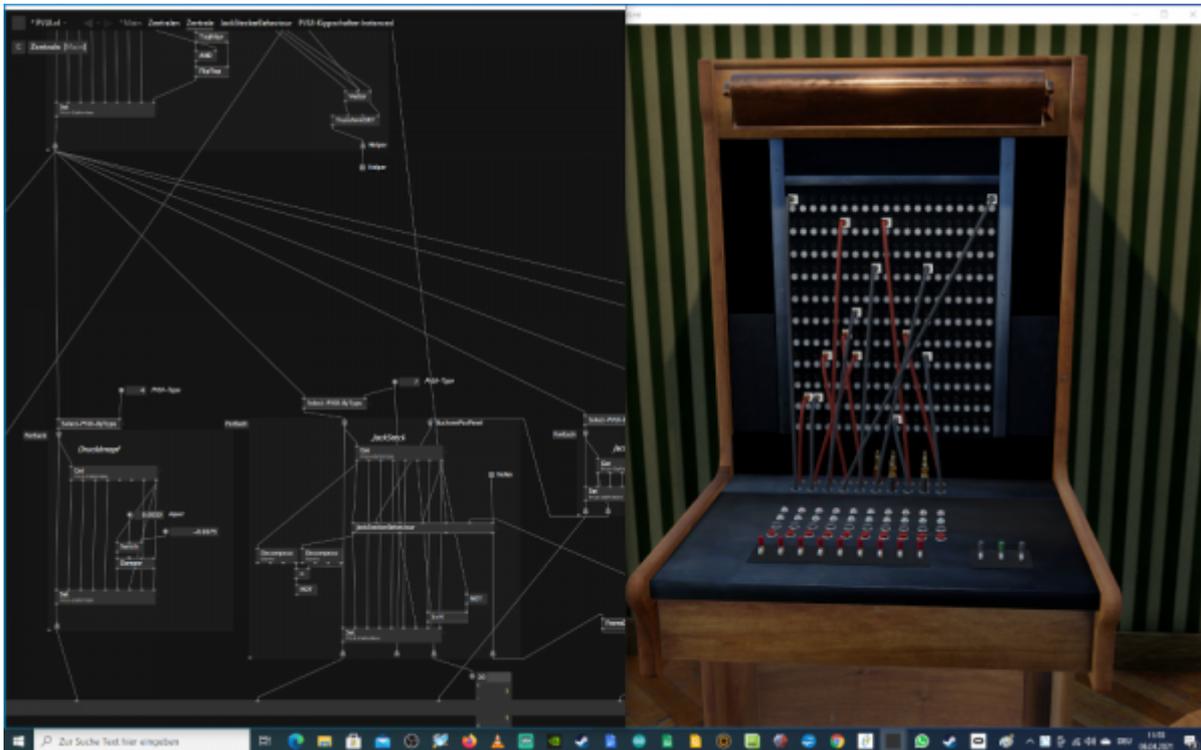
Donald Trumps Vokabularreichtum ist verglichen mit dem früherer US-Präsidenten mit Abstand am kleinsten. Auch seine Satz- und Wortlängen sind besonders kurz.

Das haben Wissenschaftler um den Computerlinguisten Gerold Schneider von der Universität Zürich herausgefunden. Das Team analysierte den Sprachgebrauch des US-Präsidenten und verglich diesen mit dem von früheren Präsidenten und Präsidentschaftskandidatinnen und -kandidaten.

<https://www.blick.ch/news/linguistik-trumps-sprache-einfach-sprunghaft-und-emotional-id16163332.htm>

Examples of our work

Speech-to-text backend for virtual telephone switchboard



<https://www.ds.uzh.ch/de/projekte/vr.html>

From:

<https://liri.linguistik.uzh.ch/wiki/> - **LiRI Wiki**

Permanent link:

<https://liri.linguistik.uzh.ch/wiki/langtech/nlp>

Last update: **2023/02/02 14:26**

