

# Der Einsatz von Distant Reading auf einem Korpus deutschsprachiger Songtexte

Thomas Schmidt, Marlene Bauer, Florian Habler, Hannes Heuberger, Florian Pilsl, Christian Wolff

Media Informatics Group, Universität Regensburg

# DHd 2020

Kontakt: Thomas.Schmidt@ur.de

# 1 Korpus-Akquise



# LyricWiki

STROLL AROUND ▾ BROWSE LYRICS ▾ HELP OUT ▾ ADVANCED SEARCH ▾

## Home

### Welcome to LyricWiki

We are a free wiki website where anyone can get reliable lyrics for any song by any artist.

Please note: this wiki is currently closed for editing. However, you are still free to browse the lyrics pages!

#### Browse Lyrics By:

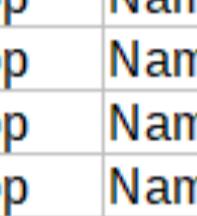
[Artist](#) [Album](#) [Song](#) [Label](#)

Artist begins with:

A • B • C • D • E • F • G • H • I • J • K • L • M • N • O • P • Q • R • S • T • U • V • W • X • Y • Z

### iTunes Top Songs

# Scraping nach vordefinierter Liste von Künstlern auf der Plattform LyricWiki

| in: Song, Bronze Songs, Language/German, and 8 more  |           |           |      |                         |                                      |   |
|--|-----------|-----------|------|-------------------------|--------------------------------------|---|
| <h1>Nena:Leuchtturm Lyrics</h1>  |           |           |      |                         |                                      | * |
| <p>Leuchtturm</p>  |           |           |      |                         |                                      |   |
| <p>This song is by <a href="#">Nena</a> and appears...</p>   |           |           |      |                         |                                      |   |
| <ul style="list-style-type: none"><li>▪ on the album <a href="#">Nena (1983)</a></li><li>▪ on the album <a href="#">99 Luftballons (1984)</a></li><li>▪ on the album <a href="#">Die Band (1991)</a></li></ul> |           |           |      |                         |                                      |   |
| <p>The German title of this song is "Leuchtturm". The English translation is</p>   |           |           |      |                         |                                      |   |
| <p>Not what you were looking for? This page is one of several.<br/><a href="#">Nena:Leuchtturm</a></p>                    |           |           |      |                         |                                      |   |
| <p>Ich geh mit dir, wohin du willst<br/>Auch bis ans Ende dieser Welt<br/>Am Meer, am Strand, wo Sonne scheint<br/>Will ich mit dir alleine sein</p>   |           |           |      |                         |                                      |   |
| <p>Komm, geh mit mir den Leuchtturm rauf<br/>Wir können die Welt von oben sehn<br/>Ein U-Boot holt uns dann hier raus<br/>Und du bist der Kapitän</p>  |           |           |      |                         |                                      |   |
| Genre  | Interpret | Album     | Jahr | Titel                   | Text                                 | * |
| Pop  | Namika    | Que Walou | 2018 | Programm                | Ich fühl' mich wie ein Pixel In ei   |   |
| Pop  | Namika    | Que Walou | 2018 | Liebe Liebe             | Hey du Ich wollt' wissen, wie's i    |   |
| Pop  | Namika    | Que Walou | 2018 | Dschungel im Kopf       | Gedanken ranken sich Um alles        |   |
| Pop  | Namika    | Que Walou | 2018 | Parkbank                | Sie hat nie die weite Welt gesel     |   |
| Pop  | Namika    | Que Walou | 2018 | DNA                     | Du und ich, auf lange Sicht, das     |   |
| Pop  | Namika    | Que Walou | 2018 | Ahmed (1960-2002)       | Ende der Achtziger, sie besucht      |   |
| Pop  | Namika    | Que Walou | 2018 | Hände                   | So viele Jahre auf der Welt, wa      |   |
| Pop  | Namika    | Que Walou | 2018 | Comic                   | Tut ja gar nicht weh, was du sa      |   |
| Pop  | Namika    | Que Walou | 2018 | Roboterliebe            | Elektrisch aufgeladen, haben ur      |   |
| Pop  | Namika    | Que Walou | 2018 | Kronleuchterlicht       | Die Schlüssel zum Lotus Neber        |   |
| Pop  | Namika    | Que Walou | 2018 | Ich will dich vermissen | Du willst alles wissen Willst mi     |   |
| Pop  | Namika    | Que Walou | 2018 | Zirkus                  | Hab' für meine hundert Besten I      |   |
| Pop  | Nena      | Nena      | 1983 | Kino                    | Jeden Abend um die gleiche Ze        |   |
| Pop  | Nena      | Nena      | 1983 | Indianer                | Indianer wie du und ich Die sind     |   |
| Pop  | Nena      | Nena      | 1983 | Vollmond                | Am Abend geht die Sonne unter        |   |
| Pop  | Nena      | Nena      | 1983 | Nur geträumt            | Ich bin so allein Ich will bei dir s |   |
| Pop  | Nena      | Nena      | 1983 | Tanz auf dem Vulkan     | Engel haben keine Schonzeit Tr       |   |
| Pop  | Nena      | Nena      | 1983 | 99 Luftballons          | Hast du etwas Zeit für mich? Da      |   |
| Pop  | Nena      | Nena      | 1983 | Zaubertrick             | Heimliche Blicke Heimliche Blic      |   |

# CSV bestehend aus Songtexten und Metadaten

\* Korpus ist auf Anfrage per Mail erhltlich

# 3 Korpus-Beschreibung & Erste Analysen

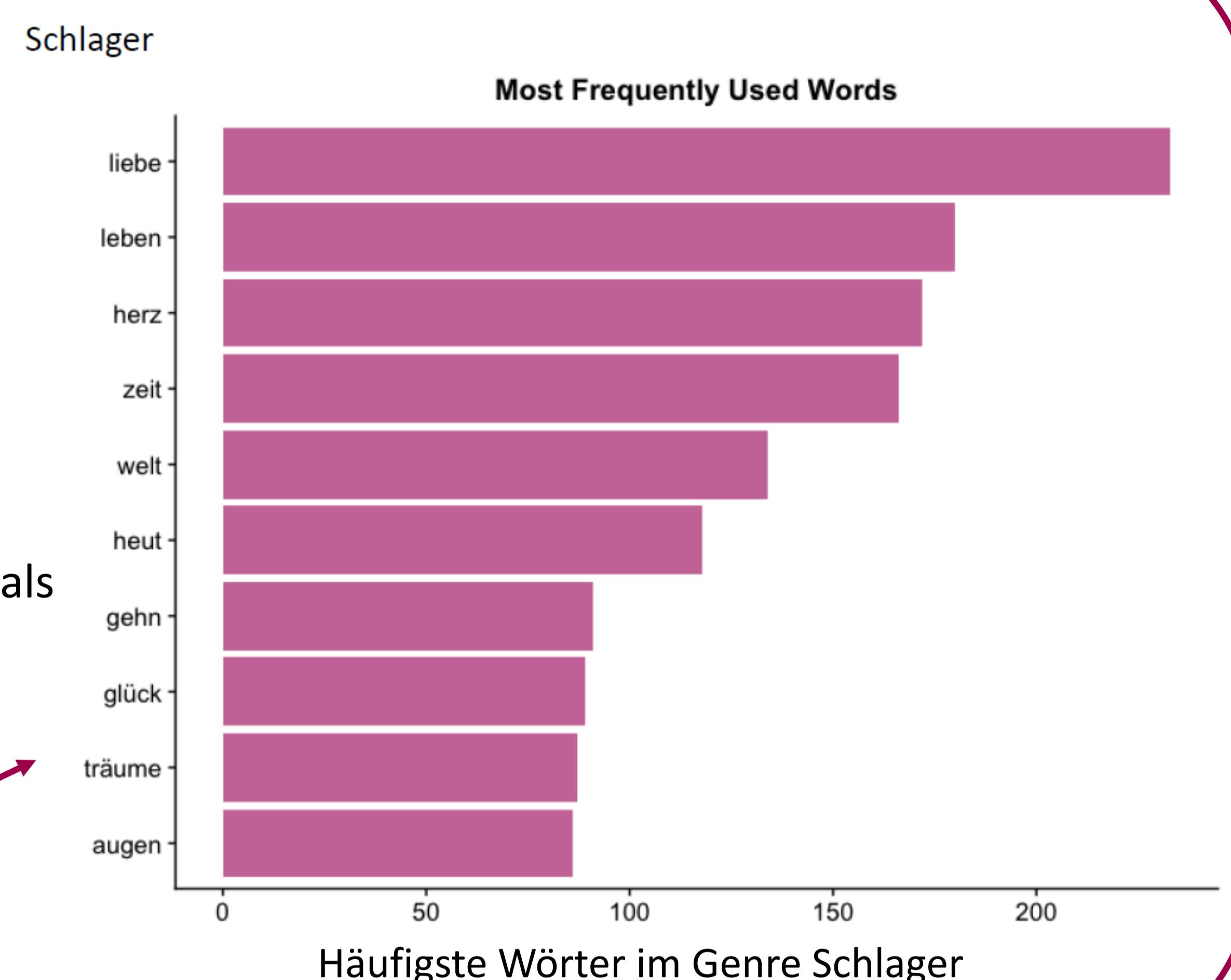
| Genre    | Künstler | Albums | Songs | Tokens  | Bekannte Vertreter                           |
|----------|----------|--------|-------|---------|--|
| Pop      | 22       | 96     | 1132  | 302614  | Nena, Rosenstolz, Herbert Grönemeyer         |
| Rap      | 33       | 129    | 1558  | 864925  | Die fantastischen Vier, Samy Deluxe, Sido    |
| Rock     | 20       | 126    | 1312  | 320751  | Udo Lindenberg, Die Ärzte, Rammstein         |
| Schlager | 16       | 83     | 634   | 147833  | Peter Maffay, Wolfgang Petry, Helene Fischer |
| Gesamt   | 91       | 434    | 4636  | 1636123 |  |

# Korpus-Beschreibung

|             | <b>Pop</b> | <b>Rap</b> | <b>Rock</b> | <b>Schlager</b> | <b>All</b> |
|-------------|------------|------------|-------------|-----------------|------------|
| <b>60er</b> | -          | -          | -           | 10              | 10         |
| <b>70er</b> | 10         | -          | 51          | 41              | 102        |
| <b>80er</b> | 101        | -          | 132         | 85              | 318        |
| <b>90er</b> | 98         | 108        | 194         | 141             | 541        |
| <b>00er</b> | 459        | 694        | 632         | 264             | 2049       |
| <b>10er</b> | 463        | 756        | 303         | 94              | 1616       |
| <b>All</b>  | 1132       | 1558       | 1312        | 634             | 4636       |

# Songverteilung per Jahrzehnt und Genre

- Mehrheitlich homogene Wortverteilungen über die Genres hinweg: Welt, Leben, Zeit als häufigste Wörter
  - Ausnahmen:
    - Rap: Umgangs- und Jugendsprache, thematische Schwerpunkte
    - Schlager: Emotionale Terme



# 4 Lexikonbasierte Sentiment Analysis

| Genre    | Häufigste positive Wörter (Häufigkeit)                                  | Häufigste negative Wörter (Häufigkeit)                                     | Gesamt-Polarität | Gesamt-Polarität normalisiert an der Anzahl der Tokens |
|----------|---|--|------------------|--|
| Pop      | liebe (628)<br>schön (255)<br>lieben (116)<br>rein (198)                | schwer (186)<br>kurz (108)<br>feuer (104)<br>kleine (87)<br>fallen (85)    | -1373.681        | -0.004539  |
| Rap      | liebe (627)<br>rein (399)<br>lieber (377)<br>schön (257)<br>reich (207) | scheiß (718)<br>scheiße (427)<br>hart (317)<br>alter (271)<br>schwer (232) | -363.398         | -0.0004201   |
| Rock     | liebe (579)<br>schön (349)<br>lieber (210)<br>lieb (150)<br>rein (146)  | schwer (144)<br>scheiß (115)<br>feuer (112)<br>kurz (108)<br>kalt (108)    | -424.503         | -0.0013234   |
| Schlager | liebe (547)<br>nah (115)<br>lieb (108)<br>lieben (104)<br>schön (102)   | feuer (114)<br>arm (78)<br>schwer (78)<br>wein (55)<br>fehlt (50)          | -344.082         | -0.0023275   |

# Häufigste Terme für das Topic „Liebe“ in allen Genres

# 5 Topic Modeling

