

Der Einsatz von Distant Reading auf einem Korpus deutschsprachiger Songtexte

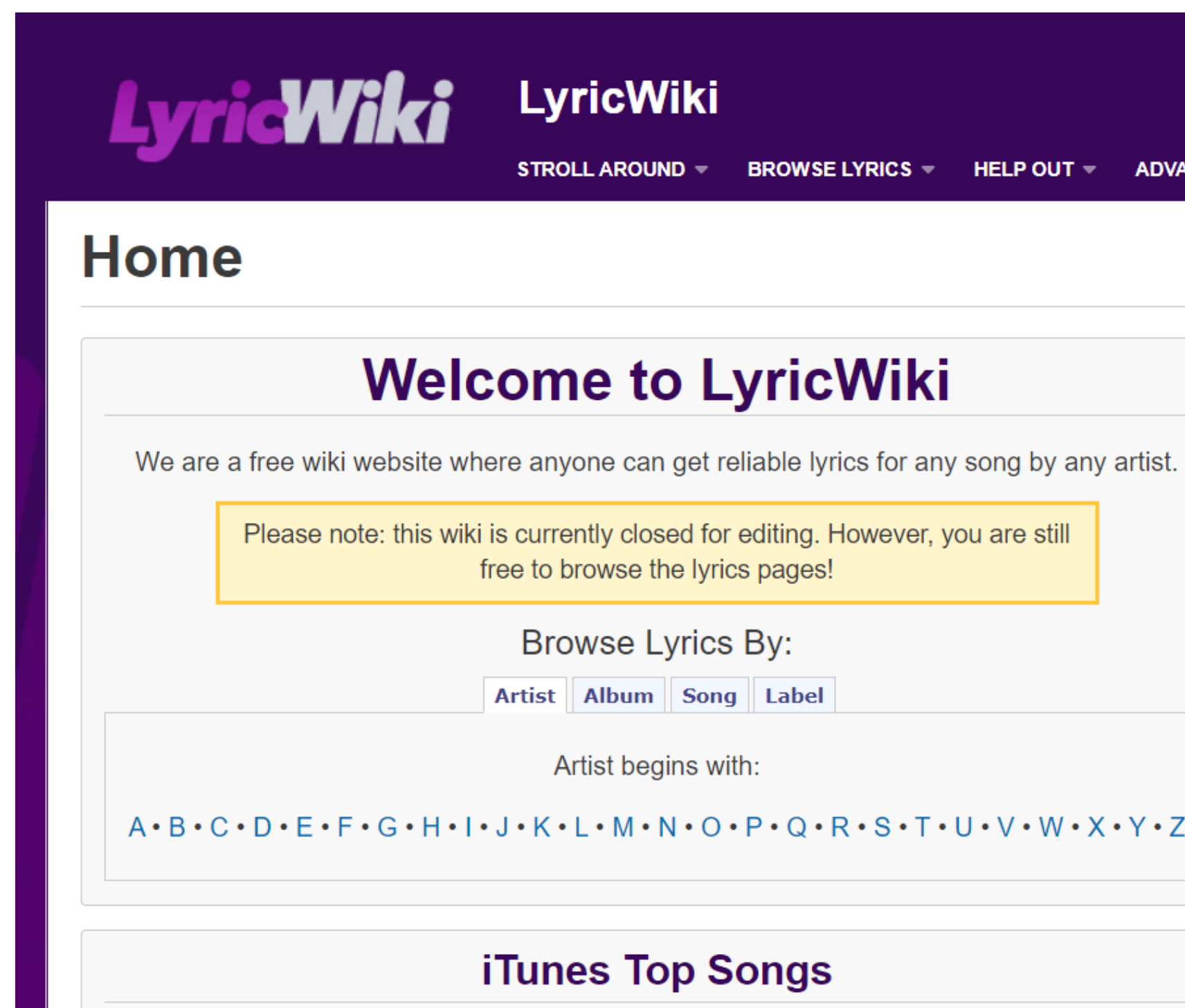
Thomas Schmidt, Marlene Bauer, Florian Habler, Hannes Heuberger, Florian Pils, Christian Wolff
Media Informatics Group, Universität Regensburg

DHd 2020

Kontakt: Thomas.Schmidt@ur.de



1 Korpus-Akquise



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 alle:	
Pop	Namika	Que Walou	2018	Parkbank	Sie hat nie die weite Welt geseh	
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 besuch	
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 mir	
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 sinc	
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	
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? D	
Pop	Nena	Nena	1983	Zaubertrick	Heimliche Blicke Heimliche Blic	

Scraping nach vordefinierter Liste von Künstlern auf der Plattform LyricWiki

CSV bestehend aus Songtexten und Metadaten

* Korpus ist auf Anfrage per Mail erhältlich

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	

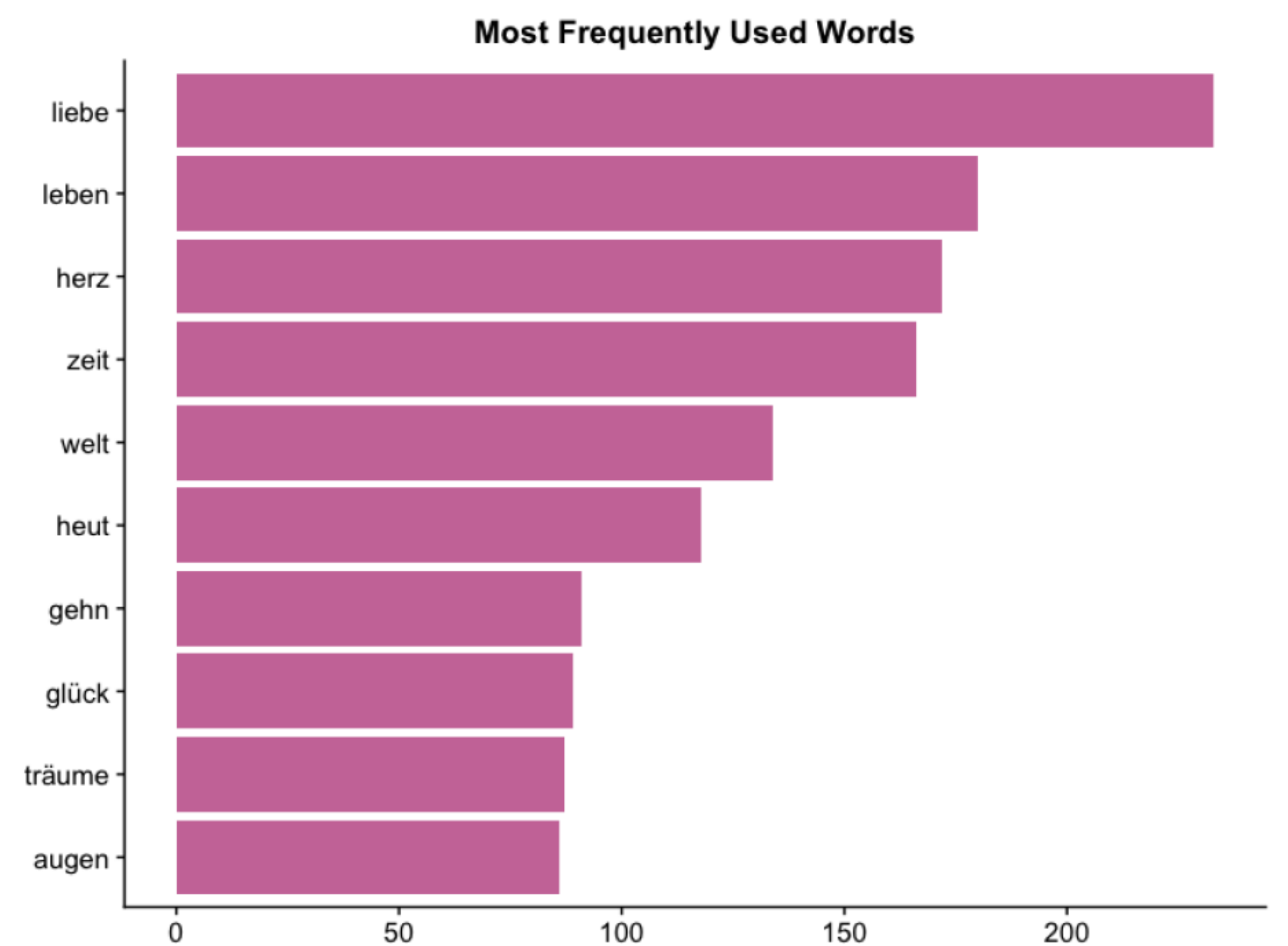
Korpus-Beschreibung

	Pop	Rap	Rock	Schlager	All
60er	-	-	-	10	10
70er	10	-	51	41	102
80er	101	-	132	85	318
90er	98	108	194	141	541
00er	459	694	632	264	2049
10er	463	756	303	94	1616
All	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

Schlager



Häufigste Wörter im Genre Schlager

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) schön (255) lieben (116) rein (198)	schwer (186) kurz (108) feuer (104) kleine (87) fallen (85)	-1373.681	-0.004539
Rap	liebe (627) rein (399) lieber (377) schön (257) reich (207)	scheiß (718) scheiße (427) hart (317) alter (271) schwer (232)	-363.398	-0.0004201
Rock	liebe (579) schön (349) lieber (210) lieb (150) rein (146)	schwer (144) scheiß (115) feuer (112) kurz (108) kalt (108)	-424.503	-0.0013234
Schlager	liebe (547) nah (115) lieb (108) lieben (104) schön (102)	feuer (114) arm (78) schwer (78) wein (55) fehlt (50)	-344.082	-0.0023275

5 Topic Modeling

