

Von DRAMALYS.EXE zu DraCor

Eine Geschichte der digitalen Dramenanalyse von 1984 bis heute



Prof. Dr. Frank Fischer

(Freie Universität Berlin)

Kurzlink zu dieser Präsentation: bit.ly/geschdra

Ringvorlesung des #RosDH · Universität Rostock · 2. Mai 2024, 17:15 Uhr

Programmankündigung:

- Nach formalistischen und strukturalistischen Vorarbeiten werden Dramentexte spätestens seit den 1980er-Jahren mit Computern erforscht. Der Vortrag lässt 40 Jahre digitale Dramenforschung Revue passieren, von den ersten Experimenten mit imperativen Programmiersprachen bis hin zu ausgewachsenen Infrastrukturen in den digitalen Ökosystemen der Gegenwart.

Inhalt

1. Vor 1984
2. 1984
3. Nach 1984
4. Zusammenfassung

Vor 1984

»Data« as of 1819

- »Thus having been provided in the fewest words, in a natural reply to as natural a question, – which yet answers a secondary purpose of attracting our attention to the difference or diversity between the characters of Cornwall and Albany, – provided **the premises and data**, as it were, having been thus afforded **for our after-insight into the mind and mood of the person, whose character, passions, and sufferings are the main subject-matter of the play**; – from Lear, the persona patiens of his drama, Shakespeare passes without delay to the second in importance, the chief agent and prime mover, and introduces Edmund to our acquaintance, preparing us with the same felicity of judgment, and in the same easy, natural way, for his character in the seemingly casual communication of its origin and occasion.«

S. T. Coleridge: Thursday, 28 January 1819 (King Lear). In: Adam Roberts (ed.): Coleridge: Lectures on Shakespeare (1811–1819). Edinburgh University Press 2016, pp. 167–174.

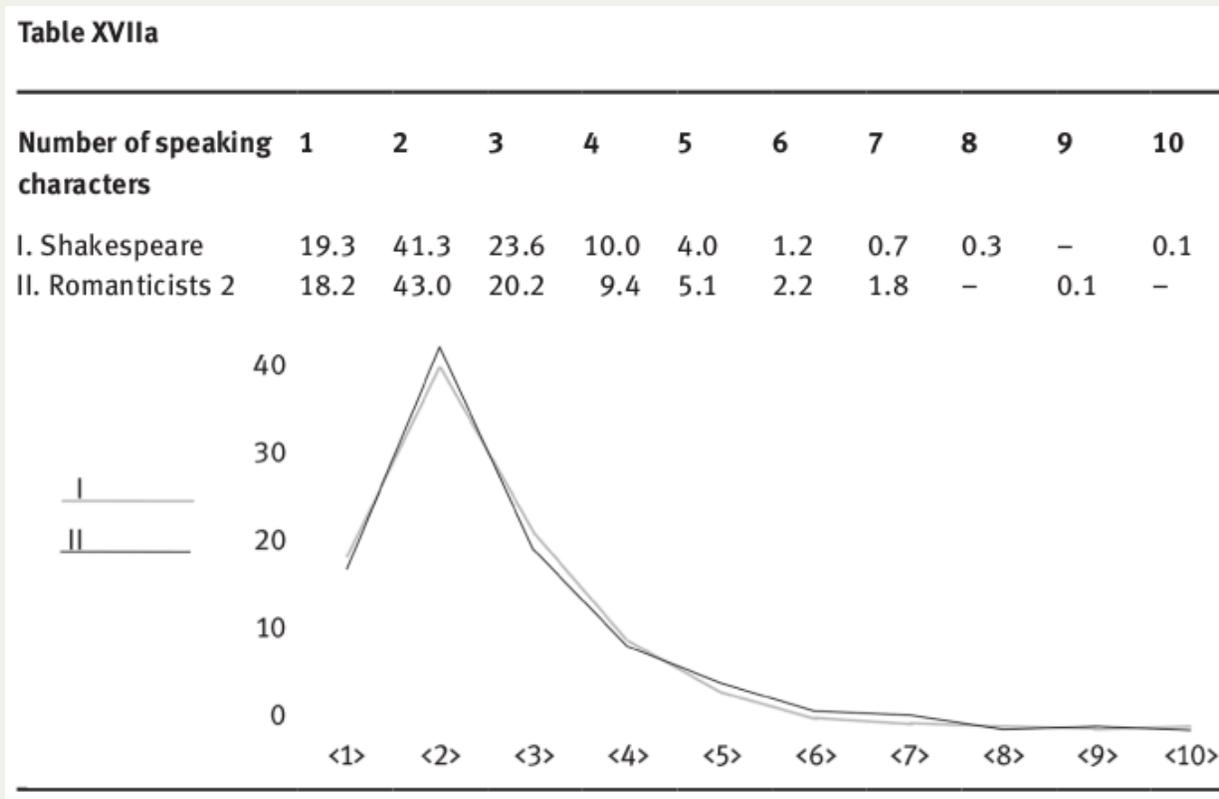
(Ein Fund von Nan Z. Da, vgl. ihren Vortrag »Literary-Critical Complexity« auf dem 2nd Workshop »Mathematics of Small Data Analysis« am Zuse-Institut Berlin, 18. Januar 2024. 🙏)

Der Moskauer Formalismus und Boris Yarkho (1889–1942)



(Bildquelle: urokiistorii.ru – via Internet Archive)

Boris Yarkho: *Speech Distribution in Five-Act Tragedies (A Question of Classicism and Romanticism)* (verfasst 1935–1938, Erstdruck 1997)



Shakespeare im Vergleich mit dem romantischen Drama (im Durchschnitt): »romantic drama is a return to Shakespeare«. – Englische Übersetzung von Craig Saunders in: Frank Fischer, Marina Akimova, Boris Orekhov (Hrsg.): *Journal of Literary Theory*. Vol. 13. No. 1: Moscow Formalism and Literary History. De Gruyter 2019, S. 13–76. (doi:10.1515/jlt-2019-0002)

»Computational thinking« bei Boris Yarkho

- »Eine große Unannehmlichkeit der von uns vorgeschlagenen komparativ-statistischen Methode ist die Unwahrscheinlichkeit, dass eine Person jemals in der Lage wäre, eine ausreichende Anzahl von Merkmalen der Klassik und der Romantik zu studieren, wenn sie dies nicht zu ihrem gesamten Lebenswerk machen will.«
- »Die Berechnung [eines bestimmten Koeffizienten] würde jedoch über 150 Berechnungen von Standardabweichungen erfordern. Da wir dafür nicht die entsprechenden Mitarbeiter zur Verfügung hatten, hätten wir damit mehr Zeit verloren, als es wert gewesen wäre.«
- »Wir haben die Technik demonstriert; der Weg ist frei für diejenigen, die mehr Glück haben als wir.«

(Übers. F. F.)

Solomon Marcus: »Poetica matematică« (1970)

(dt. »Mathematische Poetik«. Aus d. Rumän. übertr. von Edith Mândroiu. Frankfurt/M.: Athenäum 1973)

Inhalt	15
VII. Kapitel: Vergleichende Untersuchungen von poetischen Texten	
Der Binärcode und das vergleichende Studium der Gedichte	239
Poetische Segmente und Hammingdistanz	253
Eine Anwendung auf ein Gedicht von Baudelaire	254
Der Zusammenhang der grammatikalischen Kategorien	262
Der zeitliche Zusammenhang in „Familiale“ von Prévert	263
Schwacher Zusammenhang in bezug auf eine Wertemenge	264
Zusammenhängende Komponenten	266
Die schwach zusammenhängenden Komponenten der Varianten des Gedichtes „Mai am un singur dor“	267
Die Struktur des Zusammenhanges und der Konnotativitätsgrad	268
Die für eine bestimmte Kategorie bezeichnendste Variante	269
Drei Arten Isomorphismen	271
Die Aufeinanderfolge	273
Eine Abhängigkeitsbeziehung	276
Chromatische Zahl, Spaltung, Durchmesser und deren literarische Bedeutung	280
Andere Methoden	282
Literatur	284
VIII. Kapitel: Mathematische Methoden im Theaterstudium	
Einige Vorläufer	287
Ein origineller Zuschauer	289
Die Bevölkerungsdichte der Szene	292
Arten der Beziehungen zwischen den Charakteren	293
Szenischer Abstand und szenischer Durchmesser	295
Ausmaß der szenischen Gegenüberstellung	298
Eine Hammingdistanz	300
Ein Maß für die szenische Nähe	301
Ein Beispiel anhand des „Verlorenen Briefes“	301
Eine Rangordnung der Szenen eines Stückes	304
Eine neue Betrachtung des Parameters Y	306
Anzahl der Szenen, in denen nur ein einziger Charakter auftritt	307
Ein neues Maß für die Beziehungen zwischen den Charakteren	308
Die Nähe der Charaktere im „Verlorenen Brief“	310
Einige Bedeutungen	311
16	
Mathematische Poetik	
Eine Übertragung der Methode auf alle möglichen Konfigurationen von Charakteren	315
Kerne eines Stückes	320
Kerne im „Verlorenen Brief“	321
Kerne in „Letzte Nachrichten“ und „Starke Seelen“	324
Eine Analyse der Kerne in Schillers „Räubern“	326
Einige allgemeine Sätze über Kerne	333
Kurzer Überblick über die Resultate von Polti und Souriau	338
Die geometrischen Schemen von Ginstier	341
Dreiecksituationen	343
Vierecksituationen	345
Diagonalgeometrie	346
Situationen mit halboffener Geometrie	346
Kombinationen geschlossener Geometrien	347
Geometrien zweiten Grades	348
Die dramatische Form bei Steen Jansen	350
Textliche Ebene und szenische Ebene	352
Die Grenze zwischen zwei Situationen	354
Die Rangordnung der Charaktere nach Jansen	356
Ein vollständiges mathematisches Dramenmodell	357
Erläuterung des Modells	358
Korrelationen von Orten und Charakteren	360
Irreduzible Konfigurationen	362
Domination zwischen Konfigurationen	363
Leere Konfigurationen	366
Die einem Drama zugeordnete Sprache	366
Literatur	369
ANHANG	
Mathematische Begriffe und Ergebnisse	
Zusätzliche Literatur	381
Namenverzeichnis	427

Sehr instruktiv in Sachen Operationalisierung, z. B. (sinngemäß): »Zwei Figuren interagieren miteinander, wenn sie innerhalb desselben Abschnitts eines Dramas (normalerweise eine »Szene«) einen Sprechakt vollziehen.«

Strukturalismus

Manfred Pfister: »Das Drama. Theorie und Analyse« (1977)

	1	2	3	4	5	6	7	8	9	10
DIRNE	1	0	0	0	0	0	0	0	0	1
SOLDAT	1	1	0	0	0	0	0	0	0	0
STUBENMÄDCHEN	0	1	1	0	0	0	0	0	0	0
JUNGER HERR	0	0	1	1	0	0	0	0	0	0
JUNGE FRAU	0	0	0	1	1	0	0	0	0	0
EHEMANN	0	0	0	0	1	1	0	0	0	0
SÜSSES MÄDEL	0	0	0	0	0	1	1	0	0	0
DICHTER	0	0	0	0	0	0	1	1	0	0
SCHAUSPIELER	0	0	0	0	0	0	0	1	1	0
GRAF	0	0	0	0	0	0	0	0	1	1

„Der Reigen“, Arthur Schnitzler

›Konfigurationsmatrix‹ zu Schnitzlers »Reigen« (1901). – Bildquelle: DH Regensburg 2015.

1984

Computerized Drama Analysis

H. ILSEMANN

University of Hannover, Germany

Abstract

DRAMALYS.EXE, which was written for MS-DOS Computers in Turbo-Basic, reads through a drama text and, together with a character list, produces the configuration structure of the play. Entries, exits and speeches of figures are accounted for in the form of a matrix of grey, light grey and dark boxes thus giving an impressive visual shape to the play's structure. Duration and size of configurations, techniques of combining configurations in the play, and aspects of segmentation and composition can be studied in the file 'playname.TEX' that can be printed on endless paper. Simultaneously every single speech by a figure is being monitored. Length of speech in characters and words, number of speech, relation between primary and secondary text, and total words are counted. The program then summarizes those data which describe figure relations

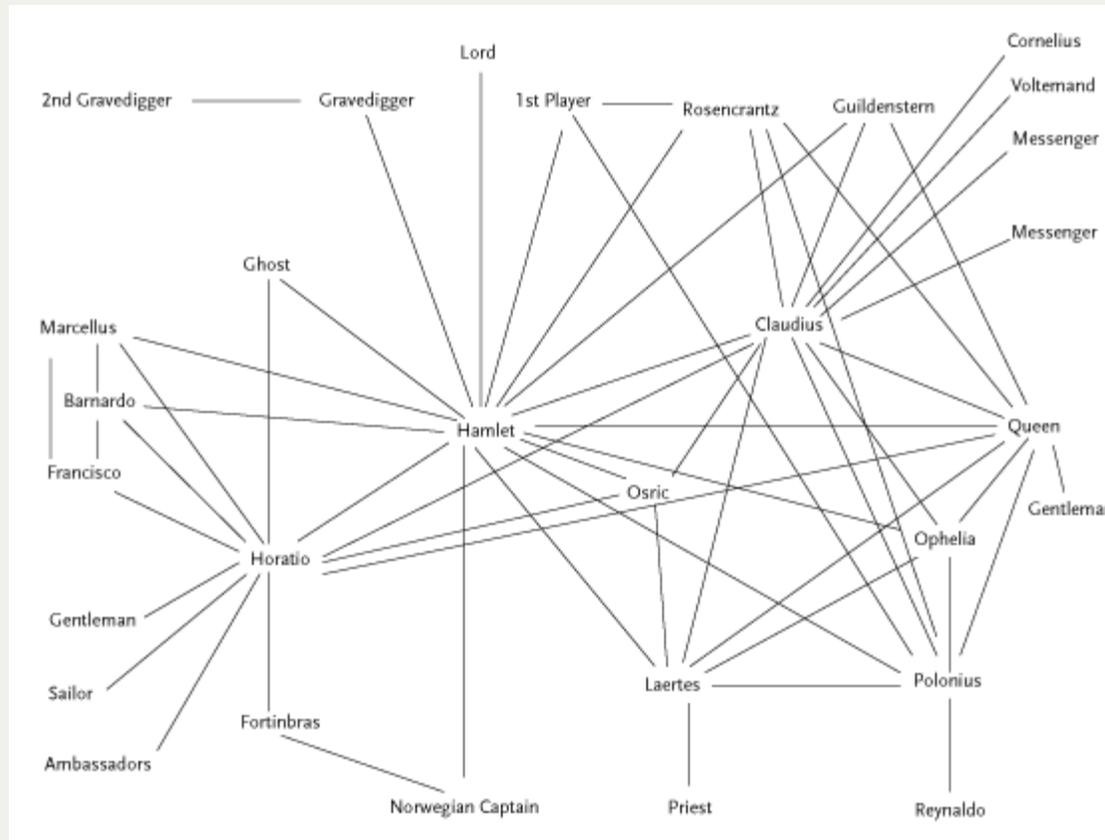
about the same time, dealt with stylistic types which he established from word and syllable lengths of various prose texts by writers as diverse as Goethe and Adenauer. Both approaches can be put to some use in the analysis of drama texts which are now being made subject to statistical investigations carried out by a computer program. DRAMALYS was first developed under CP/M in 1984 and was then rewritten in Turbo-Basic for MS-DOS in 1987³. Since then various features have been added, the latest being the graphics program DRAMGRAF⁴ that accounts for the sequential development of a play. This article, in fact, considers the question of how far one can go in the assessment and interpretation of those data that are provided by the programs. Whilst it is the overriding

Nach 1984

Deutschsprachige Dramenkorpora (ab 2007)

- Zeno.org (2007/2008)
 - deutschsprachige Volltextbibliothek, die auch knapp 700 Dramen enthielt (in eigenem XML-Format)
 - 2009 vom Forschungsverbund TextGrid erworben, in TEI konvertiert und unter Creative-Commons-Lizenz zur Verfügung gestellt: Digitale Bibliothek
- außerdem:
 - Deutsches Textarchiv (DTA)
 - Wikisource: 50 Dramen
 - Projekt Gutenberg-DE

Franco Moretti: »Network Theory, Plot Analysis« (2011)



(Quelle: Literary Lab Pamphlet 2, Abb. 2.)

»just think of this: I am discussing *Hamlet*, and saying nothing about Shakespeare's words« (S. 4)

Peer Trilcke: Social Network Analysis als Methode einer textempirischen Literaturwissenschaft (2013)

	Prinz	Kam.	Con.	Mari.	Cam.	Clau.	Pirro	Odo.	Ang.	Emil.	App.	Batt.	Ors.
Prinz	0	2	2	10	1	0	0	2	0	2	0	1	1
Kammerd.	2	0	0	0	0	0	0	0	0	0	0	0	0
Conti	2	0	0	0	0	0	0	0	0	0	0	0	0
Marinelli	10	0	0	0	0	3	1	4	1	3	2	4	4
Camillo Rota	1	0	0	0	0	0	0	0	0	0	0	0	0
Claudia G.	0	0	0	3	0	0	4	3	0	3	4	1	1
Pirro	0	0	0	1	0	4	0	2	1	0	1	0	0
Odoardo G.	2	0	0	4	0	3	2	0	0	2	0	0	2
Angelo	0	0	0	1	0	0	1	0	0	0	0	0	0
Emilia G.	2	0	0	3	0	3	0	2	0	0	1	1	0
Appiani	0	0	0	2	0	4	1	0	0	1	0	0	0
Battista	1	0	0	4	0	1	0	0	0	1	0	0	0
Orsina	1	0	0	4	0	1	0	2	0	0	0	0	0

Fig. 20

Die Visualisierung und Auswertung der Matrix bringt folgende Ergebnisse.

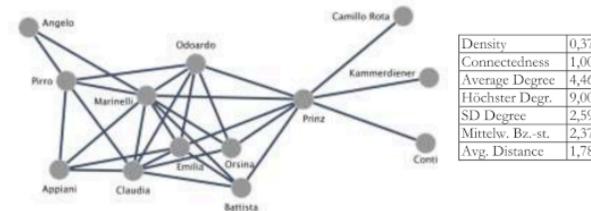


Fig. 21

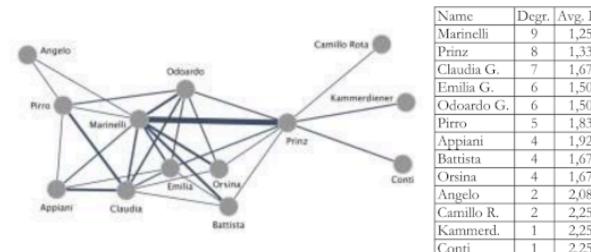
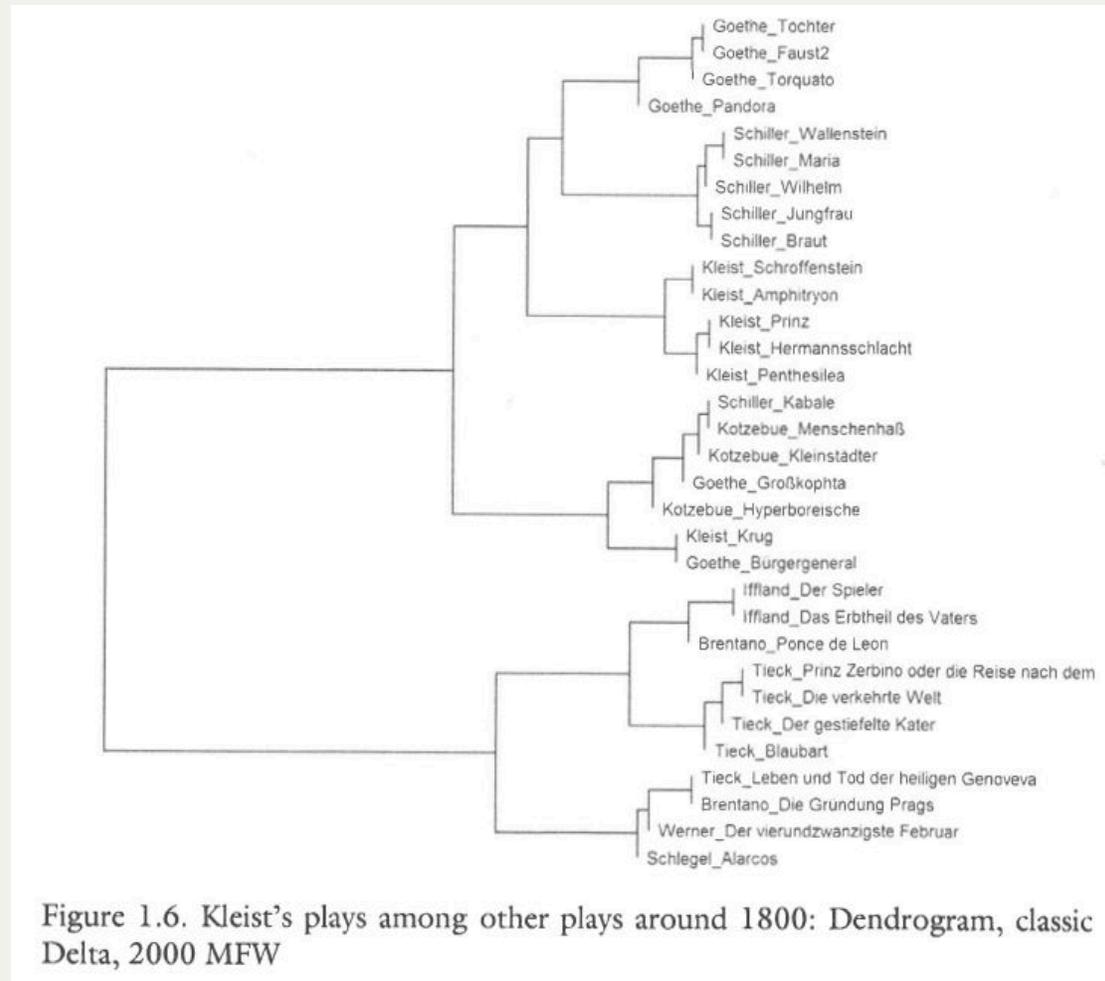


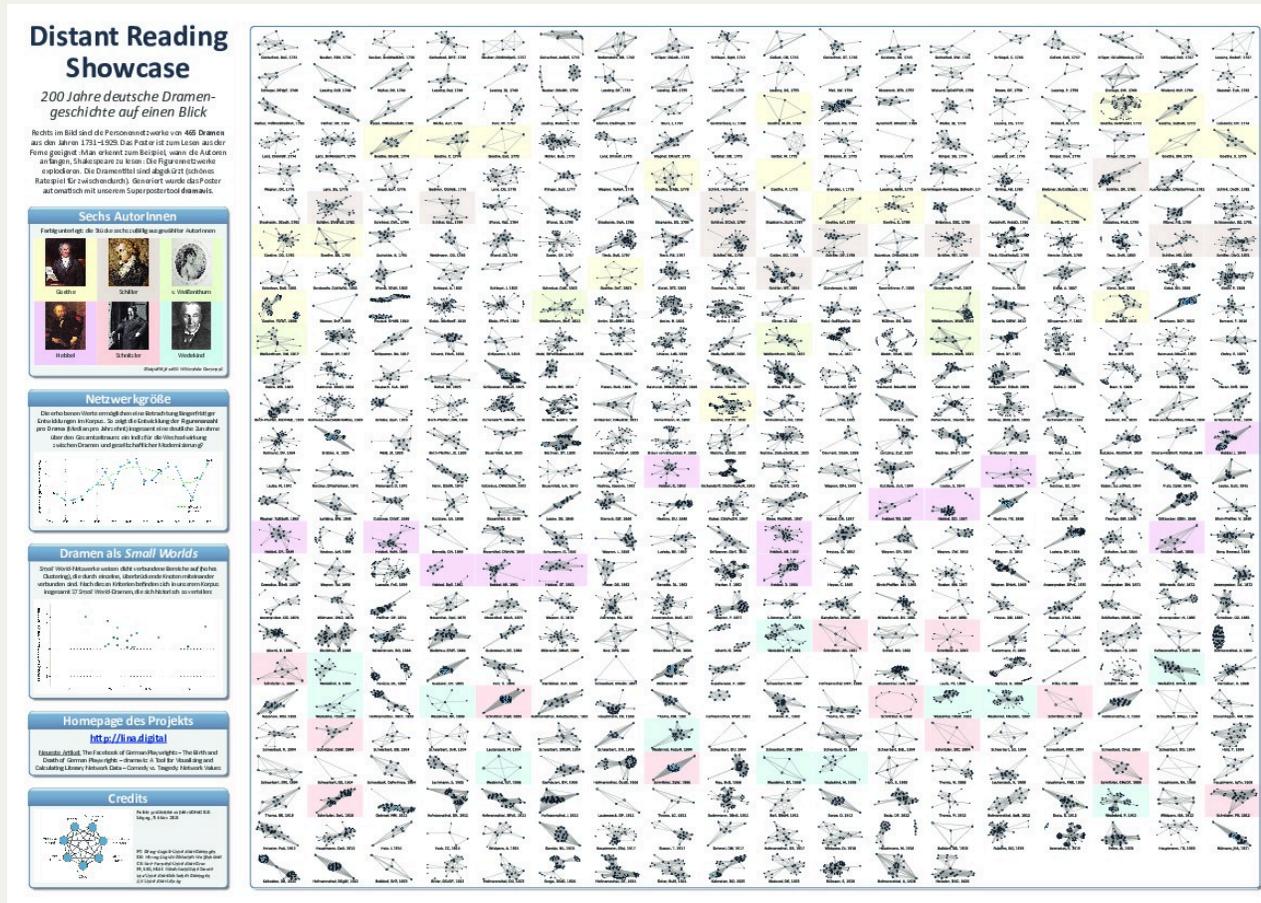
Fig. 22

Stilometrie: Heinrich von Kleist zwischen Klassik und Romantik (2014)



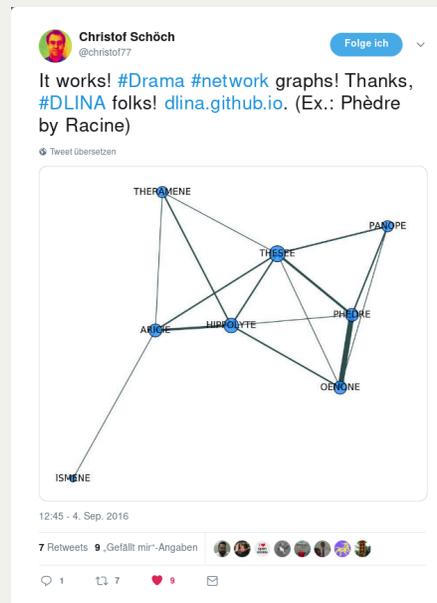
Distant-Reading Showcase (2016)

Wie verändern sich die Figurenkonstellationen im deutschen Drama?



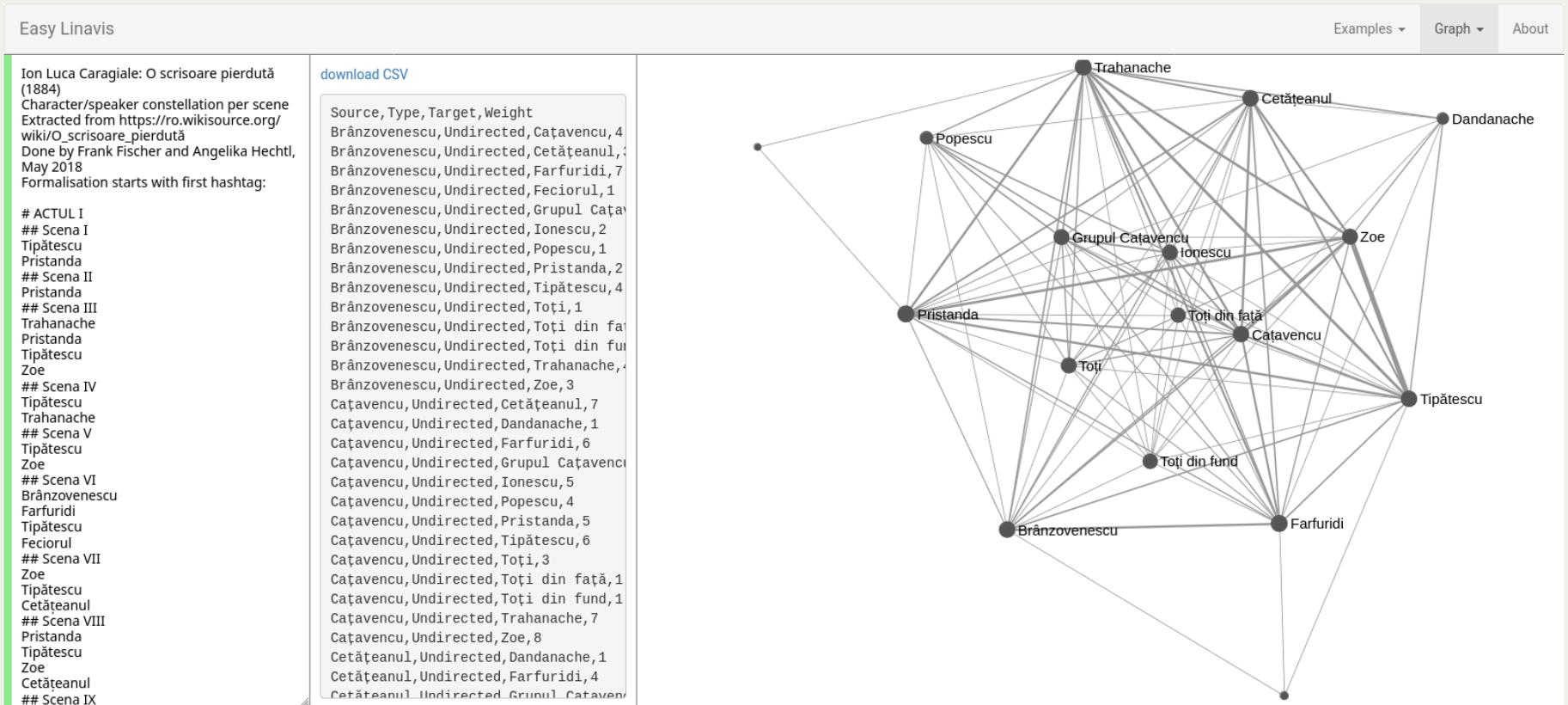
Kittel/Fischer: Dramavis (2014–2017)

- das in Python geschriebene Programm war u. a. Basis für den »Distant-Reading Showcase«
- Startpunkt: DARIAH International Digital Humanities Summer School im August 2014 in Göttingen
- Open Source, Adaptionen? – Wenige...



<https://twitter.com/christof77/status/772370040016568320>

Fischer/Milling: ezlinavis (2017)

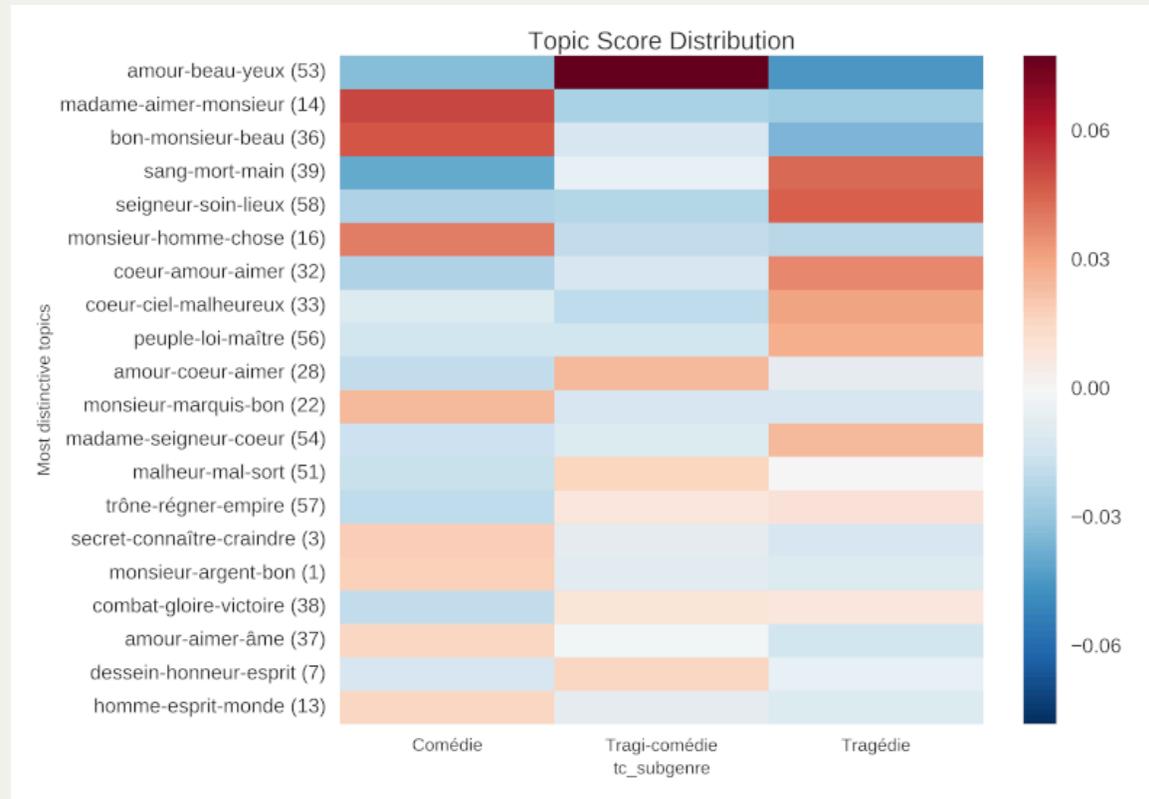


Eine der Beispieldateien: Ion Luca Caragiale: »O scrisoare pierdută« (1884),
das auch Solomon Marcus als Showcase diente.

NB.

- ezlinavis kam u. a. auch in der medizinischen Forschung zum Einsatz:
 - [doi:10.1007/s00432-022-04200-0](https://doi.org/10.1007/s00432-022-04200-0)
 - [doi:10.21873/invivo.12976](https://doi.org/10.21873/invivo.12976)

Christof Schöch: Topic Modeling Genre: An Exploration of French Classical and Enlightenment Drama (DHQ 2017)



Korpus: 391 Stücke (via Théâtre classique von Paul Fièvre).

Mark Algee-Hewitt: Distributed Character: Quantitative Models of the English Stage, 1550–1900 (NLH 2017)

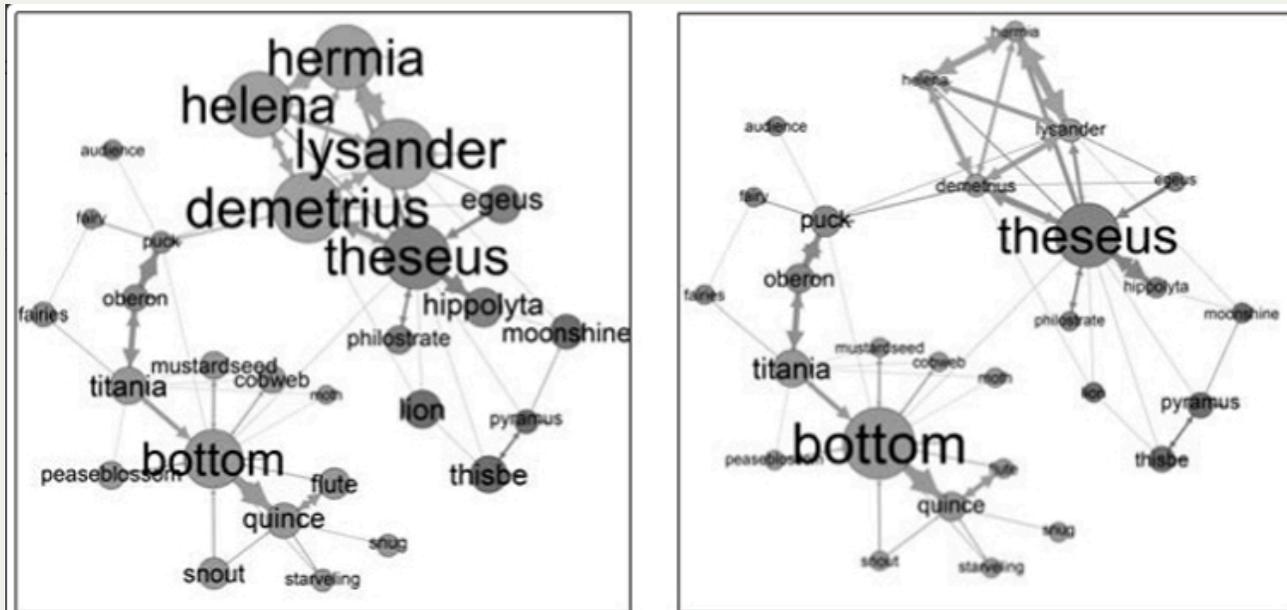


Fig 2a and 2b.

a) network of *A Midsummer Night's Dream* with nodes sized for eigenvector centrality; b) network of *A Midsummer Night's Dream* with nodes sized for betweenness centrality.

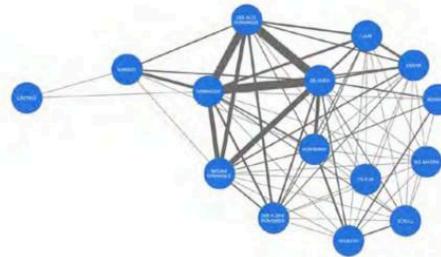
Image 2 of 11

Korpus: 3.568 Stücke (via ProQuest).

Fischer/Trilcke et al.: To Catch a Protagonist (1730–1930) (DH 2018)

Figur	Countbasiert			Netzwerkbasieret					Aggregierte Maße		Gesamt-Rank
	Häufigkeit	Sprechakte	Wörter	Degree	Weighted Degree	Betweenness	Closeness	Eigenvector	Count	Netzwerk	
DELOMER	1	1	2	2	1	2	2	1	1	2	1
DOMINIQUE	2	2	1	1	2	1	1	2	2	1	1
HOFMANN	5	5	5	3	6	3	3	7	4	4	3
MME DOMINIQUE	4	7	9	3	4	4	3	4	6	3	4
GRAP	6	4	7	5	5	6	5	5	5	5	5
ALTE DOMINIQUE	3	3	3	10	3	11	10	3	3	8	6
GRÄFIN	8	6	6	5	8	6	5	8	6	7	7
KL. DOMINIQUE	8	12	12	5	7	6	5	6	10	6	8
MARQUIS	8	8	4	13	10	5	13	9	6	11	9
NTIRATH	7	9	8	8	9	9	8	10	9	9	10
SCHULZ	12	11	10	8	11	9	8	11	12	10	11
BUDENTER	11	10	11	14	12	12	14	12	10	14	12
DIE BAUERN	13	13	15	10	13	12	10	14	14	12	13
ETUCHE	13	14	14	12	14	12	12	13	14	13	14
GÄRTNER	13	14	13	15	15	12	15	15	13	15	15

Rankings for Iffland's "Das Erbtheil des Vaters" (1802) – Spearmans Rho: 0.806.



Network graph for Iffland's "Das Erbtheil des Vaters" (1802).

Operationalisierung verschiedener »quantitativer Dominanzrelationen« (Pfister).

Trilcke/Fischer:
Literaturwissenschaft als Hackathon
(ZfdG 2018)

- »Dass die Arbeit in der ›digitalen Literaturwissenschaft‹ wesentlich dadurch geprägt ist, dass digitale Methoden eingesetzt werden, die im Methodenrepertoire der ›traditionellen Literaturwissenschaft‹ bisher *nicht* vorkommen, versteht sich von selbst. [...] So selbstverständlich diese Differenz in den Praktiken aber sein mag, so folgenreich ist sie zugleich. Nur ein Gedankenexperiment: Selbst eine exzellent ausgebildete Literaturwissenschaftlerin, ein exzellent ausgebildeter Literaturwissenschaftler hätte, wäre sie oder er zufällig in unserem temporären Potsdamer Labor aufgetaucht, vermutlich nichts von dem verstanden, was wir da taten.«

Infrastrukturelle Aspekte

- Programmiersprachen (Basic, R, Python, ...)
- Libraries (stylo, igrph, networkx, spaCy, nltk, gensim, ...)
- digitale Korpora (mit oder ohne wissenschaftliches Markup)

Was ist eigentlich mit ...

- komparatistischen Ansätzen?
- FAIR-Prinzipien? (findable, accessible, interoperable, reusable?)
- Community-Aspekten? Training?
- Linked Open Data?

DLS → CLS

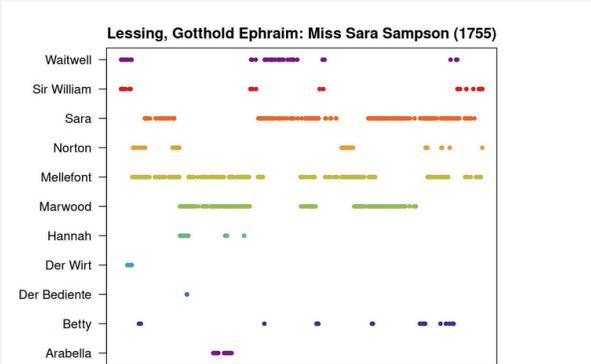
- Digital Humanities als polysemer Begriff (nach Camille Roth: Digital, digitized, and numerical humanities. In: DSH 34.3 (2019), S. 616–632, doi:10.1093/llc/fqy057)
 - **digitized**: Aufbau, Verwaltung und Verarbeitung digitalisierter Archive
 - **numerical**: mathematische Abstraktion geisteswissenschaftlicher Inhalte und Entwicklung numerischer und formaler Modelle
 - **humanities of the digital**: Untersuchung computervermittelter Interaktionen und Online-Communitys
- **DLS**: vom als Quarterly angekündigten Journal »Digital Literary Studies« von James O’Sullivan erschien 2016 nur eine Ausgabe, danach eingestellt
- **CLS**: DFG-Schwerpunktprogramm 2207 »Computational Literary Studies« (2019–), Horizon-2020-Projekt »CLS INFRA« (2021–2025), »Journal of Computational Literary Studies« (2022–)

Nils Reiter, Tim Strohmayer, Janis Pagel: DramaAnalysis (2019)

7.1 When are characters talking?

While the above displays the *length* of utterances, we can also display the position of utterances (remember the column `utteranceBegin`?). The following snippet visualizes when characters are talking, this time for Lessings Miss Sara Sampson:

```
par(mar=c(2,7,2,2))  
  
utteranceStatistics(rjmw.0) %>%  
  characterNames(rjmw.0) %>% # character names instead of ids  
  plot(main=dramaNames(rjmw.0)) # calling plot.QDUtteranceStatistics()
```



Each dot in this plot represents one utterance, the x-axis is measured in character positions. This is not really intuitive, but the flow from left to right represents the flow of the text. More technically, we again apply the function `characterNames()` to display character names instead of character ids. This is the same function as above, just applied to a different table. It can be applied to any table of the type `QDHasCharacter`. Secondly, the call to the function `plot()` gets rerouted to the function `plot.QDUtteranceStatistics()`, because the object we supply as argument is of the type `QDUtteranceStatistics`. Information about this function can be retrieved by entering `?plot.QDUtteranceStatistics`.

Screenshot aus dem Tutorial.

DraCor (2017–)

Welcome

German Drama Corpus

Number of plays	471
person	9,838 (male: 5540, female: 1724)
text	8,282,696 tokens
sp	327,903 (7,916,866 tokens)
stage	148,436 (944,423 tokens)
Last update	28.2.2019, 00:41:44

Russian Drama Corpus

Number of plays	144
person	2,481 (male: 1819, female: 575)
text	1,512,139 tokens
sp	75,457 (1,429,988 tokens)
stage	32,734 (144,438 tokens)
Last update	27.2.2019, 23:30:26

Shakespeare Drama Corpus

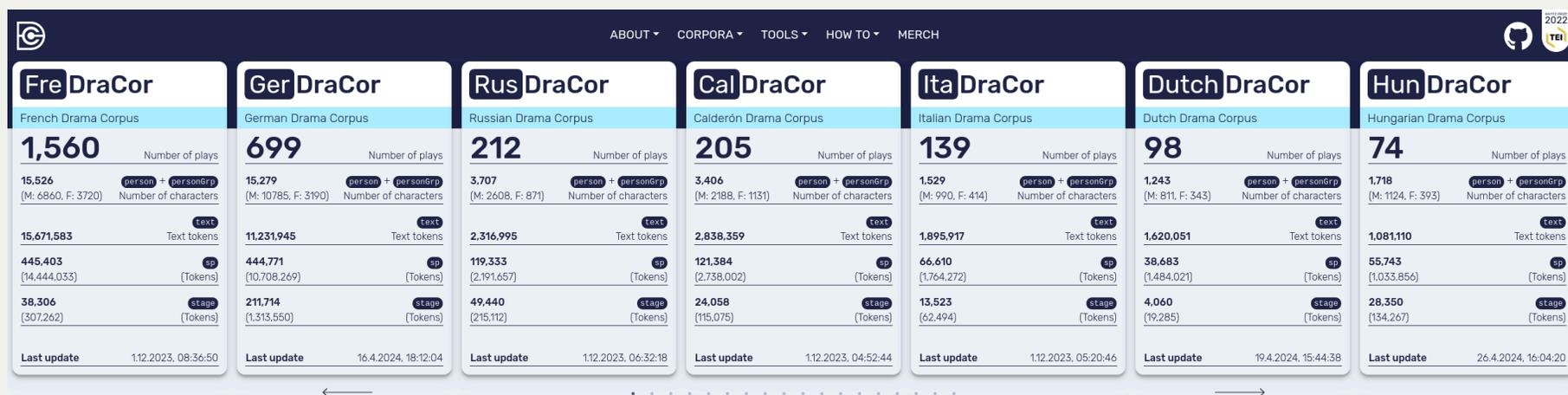
Number of plays	37
person	1,433 (male: 0, female: 0)
text	908,286 tokens
sp	31,066 (876,744 tokens)
stage	10,450 (41,230 tokens)
Last update	26.2.2019, 00:03:21



Our two in-house TEI corpora (RusDraCor and GerDraCor) hosted on dracor.org are in public-alpha state. Feel free to use them, but there are some issues that still have to be resolved before the official release, which might happen in early 2019. - ShakeDraCor was derived from the [Shakespeare Folger Library](#). - dracor.org is edited by [Frank Fischer](#) (Higher School of Economics, Moscow) and [Peer Trilcke](#) (University of Potsdam).

DraCor (Stand 2. Mai 2024)

(Design: Mark Schwindt)



If you want to cite DraCor, please use the following reference:

  Fischer, Frank, et al. (2019). Programmable Corpora: Introducing DraCor, an Infrastructure for the Research on European Drama. In *Proceedings of DH2019: "Complexities"*, Utrecht University, doi:10.5281/zenodo.4284002.

Drama Corpora Project

Unless otherwise stated, all corpora and the web design are released under Creative Commons [0 1.0](#)

[DraCor Frontend 2.2.0](#) [DraCor API 1.0.2](#) [eXist-db 6.2.0](#)

<https://dracor.org/>



DraCor auf verschiedenen Endgeräten (Mockup).

DraCor

- *Drama Corpora* (-Projekt/-Portal/-Plattform)
- Beginn im Jahr 2017 mit zwei Hauskorpora:
 - German Drama Corpus, GerDraCor (1500er–1950er)
 - Russian Drama Corpus, RusDraCor (1740er–1940er)
- Credits-Seite: > 50 Namen

Voraussetzung: Kodierungsstandard TEI (1/2)

Selden. Wie viele Buchstaben hat denn
das Alphabet der Katzen-Sprache?

Anton. Das will ich Ihnen erklären,
dennder Herr Sprachmeister hat es mich schon
gelehrt. Es hat nur 4 Buchstaben: M.
J. A. U; das heiße: Miau. Aus diesen
4 Buchstaben formiren sich aber sämtliche
Wörter, und die Verschiedenheit der Töne,

Voraussetzung: Kodierungsstandard TEI (2/2)

```
<sp who="#felden">
  <speaker>Felden.</speaker>
  <p>
    Wie viele Buchstaben hat denn das Alphabet der Katzen-Sprache?
  </p>
</sp>
<sp who="#anton">
  <speaker>Anton.</speaker>
  <p>
    Das will ich Ihnen erklären, denn der Herr Sprachmeister hat es mich
    schon gelehrt. Es hat nur 4 Buchstaben: M. I. A. U; das heißt: Miau.
    Aus diesen 4 Buchstaben formiren sich aber sämtliche Wörter, und die
    Verschiedenheit der Töne, und die Geschwindigkeit des Ausdruckes oder
    das Aushalten der Sylben formirt die Wörter [...]
  </p>
</sp>
```

Dialog aus Karl von Eckartshausens Lustspiel »Der Pudelhund« (1800) in TEI (Text Encoding Initiative), Quelle: DraCor.
Momentan sind im Standard 586 verschiedene Elemente definiert (Stand November 2023).

Wachsende Liste von Korpora aus der Community:

- FreDraCor
- GerDraCor
- RusDraCor
- CalDraCor
- ItaDraCor
- DutchDraCor
- HunDraCor
- SweDraCor
- UDraCor
- GreekDraCor
- GerShDraCor
- ShakeDraCor
- RomDraCor
- AlsDraCor
- SpanDraCor
- PolDraCor
- TatDraCor
- BashDraCor
- ...

GerDraCor

- mittlerweile 699 Stücke (Ergänzungen aus DTA, Wikisource, Gutenberg-DE, Google Books, ...)
- Community-Projekt: Beschreibung und Anwendungsbeispiele bei forTEXT (Hamburg/Darmstadt)
- Nathalie Wiedmer et al.: Annotation von Beziehungen der Figuren untereinander (Familien- und Liebesverhältnisse, dienstliche Abhängigkeiten)

Frontpage GerDraCor



[ABOUT](#) ▾ [CORPORA](#) ▾ [TOOLS](#) ▾ [HOW TO](#) ▾ [MERCH](#)



German Drama Corpus

Edited by Frank Fischer and Peer Trilcke. Features more than 550 German-language plays from the 1650s to the 1940s. For a corpus description and full credits please see the [README on GitHub](#).

Gottsched, Luise

Corpus licensed under [CC0](#)

Download a comprehensive table with metadata on all plays in the corpus: [JSON](#) [CSV](#)

Authors ▲ ▾	Title ▲ ▾	Year (normalized) ▲ ▾ ⓘ	Network Size ▲ ▾	Source ▲ ▾	ID ▲ ▾
Gottsched, Luise Adelgunde Victorie Q65863	Die Pietistery im Fischbein-Rocke oder Die Doctormäßige Frau Q15805796	1736 1736	16	TextGrid Repository TEI version	ger000462
Gottsched, Luise Adelgunde Victorie Q65863	Das Testament Ein deutsches Lustspiel in fünf Aufzügen Q42291271	1745 1745	12	TextGrid Repository TEI version	ger000220
Gottsched, Luise Adelgunde Victorie Q65863	Der Witzling Ein deutsches Nachspiel in einem Aufzuge Q55102908	1745 1745	7	TextGrid Repository TEI version	ger000395

<https://dracor.org/ger>

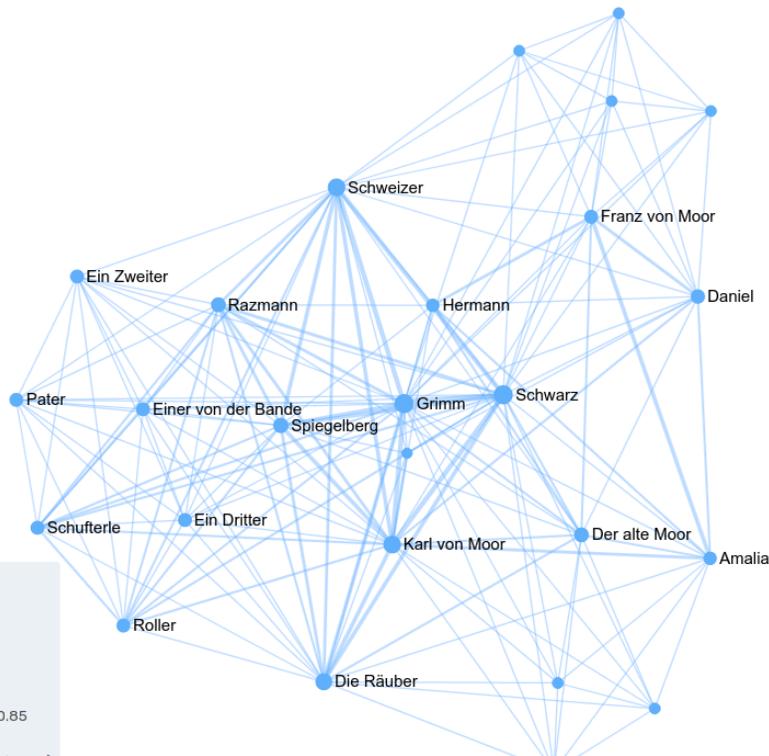
GerDraCor: »Die Räuber« (1781)

GerDraCor

Die Räuber

Schiller, Friedrich

Network Relations Speech distribution Full text Downloads



Cast list

(in order of appearance)

1. Franz von Moor
2. Der alte Moor
3. Karl von Moor
4. Spiegelberg
5. Roller
6. Grimm
7. Schwarz
8. Schufferle
9. Razmann
10. Schweizer
11. Amalia
12. Hermann
13. Daniel
14. Einer von der Bande
15. Ein Zweiter
16. Ein Dritter
17. Die Räuber 🇺🇦
18. Pater
19. Kosinsky
20. Bedienter
21. Moser
22. Bedienter
23. Volk 🇺🇦
24. Ein Räuber
25. Ein alter Räuber
26. Ein dritter Räuber

Segments: 15
All-in at segment 15 (at 100%)
Network size: 26
Density: 0.52
Diameter: 2
Average path length: 1.48
Average clustering coefficient: 0.85
Average degree: 12.92

<https://dracor.org/id/ger000008>

GerDraCor: »Das Testament« (1745)

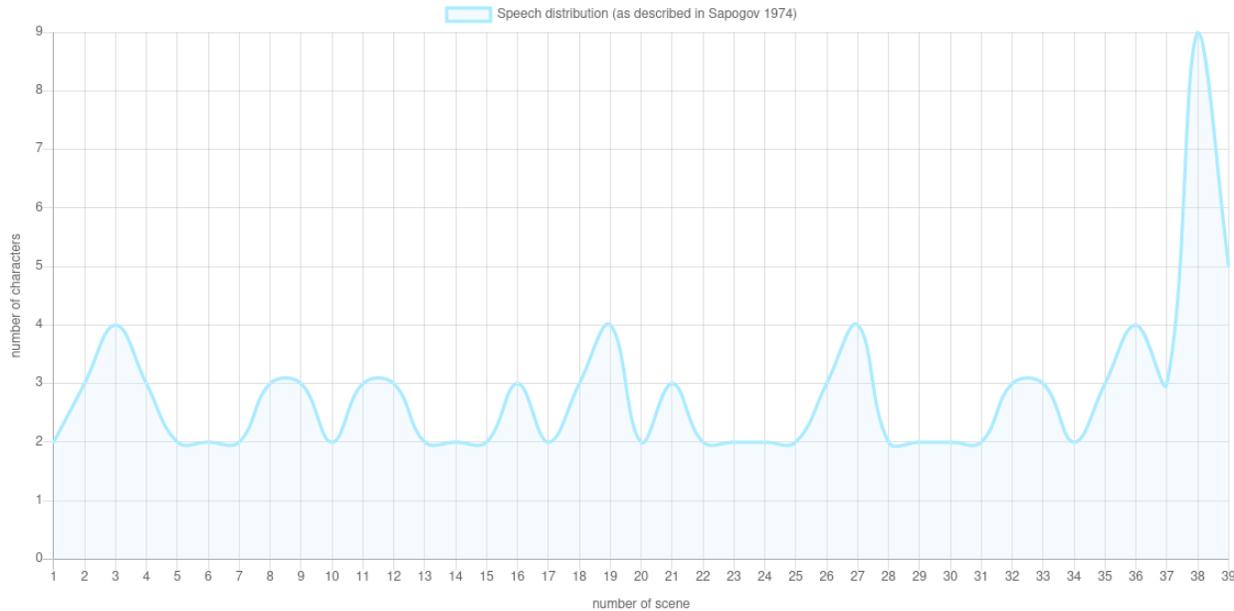
GerDraCor

Das Testament

Gottsched, Luise Adelgunde Victorie

Network Relations Speech distribution Full text Downloads

Sapogov Yarkho Trilcke/Fischer et al.



Cast list

(in order of appearance)

1. Fräulein Amalie
2. Fräulein Karoline
3. Dr. Hippokras
4. Frau von Tiefenborn
5. Heinrich
6. Herr von Kaltenbrunn
7. Herr von Kreuzweg
8. Landrat von Ziegendorf
9. Kammerherr von Ziegendorf
10. Herr von Wagehals
11. Ein Diener
12. Notarius

Cf. Sapogov 1974

<https://dracor.org/id/ger000220#speech>

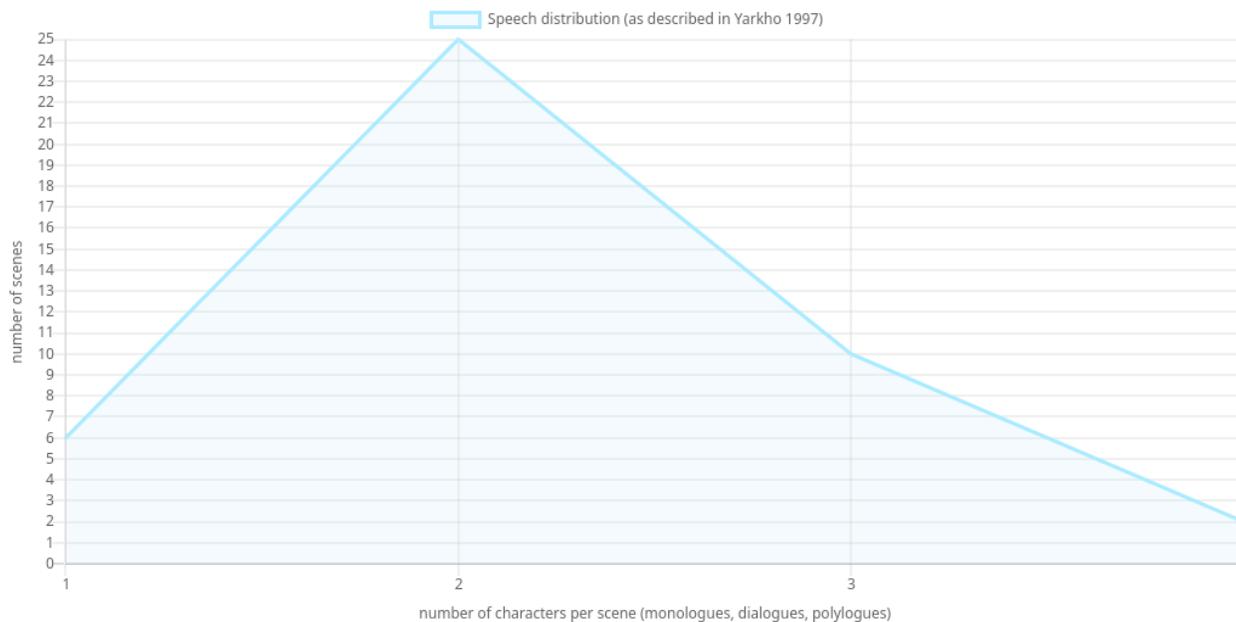
Redeverteilung nach Yarkho

Ger DraCor

Emilia Galotti

Gotthold Ephraim Lessing

Network Relations Speech distribution Full text Downloads



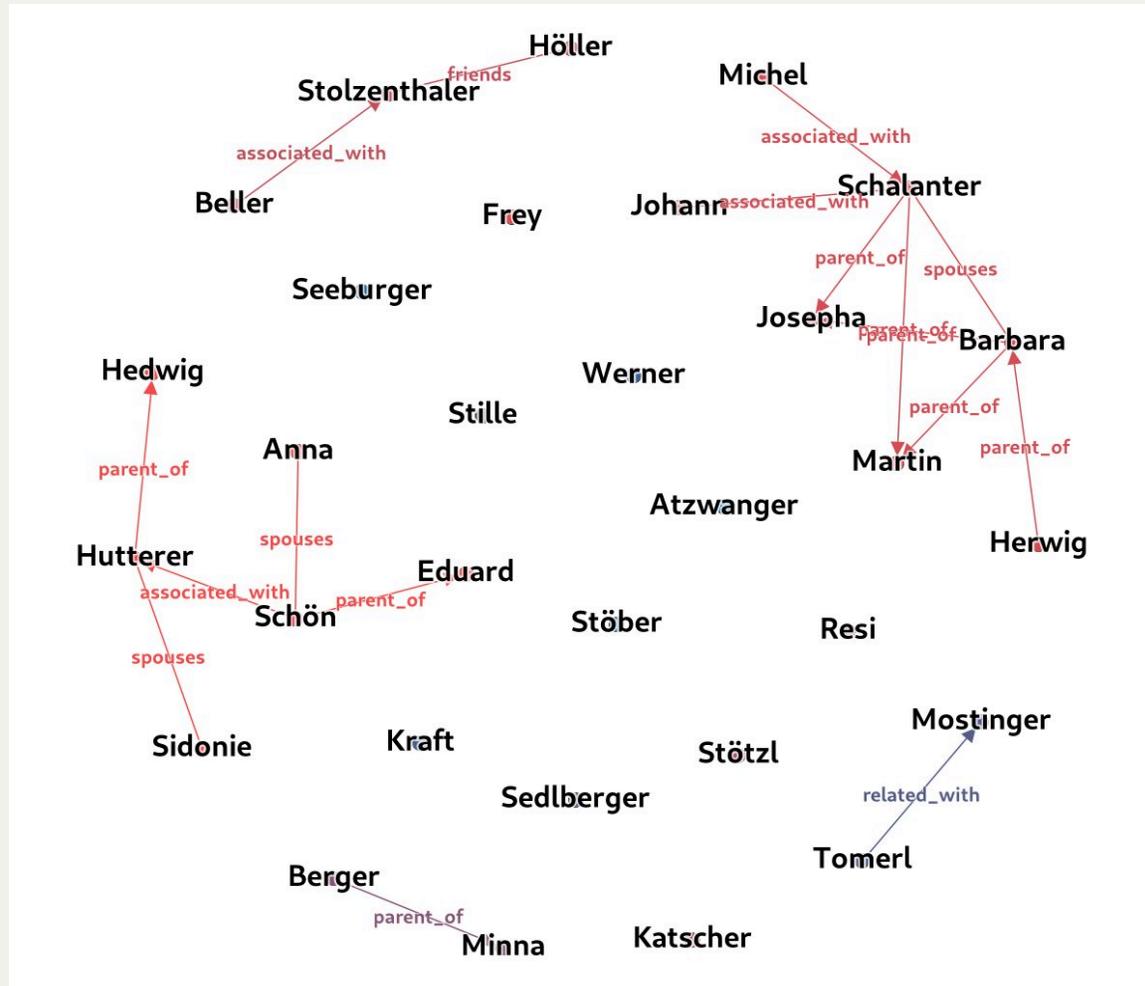
This tab shows different ways of visualising speech distribution.

- Sapogov 1974
- Yarkho 1997 (ru), 2019 (en)
- Trilcke/Fischer et al. 2017

Implementiert in **DraCor** (hier für Lessings „Emilia Galotti“):

<https://dracor.org/ger/lessing-emilia-galotti#speech>

Ludwig Anzengrubers »Das vierte Gebot« (1877)



Modellierung spezifischer Figurenrelationen.

Netzwerkdaten: <https://dracor.org/id/ger000129#relations>

DraCor-Download-Tab



Das vierte Gebot

Anzengruber, Ludwig

Network Relations Speech distribution Full text Downloads

Network data

Co-occurrence network:



Relation data (as described [here](#)):



List of characters

Including precalculated data:



Spoken text

By character:



Plain (no markup):



Full text

TEI-encoded:



Stage directions

Without speaker names:



Including speaker names:



Linked data

In RDF format:



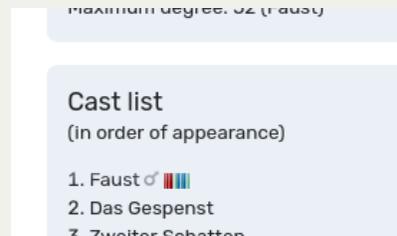
Cast list

(in order of appearance)

1. Schön
2. Anna
3. Hutterer
4. Sidonie
5. Frey
6. Hedwig
7. Eduard
8. Barbara
9. Johann
10. Michel
11. Josepha
12. Stolzenthaler
13. Schalanter
14. Martin
15. Herwig
16. Beller
17. Höller
18. Stötzl
19. Katscher
20. Resi
21. Berger
22. Minna
23. Stille
24. Mostinger
25. Sedlberger

<https://dracor.org/id/ger000129#downloads>

Annotation mythologischer und historischer Figuren



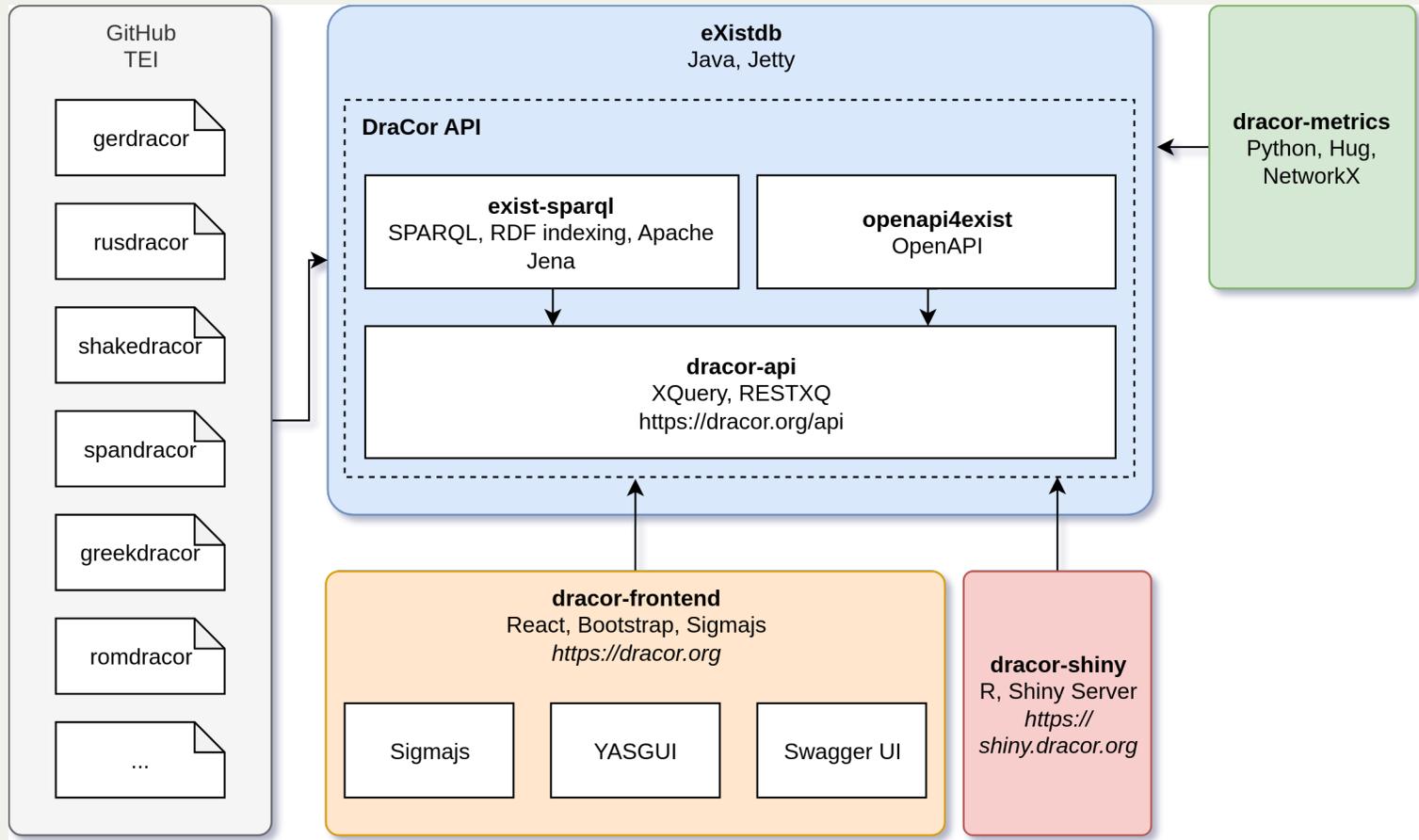
Konzept der ›Programmable Corpora‹

Publikation:

Frank Fischer et al.: **Programmable Corpora: Introducing DraCor, an Infrastructure for the Research on European Drama.** Proceedings of DH2019: »Complexities«, Utrecht, doi:10.5281/zenodo.4284002.

- Korpora als vergleichbare Objekte, die selbst Funktionen anbieten und mit anderen Datenquellen verbunden sind (u. a. via LOD)

DraCor Technology Stack



Alle Repos (Code, Daten) sind *open source*: <https://github.com/dracor-org>

DraCor-API

 ABOUT ▾ CORPORA ▾ TOOLS ▾ HOW TO ▾ MERCH 

DraCor API 0.83.1 OAS3 [/api.yaml](#) [Terms of service](#) [Contact the developer](#) [Apache 2.0](#)

Servers

public ^

- GET /info API info
- GET /corpora List available corpora
- GET /corpora/{corpusname} List corpus content
- GET /corpora/{corpusname}/metadata List of metadata for all plays in a corpus
- GET /corpora/{corpusname}/metadata/csv List of metadata for all plays in a corpus
- GET /corpora/{corpusname}/metadata.csv List of metadata for all plays in a corpus
- GET /corpora/{corpusname}/play/{playname} Get metadata and network metrics for a single play
- GET /corpora/{corpusname}/play/{playname}/metrics Get network metrics for a single play
- GET /corpora/{corpusname}/play/{playname}/tei Get TEI document of a single play
- GET /corpora/{corpusname}/play/{playname}/pdf Get PDF document for a single play

<https://dracor.org/doc/api>

DraCor-API: Beispielanfragen

<https://dracor.org/api/v1/corpora/ita/plays/goldoni-il-servitore-di-due-padroni/spoken-text?gender=FEMALE>

Sprechttext aller weiblicher Figuren in Goldonis »Il servitore di due padroni« (cioè Clarice, Smeraldina e Beatrice).

[https://dracor.org/api/v1/corpora/ger/plays/schiller-die-raeuber/spoken-text?
gender=MALE&relation=parent_of_passive](https://dracor.org/api/v1/corpora/ger/plays/schiller-die-raeuber/spoken-text?gender=MALE&relation=parent_of_passive)

Sprechttext aller Sohn-Figuren in Schillers »Räubern«.

<https://dracor.org/api/v1/corpora/shake/plays/hamlet/metrics>

Netzwerkdaten zu Shakespeares »Hamlet« (im JSON-Format).

<https://dracor.org/api/v1/corpora/ger/plays/gottschedin-die-pietisterey-im-fischbein-rocke/stage-directions>

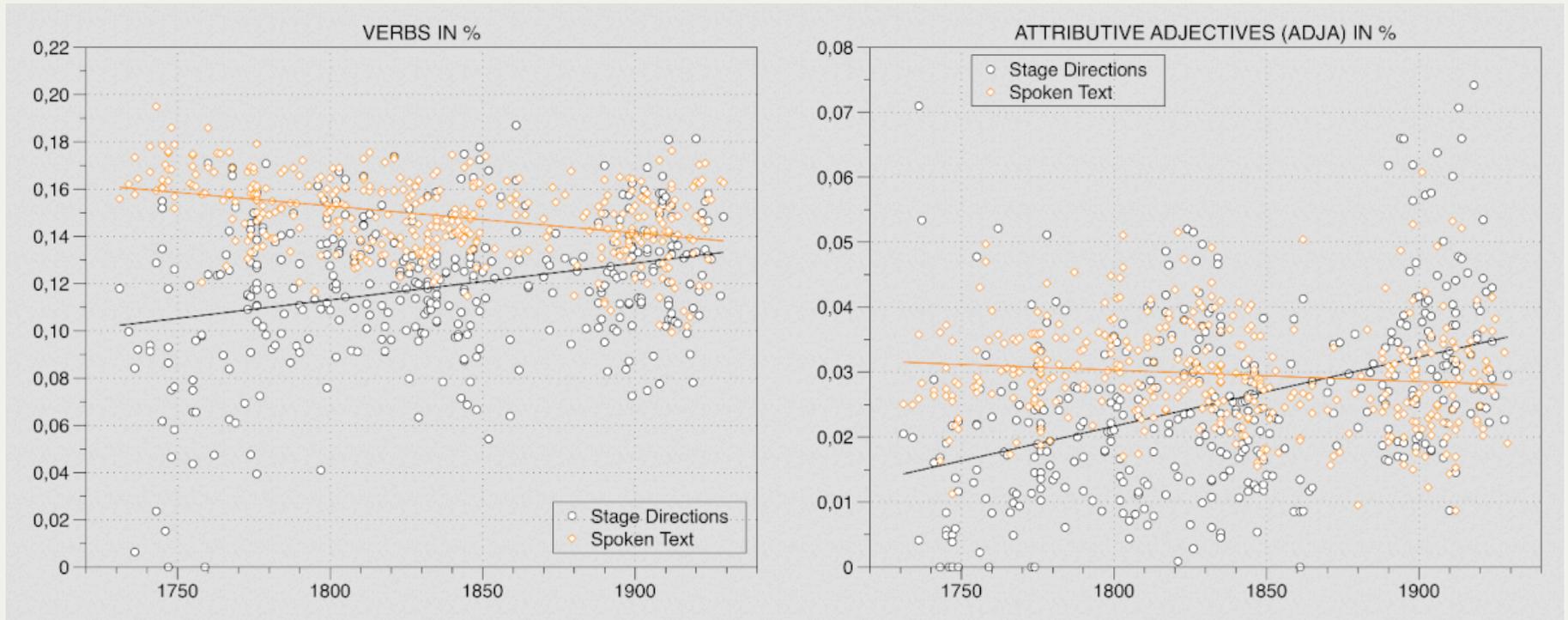
Alle Regieanweisungen aus »Die Pietisterey im Fischbein-Rocke« von Luise Gottsched.

<https://dracor.org/api/v1/corpora/ita/metadata>

Metadaten zu allen Stücken des italienischen Korpus (als JSON oder CSV).

Forschungsbeispiel:

Epifizierungstendenzen im deutschsprachigen Drama,
untersucht anhand von Regieanweisungen



Peer Trilcke, Christopher Kittel, Nils Reiter, Daria Maximova, Frank Fischer:
Opening the Stage: A Quantitative Look at Stage Directions in German Drama.
In: DH2020: »carrefours/intersections«. Book of Abstracts. University of Ottawa.

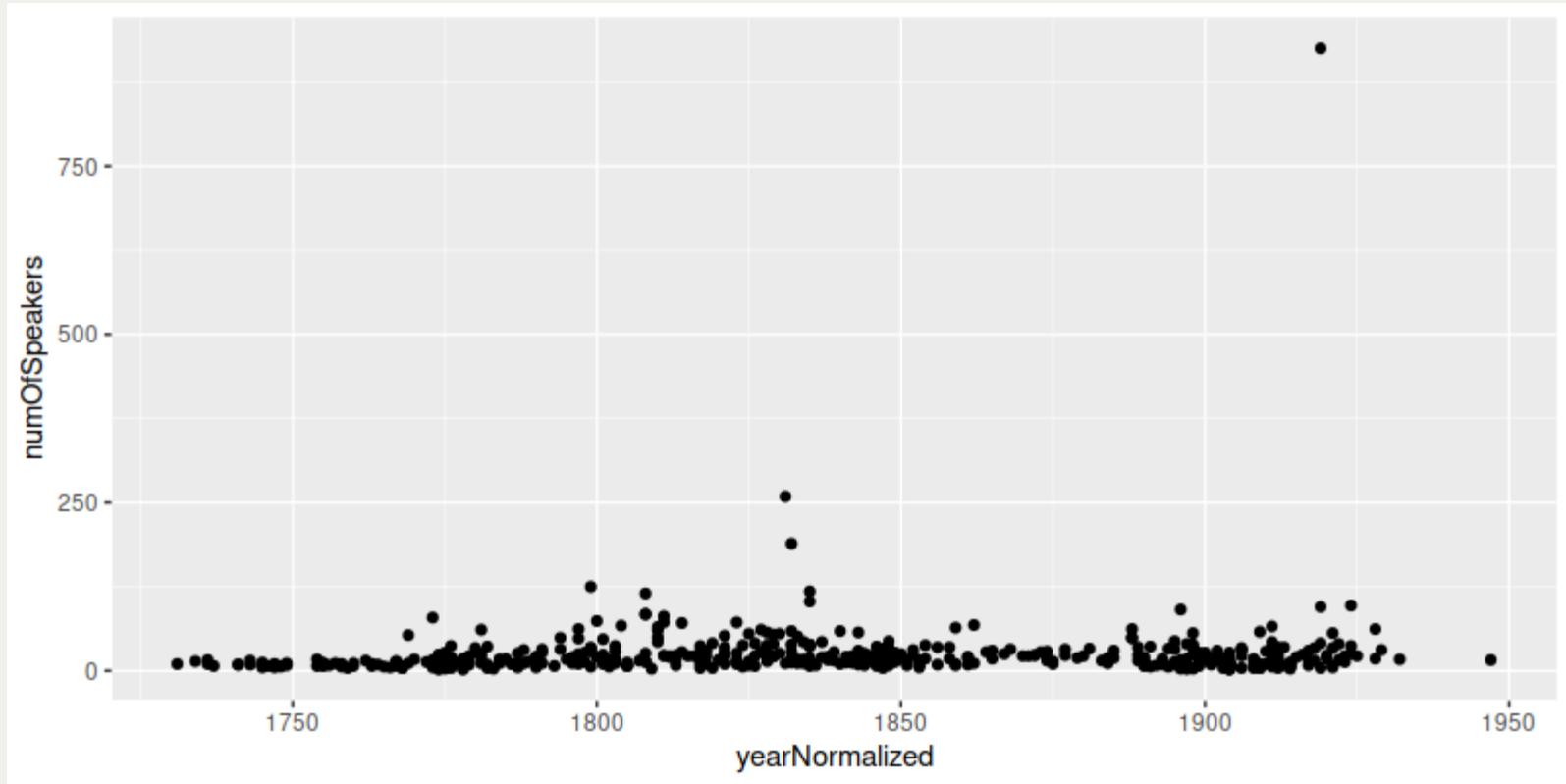
Einfache Visualisierung von Strukturdaten

Anzahl von Figuren pro Stück in chronologischer Folge (1/2)

```
library(data.table)
library(ggplot2)
gerdracor <- fread("https://dracor.org/api/v1/corpora/ger/metadata/csv")
ggplot(gerdracor[, aes(x = yearNormalized, y = numOfSpeakers)) + geom_point()
```

Beispiel für einfaches R-Script, das die Daten über die API bezieht.

Anzahl von Figuren pro Stück in chronologischer Folge (2/2)



Ausgabe in RStudio.

(Der Punkt rechts oben: Karl Kraus' »Die letzten Tage der Menschheit«, 925 Sprechinstanzen).

API-Wrapper

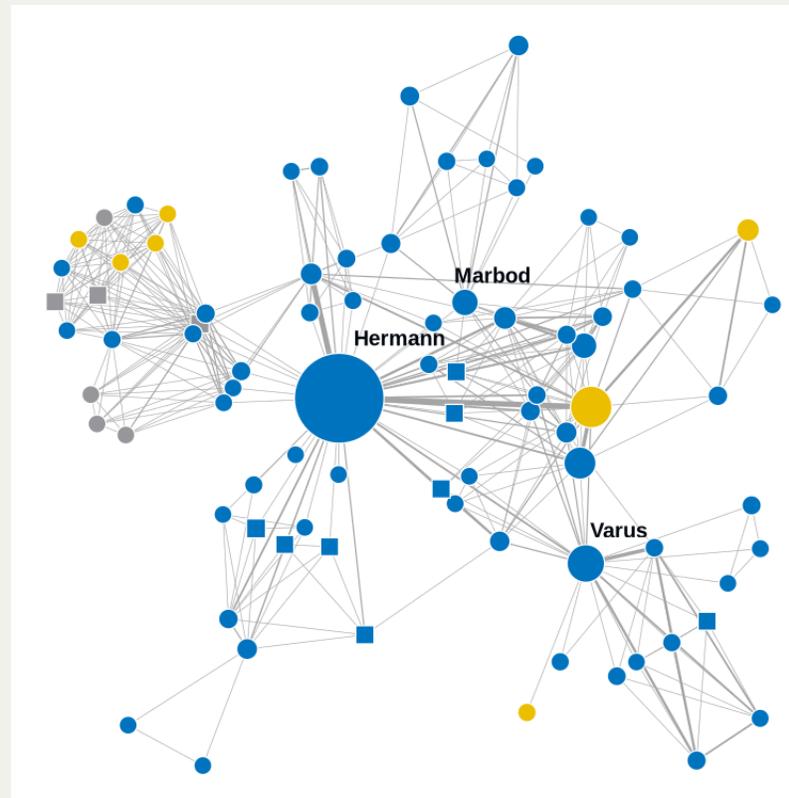
= weitere Vereinfachung:
API-Funktionen bekommen eigene Befehle

- für R: rdracor
- für Python: PyDraCor

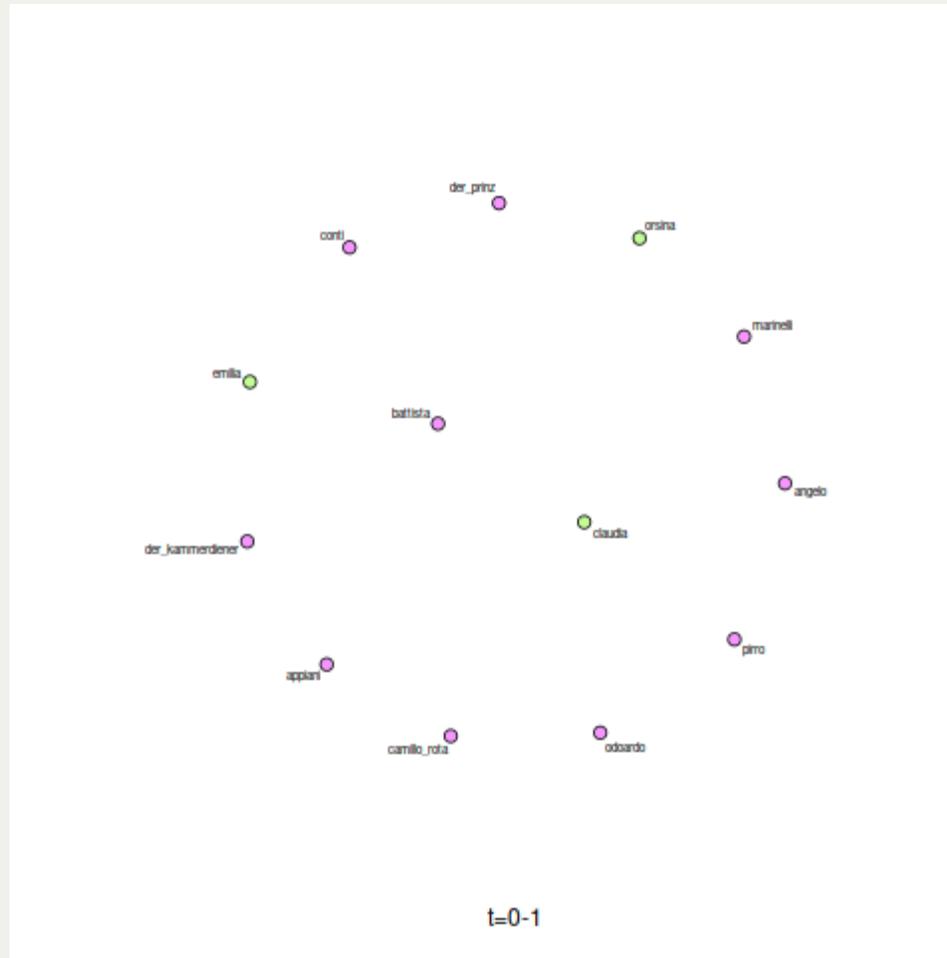
Beispiel (mit rdracor):

```
library(rdracor)
library(igraph)
hermann <- get_net_cooccur_igraph(corpus="ger", play="kleist-die-hermannsschlacht")
plot(hermann)
```

Ergebnis:



Lessings »Emilia Galotti«



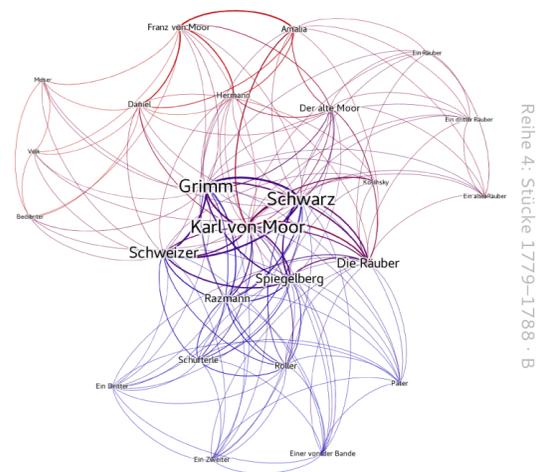
Dynamischer Graph, generiert mit dem **ndtv**-Package. Daten direkt von der **DraCor API**.
Script: Ivan Pozdniakov (Sourcecode auf [Rpubs.com](https://rpubs.com)).

Gamification



Friedrich Schiller
Die Räuber
Ein Schauspiel (1781)

4 B



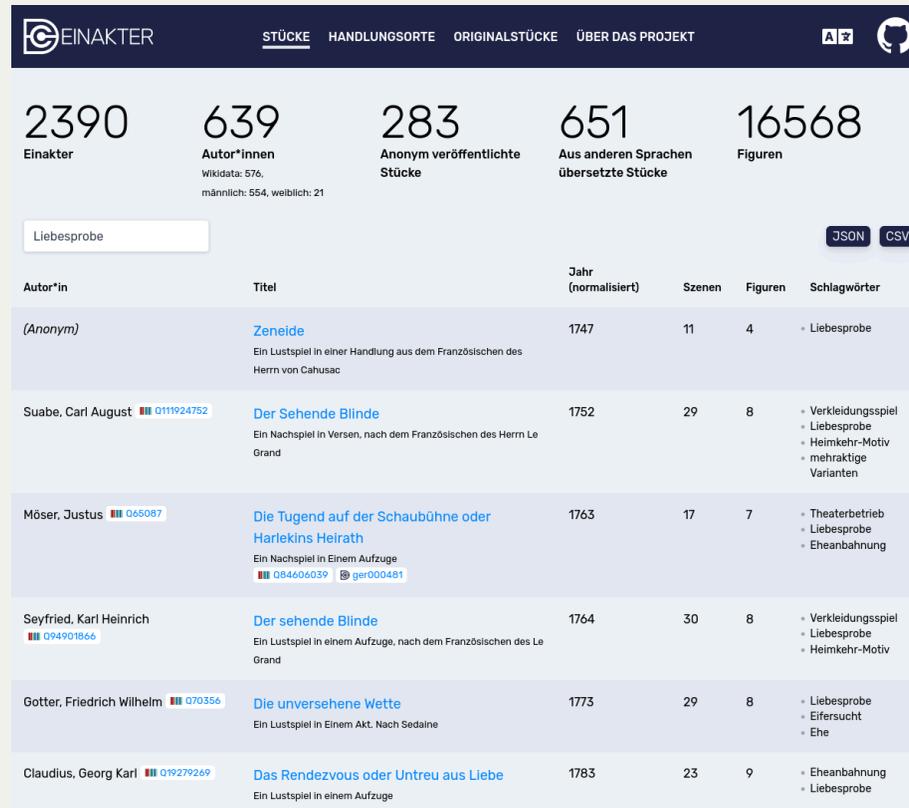
Reihe 4: Stücke 1779–1788 · B

Netzwerkgröße	25
Netzwerkdurchmesser	2
Netzwerkdichte	0,53
Clusterkoeffizient	0,85
Durchschnittliche Pfadlänge	1,47 ↓
Maximaler Grad	24 (Grimm und Schwarz)

Frank Fischer, Anika Schultz: **Dramenquartett. Eine didaktische Intervention.**
Bern: edition taberna kritika 2019. – Spielregeln: <https://dramenquartett.github.io/>

Metadatenzentriert: Einakter-Datenbank (2020)

(<https://einakter.dracor.org/>)



The screenshot shows the Einakter database interface. At the top, there are navigation links: STÜCKE, HANDLUNGSORTE, ORIGINALSTÜCKE, and ÜBER DAS PROJEKT. Below this, five statistics are displayed: 2390 Einakter, 639 Autor*innen (with Wikidata: 576 and gender breakdown: männlich: 554, weiblich: 21), 283 Anonym veröffentlichte Stücke, 651 Aus anderen Sprachen übersetzte Stücke, and 16568 Figuren. A search bar contains 'Liebesprobe'. Below the search bar, there are buttons for 'JSON' and 'CSV'. The main content is a table with columns: Autor*in, Titel, Jahr (normalisiert), Szenen, Figuren, and Schlagwörter. The table lists several plays, including 'Zeneide' by (Anonym), 'Der Sehende Blinde' by Suabe, Carl August, 'Die Tugend auf der Schaubühne oder Harlekins Heirath' by Möser, Justus, 'Der sehende Blinde' by Seyfried, Karl Heinrich, 'Die unversehene Wette' by Gotter, Friedrich Wilhelm, and 'Das Rendezvous oder Untreu aus Liebe' by Claudius, Georg Karl.

Autor*in	Titel	Jahr (normalisiert)	Szenen	Figuren	Schlagwörter
(Anonym)	Zeneide Ein Lustspiel in einer Handlung aus dem Französischen des Herrn von Cahusac	1747	11	4	• Liebesprobe
Suabe, Carl August  0111924752	Der Sehende Blinde Ein Nachspiel in Versen, nach dem Französischen des Herrn Le Grand	1752	29	8	• Verkleidungsspiel • Liebesprobe • Heimkehr-Motiv • mehraktige Varianten
Möser, Justus  065087	Die Tugend auf der Schaubühne oder Harlekins Heirath Ein Nachspiel in Einem Aufzuge  084606039  ger000481	1763	17	7	• Theaterbetrieb • Liebesprobe • Eheanbahnung
Seyfried, Karl Heinrich  094901866	Der sehende Blinde Ein Lustspiel in einem Aufzuge, nach dem Französischen des Le Grand	1764	30	8	• Verkleidungsspiel • Liebesprobe • Heimkehr-Motiv
Gotter, Friedrich Wilhelm  070356	Die unversehene Wette Ein Lustspiel in Einem Akt. Nach Sedaine	1773	29	8	• Liebesprobe • Eifersucht • Ehe
Claudius, Georg Karl  019279269	Das Rendezvous oder Untreu aus Liebe Ein Lustspiel in einem Aufzuge	1783	23	9	• Eheanbahnung • Liebesprobe

Vgl. Dîlan Canan Çakir, Frank Fischer: **Dramatische Metadaten. Die Datenbank deutschsprachiger Einakter 1740–1850.**

In: DHd2022: »Kulturen des digitalen Gedächtnisses«. 7.–11. März 2022.

Book of Abstracts. Universität Potsdam. (doi:10.5281/zenodo.6327977)

Miguel Escobar Varela: Theater as Data: Computational Journeys into Theater Research (2022)

Table of Contents

Introduction: In Pursuit of Theater's Digital Traces

Part 1: Pre-Departure Reflections

1 Towards a More Nuanced Conversation on Methodology

2 The Roles of Statistics

3 The Roles of Visualizations

Part 2: Guided Tours

4 Words as Data

5 Relationships as Data

6 Motion as Data

7 Locations as Data

Part 3: Ensuring the Journeys Continue

8 The Imperative of Open and Sustainable Data

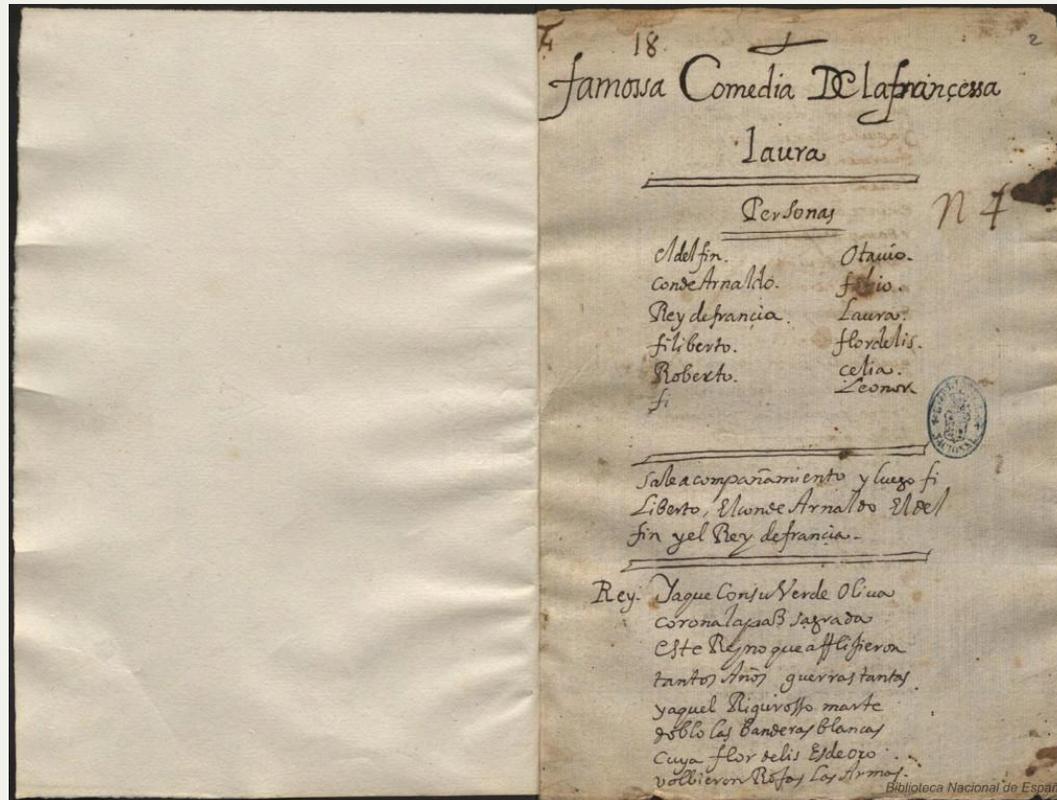
9 The Roles of Software Programming

Appendix A: Data Biographies

Appendix B: Technical Glossary

Inhaltsverzeichnis des Bandes.

Lope de Vega: »La francesa Laura« (2023)



(Bildquelle: <http://bdh.bne.es/bnearch/detalle/bdh0000216882>)

Publikation zu diesem Fund:

Álvaro Cuéllar, Germán Vega García-Luengos: «La francesa Laura». *El hallazgo de una nueva comedia del Lope de Vega último* (2023), doi:10.5565/rev/anuariolopedevega.492.

Dramen und Machine Learning/KI/GPT

- ChatGPT:
 - post-correction tasks (OCR)
 - automatic encoding (TEI)
- GPT-2:
 - Mariam Bangura, Kristina Barabashova, Anna Karnysheva, Sarah Semczuk, Yifan Wang: Automatic Generation of German Drama Texts Using Fine Tuned GPT-2 Models. arXiv:2301.03119 [cs.CL]

Zusammenfassung

- seit 1984: programmiertechnische Umsetzung vorgelegter formalistischer und strukturalistischer Ansätze
- zusätzlich NLP-basierte Verfahren
- TEI als Kodierungsstandard (mittlerweile viele Dutzend Dramenkorpora)
- Open Source-Paradigma, FAIR-Prinzipien
- breit genutzte Standardtools, die sich auch für die Dramenanalyse nutzen lassen (Gephi, Stylo etc.)
- Infrastrukturen
- APIs
- Linked Open Data

Vielen Dank.



<https://dracor.org/>

#DraCor

#ProgrammableCorpora