

# Digitale Netzwerkanalyse dramatischer Texte

Frank Fischer<sup>1</sup>, Dario Kampkaspar<sup>2</sup>, Peer Trilcke<sup>3</sup>

<sup>1</sup>Göttingen Centre for Digital Humanities

<sup>2</sup>Herzog August Bibliothek Wolfenbüttel

<sup>3</sup>Georg-August-Universität Göttingen

<sup>1</sup>frank.fischer@zentr.uni-goettingen.de, <sup>2</sup>kampkaspar@hab.de,

<sup>3</sup>trilcke@uni-goettingen.de

DHd 2015 – Graz

# Erkenntnisversprechen strukturanalytischer Ansätze

- mehr oder weniger abstrakte Modelle der (semantischen, kompositorischen etc.) Organisation literarischer Texte
- literarhistorische Aussagen zum Strukturwandel literarischer Texte
- der Gattungsaspekt, etwa die Frage, ob es netzwerkstrukturelle Unterschiede zwischen Tragödie und Komödie gibt

## Vorläufer der literarischen Netzwerkanalyse

*“Stellen wir uns einen Zuschauer vor, der im Laufe der Theatervorstellung nur vermag, die Auf- und Abtritte der Schauspieler zu beobachten und jeden Charakter zu erkennen, also auch zwei verschiedenartige Charaktere voneinander zu unterscheiden. Die Informationsmenge, die unser Zuschauer empfängt, nennen wir **die Szenenstruktur des Stückes.**”*

Solomon Marcus: *Mathematische Poetik*, Frankfurt/M. 1973, S. 289

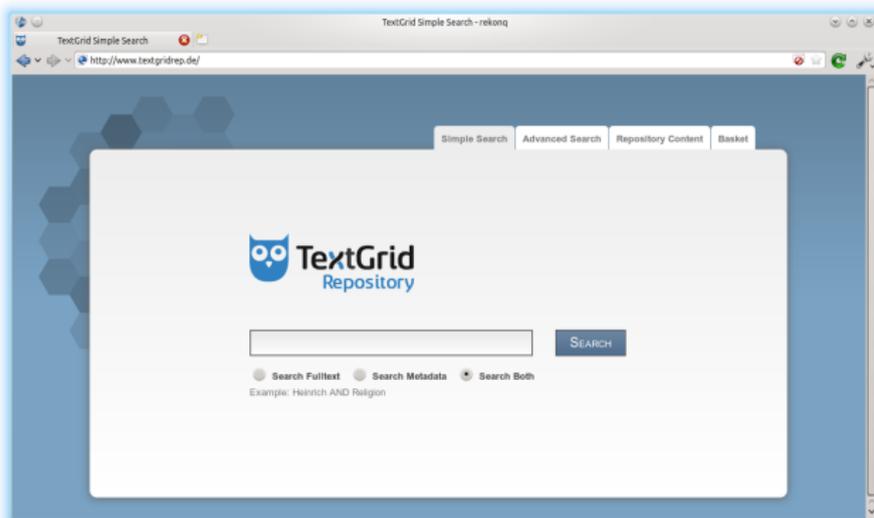
# Unser Ansatz

- Anknüpfen an Marcus: Interaktionsstruktur
- Dezidierter Bezug auf Netzwerkanalyse
- Ziel: Automatisierung der Analyse
- Status: work in progress!

# Bestehende Korpora deutschsprachiger Dramen

- Deutsches Textarchiv: 54 Dramen (Belletristik :: Drama)
- TextGrid Repository: 690 Dramen (genre:"drama")
- Wikisource: 50 Dramen (Kategorie:Drama)
- Projekt Gutenberg-DE: 641 Dramen (Dramatik)

# TextGrid Repository

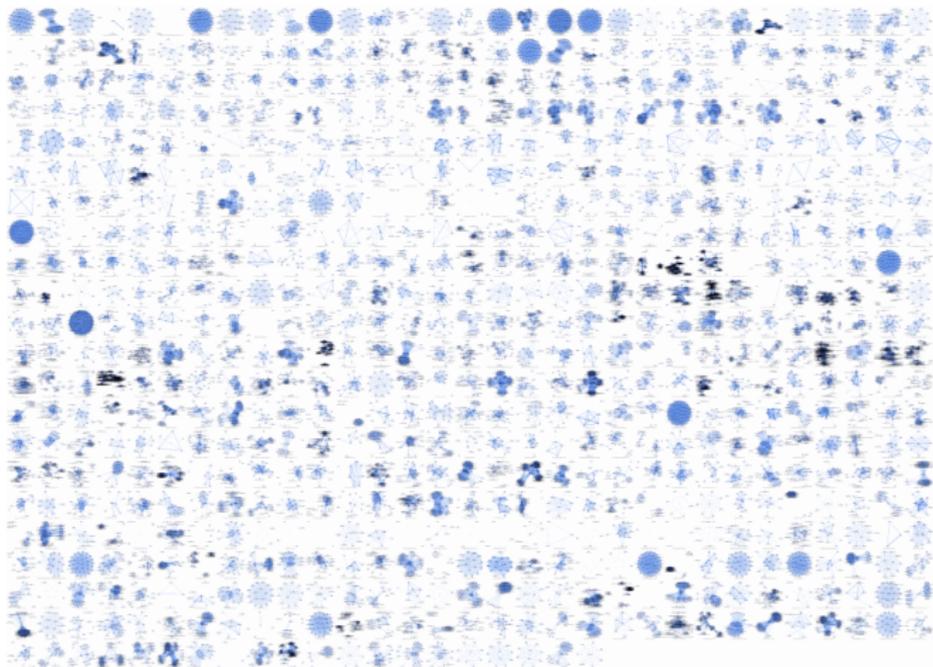


textgridrep.de / genre:drama / tei:speaker/string()

# TextGrid Repository

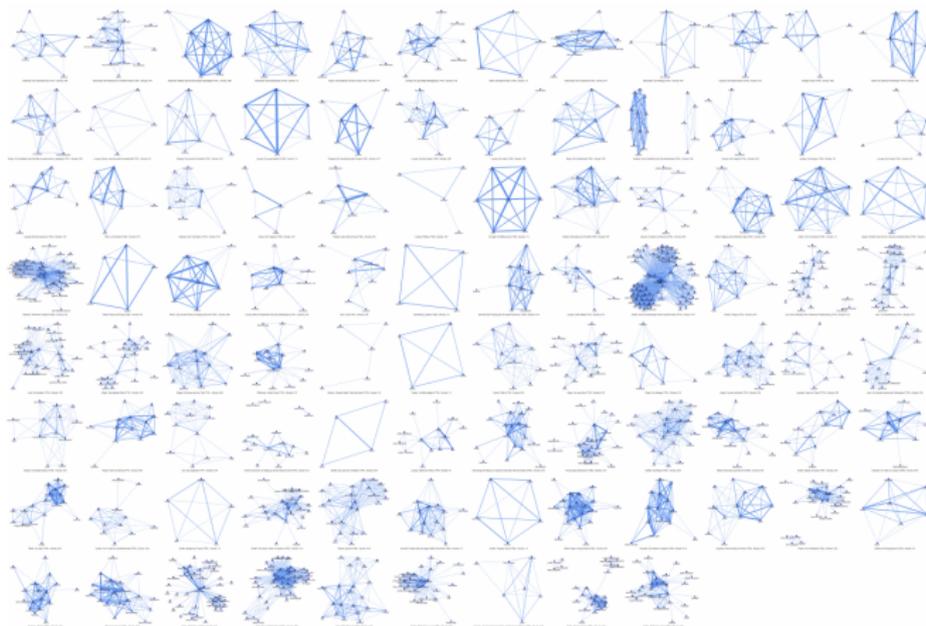
- in basales TEI transformierte Texte inkl. <speaker>-Tags
- 690 meist deutschsprachige Dramen (darunter eine Handvoll griechischer Tragödien und Shakespeare; auch einige Opernlibretti)
- Zeitraum: 5. Jh. v. Z. sowie 1510–1931 (Angaben im <note>-Tag)
- aus dem 18. Jh. (seit Gottscheds *Cato*): ca. 150 Dramen und Libretti (je nachdem, ob man von Entstehungszeit, Uraufführung oder Erstdruck ausgeht)
- Auswertung auf Basis von 93 Stücken zwischen 1731–1799 (Auslassung von Fragmenten, Libretti und Stücken ohne Szenensegmentierung)
- abgezwigte Teilkorpora: Tragödien und Komödien
- Ziele: strukturanalytische Beschreibung des Modernisierungsprozesses seit Gottsched, etwa des Einflusses der Shakespeare-Rezeption seit ca. 1770; geschlossenes vs. offenes Drama (nach Klotz 1968; vgl. Trilcke 2013)

# Aufbereitung des Korpus



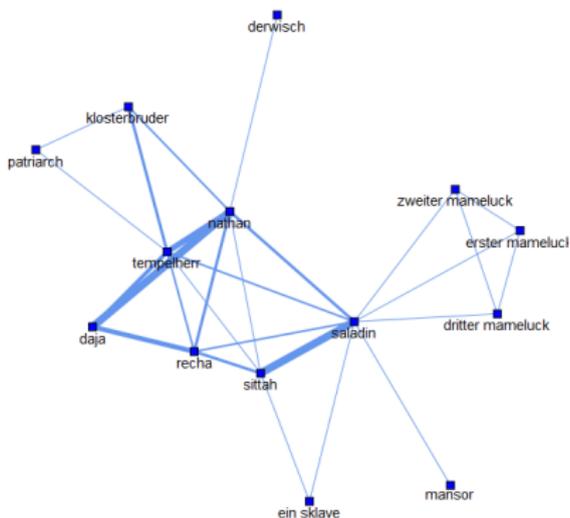
671 Stücke in all their dubious beauty

# Aufbereitung des Korpus



93 Stücke (1731–1799)

# Netzwerkgraph in schön



Lessing: *Nathan der Weise* (1779), Netzwerk gemäß ausgelesenem  
tei:speaker/string(), nachbearbeitet



# Probleme bei Personennamen

Wenn `xml:id/@who` nicht vergeben sind, muss man auf die Strings zurückgreifen, wobei sich folgende Probleme ergeben:

- übernommene Druckfehler
- OCR-Fehler
- Abkürzungen aus editorischen Gründen
- Inkonsistenzen bei der Sprecherangabe (“Odoardo” vs. “Odoardo Galotti” usw.)
- keine weitere Qualifizierung von Kollektiva (“ALLE”, “BEIDE” usw.)
- zur Fehlersuche dienen die automatisch erzeugten Graphen, als wichtiger Teil der Prüfschleife bei der Korpusaufbereitung

# DTA vs. TextGrid Rep

```

<sp who="#FAU">
  <speaker>
    |   <hi rendition="#g">Faust</hi>.
  </speaker><lb/>
  <p>Habe nun, ach! Philosophie,<lb/>
  Juristerey und Medicin,<lb/>
  Und leider auch Theologie!<lb/>
  Durchaus studirt, mit heißem Bemühn.<lb/>
  [...]
```

```

<sp>
  <speaker xml:id="tg1386.2.7">FAUST.</speaker><lg>
  <l rend="zenoPLm4n4" xml:id="tg1386.2.8">Habe nun, ach! Philosophie,</l>
  <l rend="zenoPLm4n4" xml:id="tg1386.2.9">Juristerei und Medizin,</l>
  <l rend="zenoPLm4n4" xml:id="tg1386.2.10">Und leider auch Theologie</l>
  <l rend="zenoPLm4n4" xml:id="tg1386.2.11">Durchaus studiert, mit heißem Bemühn.</l>
  [...]
```

Beginn des *Faust I* (1808), oben: DTA, unten: TextGrid Rep

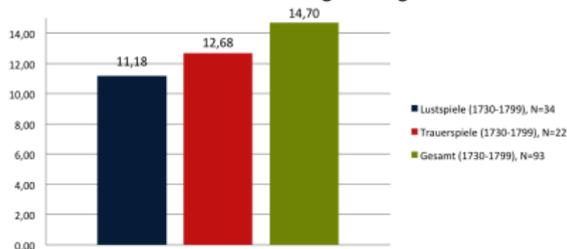
# Auswertung – erste Beobachtungen

Auf der Grundlage der bisher erhobenen und korrigierten Daten lassen sich bestimmte Fragen/Thesen diskutieren. Trotz der noch zu verbreiternden Datengrundlage zeichnen sich schon Ergebnisse ab, veranschaulicht an diesen vier Punkten:

- 1 Gattungscharakteristika
- 2 Historischer Strukturwandel
- 3 Autoren – Norm und Abweichung
- 4 Zentralitätswerte

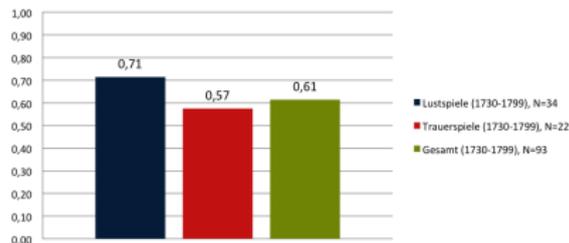
# Gattungscharakteristika

Abb. 01: Anzahl Figuren Avg.



Anzahl der Figuren (Mittel)

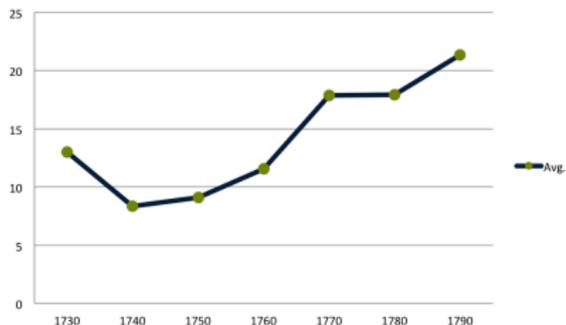
Abb. 02: Dichte Avg.



Dichte (Mittel)

# Historischer Strukturwandel (1)

Abb. 03: Anzahl Figuren Avg.



Anzahl der Figuren (Mittel)

Abb. 04: Average Degree Avg.

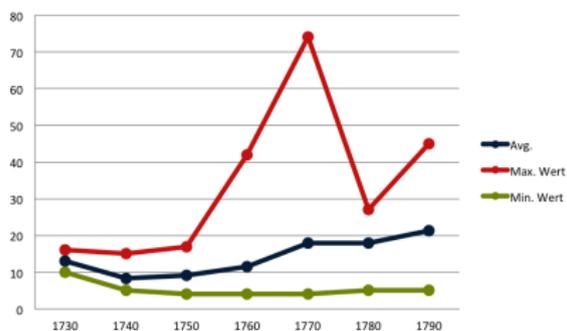


Average Degree (Mittel)

(jeweils auf Jahrzehnte gerechnet)

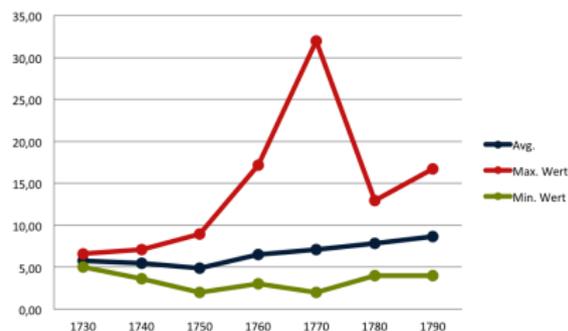
# Historischer Strukturwandel (2)

Abb. 05: Anzahl Figuren Avg./Max./Min.



Anzahl der Figuren (Min./Mittel/Max.)

Abb. 06: Average Degree Avg./Max./Min.

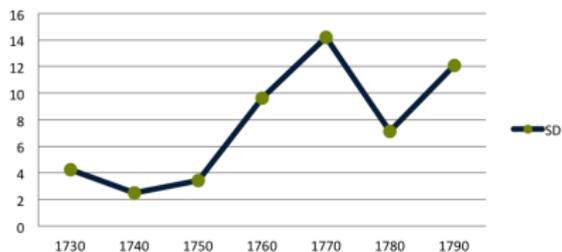


Average Degree (Min./Mittel/Max.)

(jeweils auf Jahrzehnte gerechnet)

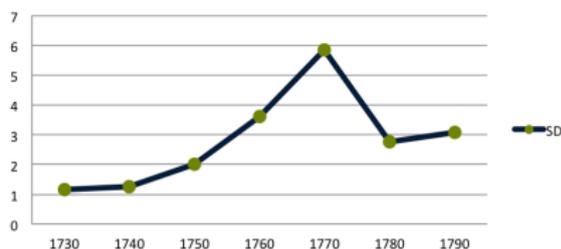
# Historischer Strukturwandel (3)

Abb. 07: Anzahl Figuren SD



Anzahl der Figuren (SD)

Abb. 08: Average Degree SD



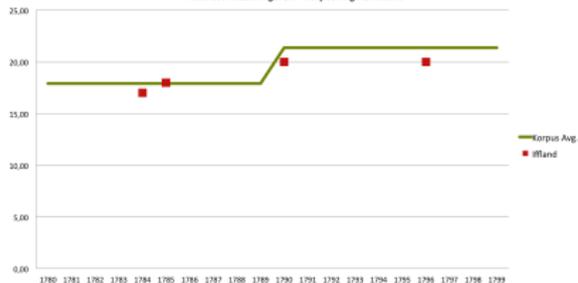
Average Degree (SD)

(jeweils auf Jahrzehnte gerechnet)

# Autoren – Norm und Abweichung

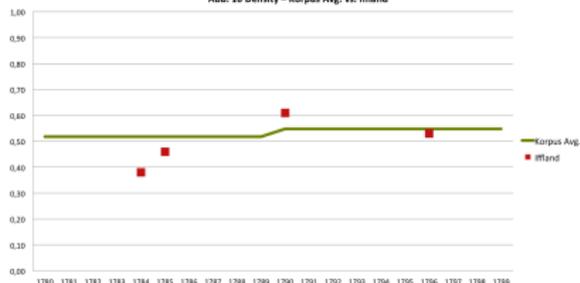
## Beispiel: Iffland

Abb. 09: Anzahl Figuren – Korpus Avg. vs. Iffland



Anzahl der Figuren

Abb. 10 Density – Korpus Avg. vs. Iffland



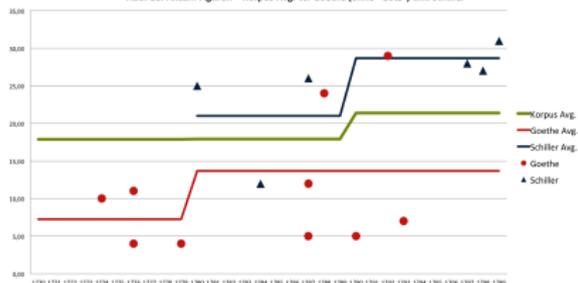
Dichte

(Werte des Gesamtkorpus auf Jahrzehnte gerechnet)

# Autoren – Norm und Abweichung

## Beispiel: Goethe und Schiller

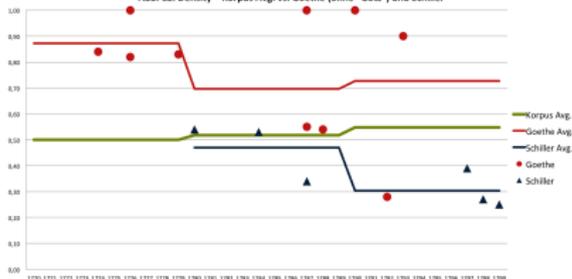
Abb. 11: Anzahl Figuren – Korpus Avg. vs. Goethe (ohne "Götz") und Schiller



Anzahl der Figuren

(Werte des Gesamtkorpus auf Jahrzehnte gerechnet)  
 (Goethes Werke hier ohne den "Götz von Berlichingen")

Abb. 12: Density – Korpus Avg. vs. Goethe (ohne "Götz") und Schiller



Dichte

# Zentralitätswerte (1)

## Goethe: Iphigenie auf Tauris (1787)

Anzahl Figuren: 5

Max Degree: 4

Average Degree: 4.0

Density: 1.0

Average Path Length: 1.0

Figur	Betw. Centr.	Degree	Average Dist.	Closeness Centr.
iphigenie	0.0	4	1.0	1.0
thoas	0.0	4	1.0	1.0
pylades	0.0	4	1.0	1.0
orest	0.0	4	1.0	1.0
arkas	0.0	4	1.0	1.0

## Zentralitätswerte (2)

### Lessing: Nathan der Weise (1779)

Anzahl Figuren: 14

Max Degree: 9

Average Degree: 3.86

Density: 0.3

Average Path Length: 1.93

Figur	Betw. Centr.	Degree	Average Dist.	Closeness Centr.
daja	0.0	3	2.15	0.46
nathan	18.83	7	1.46	0.68
recha	2.33	5	1.62	0.62
tempelherr	16.83	7	1.46	0.68
saladin	42.5	9	1.31	0.76
sittah	3.5	5	1.62	0.62
derwisch	0.0	1	2.38	0.42
klosterbruder	1.0	3	2.15	0.46
patriarch	0.0	2	2.31	0.43
sklave	0.0	2	2.15	0.46
zweiter mameluck	0.0	3	2.08	0.48
erster mameluck	0.0	3	2.08	0.48
dritter mameluck	0.0	3	2.08	0.48
mansor	0.0	1	2.23	0.45

# Zusammenfassung

- Grundlegung einer literaturwissenschaftlichen Netzwerkanalyse anhand älterer Vorarbeiten und Überführung dieser Ideen in eine digitale Analyse großer Korpora
- Aufbau eines Workflows zur Extraktion basaler Netzwerkdaten für Dramentexte aus rudimentär gemarkupten Korpora
- erste Auswertungen zur Gattungsdifferenzierung und literarischen Evolution

## To do

- noch zu wenige und teils zu unsaubere Daten
- bisher nur Erfassung rudimentärer Strukturinformationen, aber Interaktionen in Dramen sind viel komplexer
- Auswertung und Vergleich figurenbezogener Werte (Degree, Betweenness Centrality, Closeness Centrality bzw. Average Distance)
- Vergleich mit Dramenkorpora anderer Sprachen

# Literatur

- Klotz, Volker: *Geschlossene und offene Form im Drama*. 3. Aufl. München 1968.
- Marcus, Solomon: *Mathematische Poetik*. Frankfurt/M. 1973.
- Moretti, Franco: *Network Theory, Plot Analysis*. Stanford Literary Lab Pamphlets 2 (01.05.2011), URL: <http://litlab.stanford.edu/LiteraryLabPamphlet2.pdf>
- Pfister, Manfred: *Das Drama. Theorie und Analyse*. München 1977 u. ö.
- Titzmann, Michael: *Strukturelle Textanalyse. Theorie und Praxis der Interpretation*. München 1977 u. ö.
- Trilcke, Peer: *Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft*. In: Philip Ajouri, Katja Mellmann, Christoph Rauen (Hg.): *Empirie in der Literaturwissenschaft*. Münster 2013, 201–247.
- Wasserman, Stanley; Faust, Katherine: *Social Network Analysis. Methods and Applications*. New York 1994.