

Wann findet die deutsche Literatur statt?

Frank Fischer¹ & Jannik Strötgen²

¹Göttingen Centre for Digital Humanities

²Institut für Informatik, Ruprecht-Karls-Universität Heidelberg

¹frank.fischer@zentr.uni-goettingen.de, ²stroetgen@uni-hd.de

DHd 2015 – Graz, Austria

Suomen kirjallisuus

*“Die meisten **finnischen Romane** spielen im **Sommer**.”*

Stefan Moster, *Im Gespräch*, Deutschlandradio Kultur, 8.10.2014

Durchs Gebirg

“Den 20. Jänner ging Lenz durchs Gebirg.”

Georg Büchner: *Lenz* (ersch. 1839)

Bloomsday

*“Miss Dunne clicked on the keyboard:
— 16 June 1904.”*

Jamey Joyce: *Ulysses* (ersch. 1922)

Effi Briests Hochzeitstag

*“Ende August war da, der Hochzeitstag
(3. Oktober) rückte näher, [...]”*

Theodor Fontane: *Effi Briest* (ersch. 1896)

Hunger nach Gedrucktem

“Und man ließ mich besonders lange warten an diesem Donnerstag, dem 27. Juli, zwei geschlagene Stunden im Vorzimmer stehend warten; ich erinnere mich auch an dieses Datum aus einem bestimmten Grunde so genau, denn in diesem Vorzimmer, wo ich selbstverständlich, ohne mich niedersetzen zu dürfen zwei Stunden mir die Beine in den Leib stehen mußte, hing ein Kalender, und ich vermag Ihnen nicht zu erklären, wie in meinem Hunger nach Gedrucktem, nach Geschriebenem ich diese eine Zahl, diese wenigen Worte '27. Juli' an der Wand anstarrte und anstarrte; ich fraß sie gleichsam in mein Gehirn hinein.”

Stefan Zweig: *Schachnovelle* (ersch. 1942)

Gefangen im Fiktionsfeld?

“Das Datum spielt keine andere Rolle als jede andere Charakterisierung eines Romantages auch, es ist nichts als ein Stück des Wirklichkeitsstoffes und ist in der Fiktion ebenso fiktiv wie Haus und Straße, Feld und Wald, wie die Städte Mannheim und Köln, [...]. Denn sobald Zeit und Ort das Erlebnisfeld fiktiver Personen, d. h. ein Fiktionsfeld abgeben, ist es um ihre 'Wirklichkeit' getan, mag auch das Fiktionsfeld Bestandteile aufweisen, die einer mehr oder weniger allgemein bekannten Wirklichkeit entstammen.”

Käte Hamburger: *Die Logik der Dichtung*, Stuttgart 1957, S. 57

Gutenberg-DE

- Basis: Gutenberg-DE Edition 13 DVD-ROM (Oktober 2013)
- Teilkorpus mit deutschsprachigen fiktionalen Texten (891 MB reiner Text)
 - extrahiert wurden alle Texte der Typen: `<meta name="type" content="{fiction,narrative,novelette}"/>`
 - daraus eliminiert wurden Übersetzungen (in den Metadaten sind zum Teil Übersetzer angegeben: `<meta name="translator" content="[...]" />`)
 - wegen fehlender oder falscher Metadaten war eine manuelle Nachbearbeitung nötig
- insgesamt 549 Autoren und 2735 Werke von ca. 1510 bis in die 1940er-Jahre (Schwerpunkt liegt zwischen 1840–1930)

Gutenberg-DE (Top-30-Autoren $\hat{=}$ 35,32%)

Autor innen	Werkumfang in Byte	Anteil am Gesamtkorpus
Karl May	29856696	3,19%
Robert Kraft	20107708	2,15%
Friedrich Gerstäcker	16430129	1,76%
Wilhelm Raabe	14636947	1,57%
Friedrich Spielhagen	13482907	1,44%
Karl Gutzkow	13292673	1,42%
Jakob Wassermann	12695650	1,36%
Rudolph Stratz	12316073	1,32%
Ludwig Ganghofer	11727769	1,25%
Joseph von Lauff	11289743	1,21%
Jean Paul	11246711	1,20%
Georg Ebers	11139138	1,19%
Ludwig Tieck	10689713	1,14%
Gustav Freytag	10546794	1,13%
Jeremias Gotthelf	10437136	1,12%
Berthold Auerbach	10290698	1,10%
Felix Dahn	9994799	1,07%
Willibald Alexis	9068887	0,97%
F. W. Hackländer	8535643	0,91%
Oskar Meding	8337665	0,89%
Christoph Martin Wieland	8188018	0,88%
Theodor Fontane	8078790	0,86%
E. T. A. Hoffmann	8067688	0,86%
Balduin Möllhausen	7891664	0,84%
Adalbert Stifter	7120172	0,76%
Peter Rosegger	7069315	0,76%
Gottfried Keller	6918867	0,74%
Hermann Löns	6911193	0,74%
E. Marlitt	6906046	0,74%
Ernst Weiß	6867008	0,73%

Extraktion normalisierter Datums- und Monatsangaben



Extraktion normalisierter Datums- und Monatsangaben

Temporal Tagging

- Extraktion und Normalisierung von Zeitausdrücken

Zeitausdrücke nach TimeML

- date (heute, am Montag, 10. August, am nächsten Tag, Juni)
- time (am Morgen, um 13 Uhr)
- duration (5 Jahre, 2 Stunden)
- set (jährlich, einmal am Tag)

Temporal Tagger

- HeidelTime - frei verfügbar, 11 Sprachen, 4 Domänen

Extraktion normalisierter Datums- und Monatsangaben

Temporal Tagging

- Extraktion und Normalisierung von Zeitausdrücken

Zeitausdrücke nach TimeML

- date (heute, am Montag, **10. August**, am nächsten Tag, **Juni**)
- time (am Morgen, um 13 Uhr)
- duration (5 Jahre, 2 Stunden)
- set (jährlich, einmal am Tag)

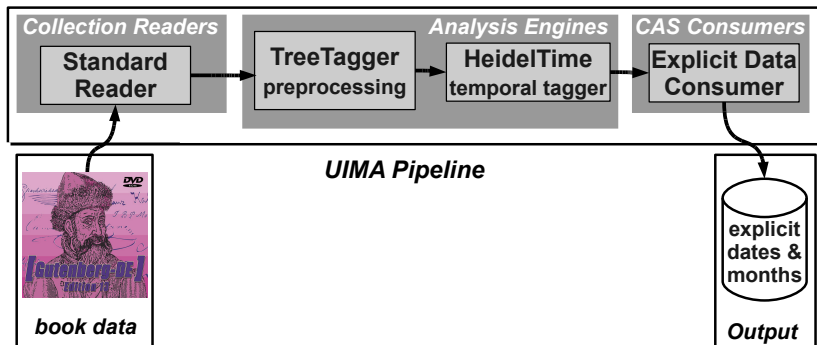
Temporal Tagger

- HeidelTime - frei verfügbar, 11 Sprachen, 4 Domänen

Extraktion normalisierter Datums- und Monatsangaben



Extraktion normalisierter Datums- und Monatsangaben

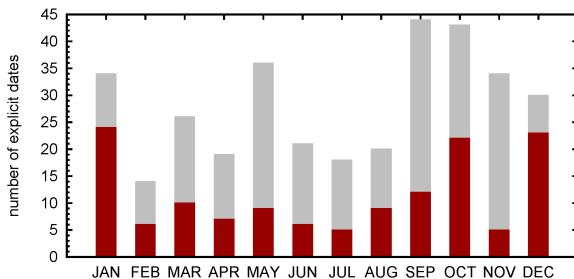


Was wir aus der Pipeline bekommen ...

Statistiken über und Textauszüge mit
expliziten Datums- und Monatsangaben

Was wir aus der Pipeline bekommen ...

“Monate vs. Tage” Auswertungen



Was wir aus der Pipeline bekommen ...

“Explizite Tage” Heatmaps

```

EXPLICIT DAYS: 81
MONTH: 01 21321111---2-211-121-1--1-----
MONTH: 02 1-11-----1--1--1-----
MONTH: 03 -----2---1---2-----2--1-2--
MONTH: 04 1-----1-2-----1-----
MONTH: 05 321--2----1-----
MONTH: 06 2-----1----1--1-1-
MONTH: 07 2-1-----1-----1---
MONTH: 08 -----12-2112-----
MONTH: 09 -3---1-1--2-----3----1--1
MONTH: 10 7-82-----4-----1-----
MONTH: 11 -----1--12-----1---
MONTH: 12 -6-1--1--12-----1--4-1--123

```

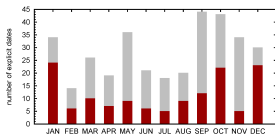
Was wir aus der Pipeline bekommen ...

Textpassagen für “explizite Tage”

01-01	Die Poggenpuhls (1896)	1. Januar	SATZ: Und so vernimm denn, am Silversterabend noch ein Bettler (allerdings ein glücklicher, denn wir brachten es im Kasino auf sieben Bowlen in Großformat) und am 1. Januar früh ein Gott, ein Krösus.
01-01	Mathilde Möhring (1908)	erste Januar	SATZ: Der erste Januar war ein wundervoller Wintertag, alles überreift und übereist, aber nicht sehr kalt und eine helle Sonne am blauen Himmel.
01-02	Meine Kinderjahre (1894)	zweiten Januar	SATZ: Silvester war wie herkömmlich gefeiert worden, und für den zweiten Januar stand ein neues Vergnügen in Sicht, von dem ich mir ganz besonders viel versprach:
01-03	Die Poggenpuhls (1896)	dritte Januar	SATZ: Es war ein Wintertag, der dritte Januar.
01-03	Effi Briest (1895)	3. Januar	SATZ: Seit Silvesternacht ging ein scharfer Nordost, der sich in den nächsten Tagen fast bis zum Sturm steigerte, und am 3. Januar nachmittags hieß es, daß ein Schiff draußen mit der Einfahrt nicht zustande gekommen und hundert Schritt vor der Mole gescheitert sei;
01-03	Meine Kinderjahre (1894)	dritten Januar	SATZ: mein Freund Wilhelm Krause, der schon als Schüler und Pensionär des bekannten Direktors v. Klöden die Gewerbeschule besuchte, mußte am dritten Januar wieder in Berlin sein, und seitens seines Vaters, des Kommerzienrats, war mit einigen Freunden verabredet worden, dem liebenswürdigen Jungen bis nach dem jenseitigen Ufer hinüber, von wo dann die Fahrpost ging, das Geleit zu geben.

Was wir aus der Pipeline bekommen ...

Monate vs. Tage



Tage Heatmaps

EXPLICIT DAYS: 81

```

MONTH: 01 21321111--2-211-121-1-1-----
MONTH: 02 1-11-----1-1-1-----
MONTH: 03 -----2--1--2--2--2-1-2--
MONTH: 04 1-----1-2-----1-----
MONTH: 05 321-2----1-----
MONTH: 06 2-----1-----1-----1-1-1-
MONTH: 07 2-1-----1-----1-----
MONTH: 08 -----12-2112-----
MONTH: 09 -3--1-1-1-2-----3--1-1-
MONTH: 10 7-82-----4-----1-----
MONTH: 11 -----1-12-----1-----
MONTH: 12 -6-1-1-12-----1-4-1-123
  
```

Tage – Texte

- I-01 Sie Poggenpohl (1896) 1. Januar
SATZ: Und so vermach dem, ein Silberstrahlband noch ein matter Glanzlichtung ein glückliches, denn wir brachten es im Kasten auf sieben Bänden in Goldfäden) und am 1. Januar FROB ein Satz, ein Kreuz.
- 01-01 Mathilde Möhring (2008) erste Januar
SATZ: Der erste Januar war ein sonderbarer Mittertag, alles überhitzt und überreizt, aber nicht oder halb und eine helle Sonne an Blauen Himmel.
- 01-02 Meine Kinderjähre (1884) zweiten Januar
SATZ: Silvester war wie herkömmlich gefeiert worden, und für den zweiten Januar stand mir neues Verlangen in Sicht, von dem ich mir ganz besonders viel versprach.
- 01-03 Sie Poggenpohl (1896) dritte Januar
SATZ: In der aus Mittertag, der dritte Januar.
01-03 Effi Briest (1895) 3. Januar
SATZ: Seit Silvesterabend ging ein sicherer Herbst, der mich in dem nächsten Tagen fast bis zum Sternsteigerte, und am 3. Januar nachmittags ließ es, daß die Welt endlich mit der Eifersucht nicht zufriede gekommen und Hundert Schritt vor der Welt geschickter sei.
- 01-03 Meine Kinderjähre (1884) dritten Januar
SATZ: mein Freund Wilhelme Krause, der schon als Schüler und Ausbilder des bekannten Bruckers v. Klippe die amerikanische Beweise, hatte am dritten Januar wieder in Berlin sein, und seinen seinen Vater, des Kommerzienrats, war mit einigen Freunden verabredet worden, des Lebensunfähigen Jungen bis nach dem jeweiligen oder hundert, von wo dem die Rückpost ging, das selbst zu gehen.

alle Statistiken pro Werk, Autor, Epoche, ...

Im wunderschönen Monat Mai

*“Im wunderschönen Monat Mai,
Als alle Knospen sprangen,
Da ist in meinem Herzen
Die Liebe aufgegangen.*

*Im wunderschönen Monat Mai,
Als alle Vögel sangen,
Da hab ich ihr gestanden
Mein Sehnen und Verlangen.”*

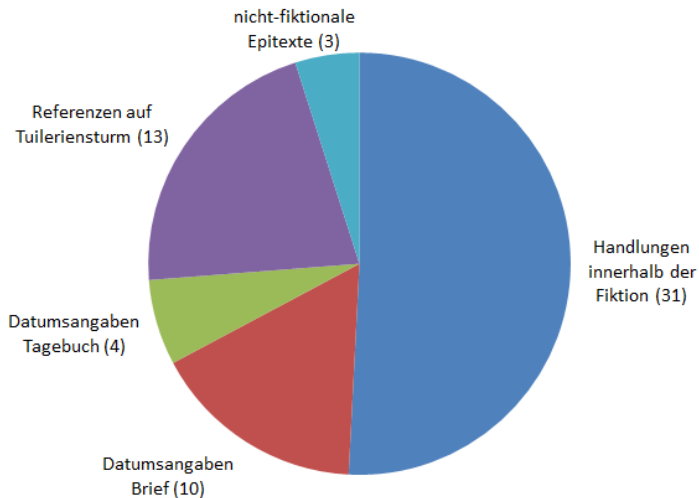
Heinrich Heine: *Buch der Lieder* (ersch. 1827)

Heatmap

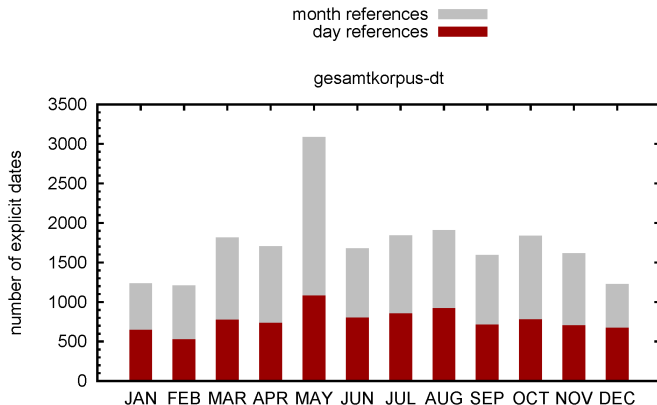
MONTH: 01	■333222323131323223222222222131	
MONTH: 02	43322222133212332332212322231	
MONTH: 03	■3334322233243422433■3232322252	(21.)
MONTH: 04	■33233223432223223332322223323	
MONTH: 05	■3544332353■4353232424323223244	(12.)
MONTH: 06	■33233323333324432343324433233	
MONTH: 07	■444332333243■52333432224223223	(14.)
MONTH: 08	■3■442232■2444■3344533332323222	(3., 10., 15.)
MONTH: 09	■54433233332234233233221222323	
MONTH: 10	■353322224223552253432222222133	
MONTH: 11	■44233333■23225213232222222224	(10.)
MONTH: 12	55223412132232321322333■2122224	(24.)

Grün hinterlegt: Tage mit 50+ Vorkommen im Korpus

Semantiken des 10. August

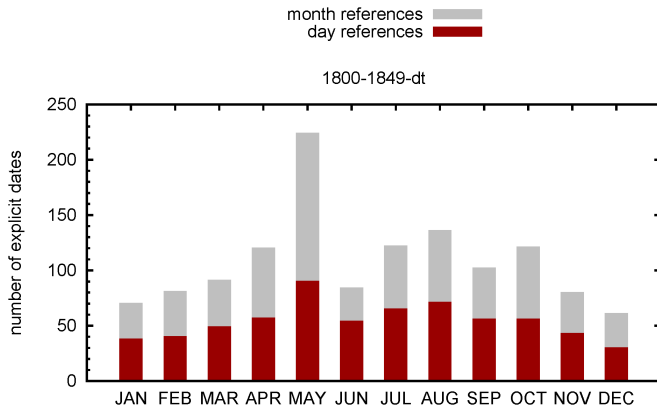


Datums- vs. Monatsnennungen



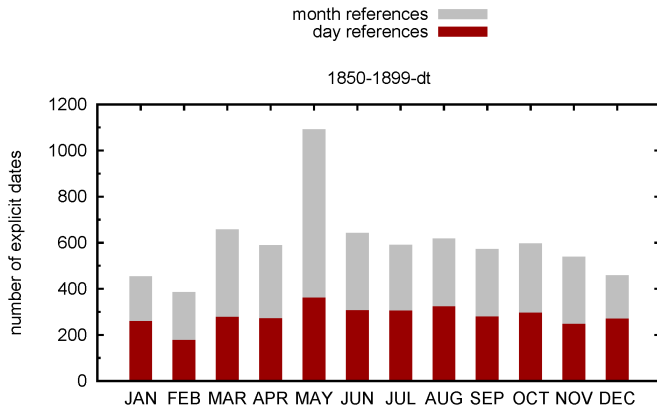
Gesamtkorpus deutsche Literatur: Mai gewinnt klar, Frühling und Sommer solide, Herbst schwächer, November bis Februar im Sinkflug

Teilkorpus 1800–1849



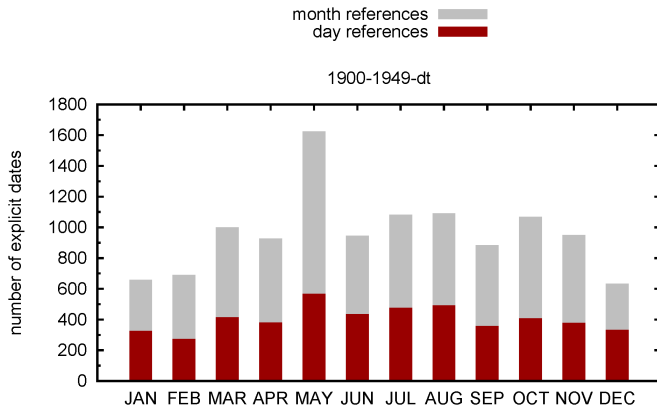
Und noch mal der Mai!

Teilkorpus 1850–1899



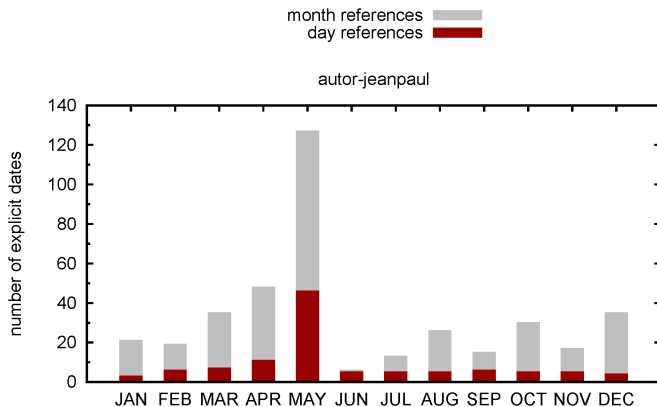
Und noch mal!

Teilkorpus 1900–1949



Und noch mal!

Jean Paul



“Im Kalender unserer Phantasie fällt der Frühlingsanfang nicht in den 21sten März, sondern in den ersten Mai; [...]” (*Palingenesien*, 1798)

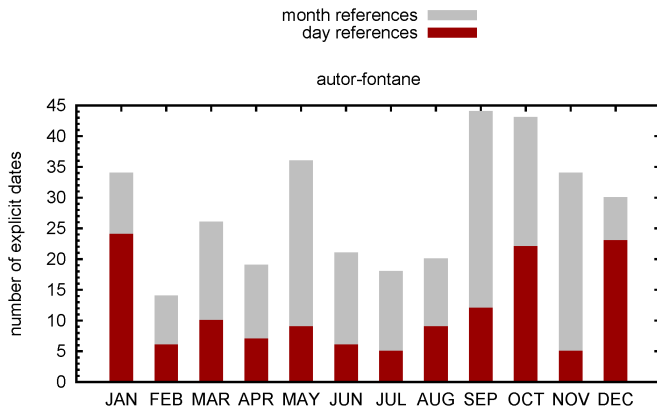
Wann spielt *Effi Briest*?



R. W. Fassbinder: *Fontane Effi Briest* (1974)

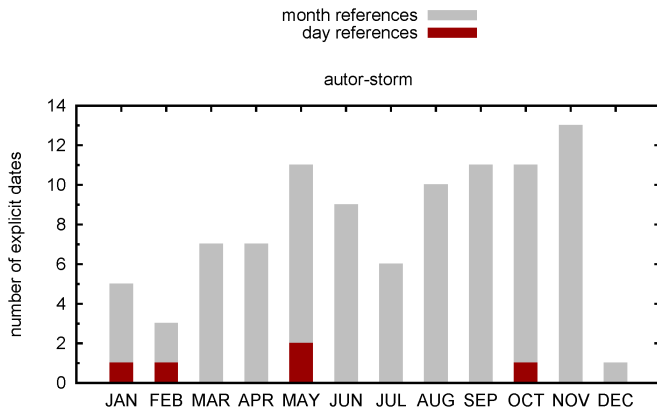
Bildschirmfotos © Basis-Film Verleih

Theodor Fontane



Ein Herbst- und Winterautor?

Theodor Storm



Zeitliche Verortung fast ausschließlich über Monatsnamen

Theodor Storm: *Der Schimmelreiter* (1888)

Alle 19 Monatsnennungen in der Novelle (alle unspezifisch):

- "an einem October-Nachmittag"
- "zu Ende October"
- "Im Februar bei dauerndem Frostwetter"
- "seit November"
- "Es war an einem Maiabend; aber es war Novemberwetter"
- "Es war im Januar"
- "es war zu Anfang des Septembers"
- "die goldene Septembersonne glitzerte"
- "Es war zu Ende März"
- "lauter Märzenluft"
- "Inzwischen war schon Ende März"
- "noch bis gegen Ende November"
- "Als es in den October hineinging"
- "Zu Ende November"
- "gegen Ende des März"
- "So war der September gekommen"
- "Zu Ende des Septembers"
- "als aber nach Ende September"
- "Es war vor Allerheiligen, im October."

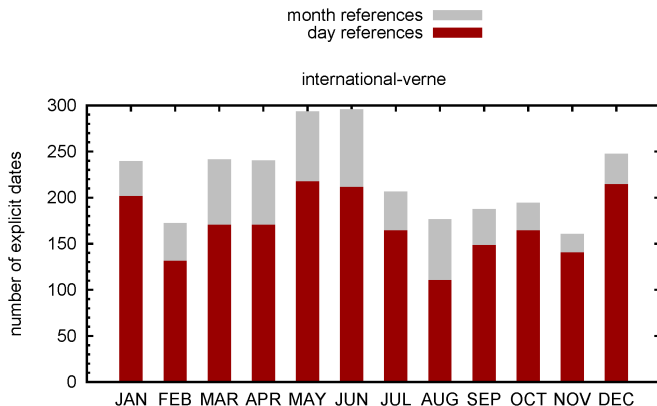
Gutenberg-DE (# Datumsvorkommen pro Werk)

#	Autor_in: Werk (Jahr)	Gattung
322	Karl Bleibtreu: Bismarck – Band 4 (1915)	Historischer Roman
288	Heinrich Albert Oppermann: Hundert Jahre (1870)	Historischer Roman
204	Lily Braun: Die Liebesbriefe der Marquise (1912)	Briefroman
170	Arthur Schurig: Seltsame Liebesleute (1920)	Briefroman
124	Robert Ascher: Der Schuhmeier (1933)	Romanbiografie
91	Karl Bleibtreu: Bismarck – Band 2 (1915)	Historischer Roman
91	Marie Nathusius: Tagebuch eines armen Fräuleins (1854)	Tagebuchroman
84	Bertha von Suttner: Die Waffen nieder! (1889)	Tendenzroman
79	Albert Emil Brachvogel: Oberst von Steuben (1875/1920)	Historischer Roman
74	M. A. v. Thümmel: Reise in die mittägl. Prov. von Frankr. (1791)	Reiseroman
69	Sir John Retcliffe: Sebastopol – Band 1 (1855)	Historischer Roman
66	J. W. von Goethe: Die Leiden des jungen Werther (1774)	Briefroman
60	Karl Bleibtreu: Bismarck – Band 3 (1915)	Historischer Roman
56	Karl von Holtei: Die Vagabunden (1851)	Bildungsroman
54	Heinrich Lersch: Die Pioniere von Eilenburg (1934)	Historischer Roman
53	Kurd Laßwitz: Auf zwei Planeten (1897)	Science-Fiction
52	Felicitas Rose: Heideschulmeister Uwe Karsten (1909)	Heimatroman
51	Karl Bleibtreu: Bismarck – Band 1 (1915)	Historischer Roman
46	Edlef Köppen: Heeresbericht (1930)	Historischer Roman
46	Franz Kafka: Die Acht Oktavhefte (1916)	Aufzeichn., Entwürfe
45	Philipp Galen: Der Strandvogt von Jasmund (1859)	Historischer Roman
41	F. X. Bronner: Ein Mönchsleben a. d. empfinds. Zeit – B. 2 (1850)	Autobiograf. Roman
40	Ludwig Rellstab: 1812 (1834)	Historischer Roman
39	Margarete Böhme: Tagebuch einer Verlorenen (1905)	Tagebuchroman
38	Alfred Schirokauer: Lassalle (1912)	Romanbiografie
37	Jura Soyfer: So starb eine Partei (1934)	Tendenzroman
36	Alfred Schirokauer: Mirabeau (1921)	Romanbiografie
35	Elisabeth von Heyking: Briefe, die ihn nicht erreichten (1903)	Briefroman
35	Ernst Constantin: Das warme Polarland (1891)	Abenteuroman
35	Friedrich Spielhagen: Frei geboren (1900)	Tagebuchroman

Gutenberg-DE (Werke mit nur 1 Datumsangabe)

Autor_in: Werk (Jahr)	Tag
Johann Karl Wezel: Lebensgeschichte Tobias Knauts (1773–76)	24. Juni
Novalis: Heinrich von Ofterdingen (1802)	16. März
E. T. A. Hoffmann: Der Sandmann (1816)	30. Oktober
Eduard Mörike: Mozart auf der Reise nach Prag (1855)	14. September
Gottfried Keller: Der Landvogt von Greifensee (1877)	31. Mai
Marie v. Ebner-Eschenbach: Lotti, die Uhrmacherin (1880)	12. Mai
Theodor Storm: Zur Chronik von Grieshuus (1884)	24. Januar
E. Marlitt: Das Eulenhäus (1888)	6. Juni
Carl Spitteler: Imago (1906)	4. Dezember
Fanny zu Reventlow: Der Geldkomplex (1916)	10. August
Eduard von Keyserling: Fürstinnen (1917)	20. August
Franz Kafka: Das Schloss (1926)	3. Juli
Arthur Schnitzler: Spiel im Morgengrauen (1927)	8. Juni
Robert Musil: Der Mann ohne Eigenschaften (1930)	2. Dezember
Stefan Zweig: Schachnovelle (1942)	27. Juli

Jules Verne, der Kalenderboy der Weltliteratur



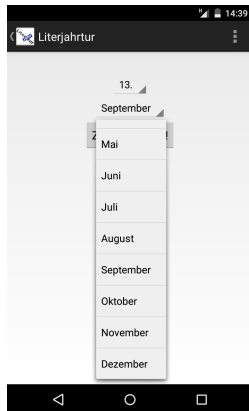
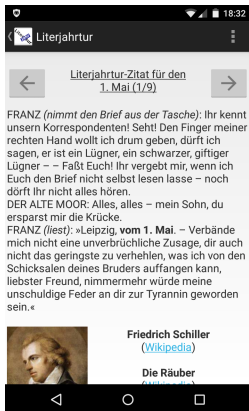
Die Abenteuerromane von Jules Verne haben
stets eine tagesgenaue Chronologie

Fiktive Datumsangaben

Welche Rolle spielen fiktive Daten?

- Erich Kästner: *Der 35. Mai oder Konrad reitet in die Südsee* (1931)
- Shakespeares 80. April in *The Winter's Tale*, Autolycus' Ballade im 4. Akt: "Here's another ballad of a fish, that appeared upon / the coast on Wednesday the four-score of April, / forty thousand fathom above water, [...]" (umgerechnet: "19. Juni")
- Bedeutung des 30. Februar, etwa als Synonym für den St. Nimmerleinstag
- "›Wann sind Sie geboren, Mr. Starke?‹ ›Am 31. Februar 1993.‹" (Robert Kraft: *Ein moderner Lederstrumpf*, 1902)
- 31. September (Jean Paul: *Flegeljahre*, 1804–05)
- ...

Android-App: Literjahrturn



Download der Beta-Version im Play Store:

<https://goo.gl/rF4NbJ>

Zusammenfassung

- Aufbau eines Workflows zur Monats- und Datumsanalyse großer literarischer Korpora (für mehrere Sprachen)
- herausgehobene Bedeutung des Monats Mai im benutzten Korpus deutscher Prosaliteratur
- Präferenzen einzelner Autoren für bestimmte Jahresabschnitte als Handlungs- und Referenzgrößen
- to do: Vergleiche mit anderen Nationalliteraturen auf breiterer Datenbasis
- to do: weitere Qualifizierung der Daten (was ist Handlung? was sind Referenzen auf historische Geschehnisse? usw.)

Literatur

- Matthew Jockers: *Macroanalysis. Digital Methods and Literary History*. Chicago: University of Illinois Press, 2013.
- James Pustejovsky, Jose M. Castano, Robert Ingria, Roser Sauri, Robert J. Gaizauskas, Andrea Setzer, Graham Katz and Dragomir R. Radev. *TimeML: Robust Specification of Event and Temporal Expressions in Text*. In: *New Directions in Question Answering*, pp. 28–34, 2003.
- Jannik Strötgen, Michael Gertz. *Temporal Tagging on Different Domains: Challenges, Strategies, and Gold Standards*. In: *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012)*, pp. 3746–3753, 2012.

Danke!



“Literjahr” im Play Store