

# Digital Humanities

## *Distant Reading* – Übungen mit *Voyant*

Diese Unterrichtseinheit führt ein in die Idee der *Digital Humanities* und in das Konzept des *Distant Reading*. In einer ersten Lerneinheit<sup>1</sup> geht es um die Grundidee und um die Möglichkeiten der Analyse literarischer Texte mit den *Voyant Tools*. Diese Analyselandschaft ermöglicht es, Einzeltexte oder ganze Korpora zu analysieren und zu visualisieren.

In der zweiten Lerneinheit *Distant Reading und Faust I* werden konkrete Aufgaben zu Goethes *Faust I* gelöst.

### 1. Vorbereitungslektüre (Hausaufgaben)

Aufgabe 1	Lies Franco Morettis Aufsatz <i>Die Schlachtbank der Literatur</i> .
Aufgabe 2	Lies den Blogpost <a href="#">Digital Humanities und Distant Reading</a> <sup>2</sup> .

### 2. Organisatorische Vorbereitungen

Aufgabe 3	Erstelle in deinem freigegebenen Aufsatzordner auf <i>OneDrive</i> einen Unterordner mit dem Namen <i>Distant Reading Vorname</i> . Hier speicherst du alle Hilfsdateien deiner Arbeit ab. Trage den Link in die Aufsatzübersicht ein.
Aufgabe 4	Dokumentiere deine Arbeit in einem neuen <i>Word</i> -Dokument mit dem Dateinamen <i>Distant Reading Vorname</i> . Halte hier deine Lösungen und Erkenntnisse fest. Auch die Hilfsdateien und Screenshots legst du systematisch hier ab. Speichere zunächst Morettis Aufsatz, den Link zum Blogpost und diese Aufgaben in deinem Projektordner bzw. in deinen Notizen ab.
Aufgabe 5	Lade dir die Textversion (.txt) eines noch ungelesenen Werks der deutschen Literatur herunter, z. B. vom <a href="#">TextGrid Repository</a> oder von <a href="#">Wikisource</a> oder vom <a href="#">Projekt Gutenberg</a> . Speichere die Datei in deinem <i>OneDrive</i> Ordner <i>Distant Reading Vorname</i> .

<sup>1</sup> Diese erste Unterrichtseinheit basiert auf der Idee, die Anne McGrail im Mai 2017 in ihrem Blog *McGrail's Blog about Digital Humanities at the Community College* publiziert hat: [Doing DH at the CC](#).

Die zweite Grundlage ist die Lerneinheit [Visualisierung mit Voyant](#) von Marie Flüh, forText, Juni 2019.

<sup>2</sup> Eine kurze Einführung in die *Voyant Tools* findet sich im Post [Distant Reading mit Voyant](#).

Aufgabe 6	Öffne den Text und lösche am Anfang und am Schluss alle zusätzlichen Informationen wie die Lizenz, Titel oder Autor. Analysiert werden soll nur der eigentliche literarische Text ohne diese Metadaten.
-----------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

The screenshot shows the TextGrid Repository interface. At the top, there is a search bar containing 'Goethe Faust' and a search icon. To the right of the search bar are links for 'Regal 0', 'Anmelden', and 'English'. Below the search bar, there are dropdown menus for 'Explore' and 'Hilfe'. The main content area displays the title 'Faust. Der Tragödie erster Teil' in blue text. To the left of the title, there are metadata fields: 'Dateityp' (text/tg.edition+tg.aggregation+xn), 'Autor' (Goethe, Johann Wolfgang (GND)), and 'Publikationsdatum' (1948). To the right of the title, there are two buttons: '+ Zum Regal hinzufügen' and 'Herunterladen'.

Später kann es unter Umständen Sinn machen, wenn der Text in Segmente aufgeteilt wird, zum Beispiel in die fünf Teile, die den fünf Akten entsprechen, in die einzelnen Szenen oder in die drei oder vier Teile einer grösseren Erzählung. Diese können dann separat und im Vergleich zu anderen Szenen analysiert werden. Jede Szenennummer sollte dann unter dem Szenentitel abgespeichert und später hochgeladen werden. Der Volltext soll aber auch in einer einzigen Datei für die Analyse erhalten bleiben. Hier arbeiten wir nur mit einem Volltext!

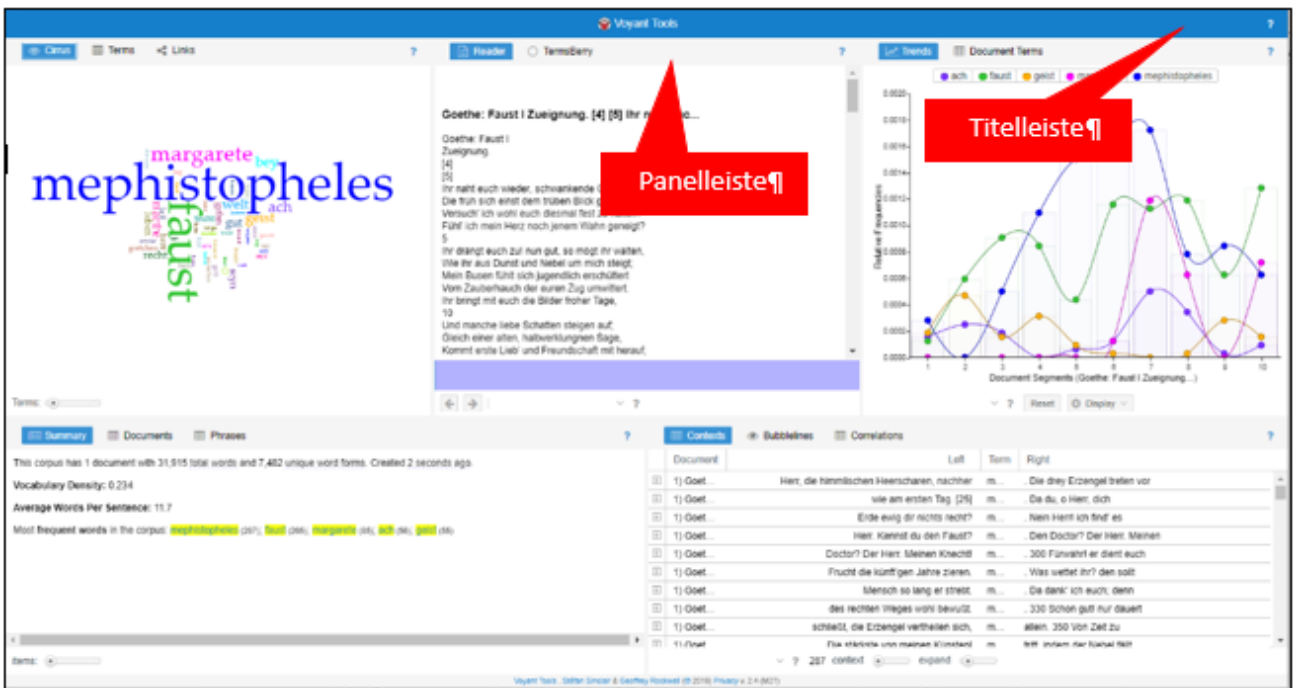
Aufgabe 7	Versuche nun Voraussagen über den Inhalt des noch nicht gelesenen Texts zu machen. Halte diese in deinen Notizen fest.
-----------	------------------------------------------------------------------------------------------------------------------------

### 3. Die *Voyant Tools*

Aufgabe 8	<p>Öffne die Startseite von <i>Voyant</i>: <a href="https://voyant-tools.org/">https://voyant-tools.org/</a>. Kontrolliere, ob die blaue <i>Reveal</i>-Schaltfläche sichtbar ist. Falls nicht, versuchst du es mit einem anderen Browser, in dem <i>Java Script</i> aktiviert ist.</p> <p>Nun wählst du deinen Text aus, kopierst ihn in das Textfeld <i>Add Texts</i> und lädst ihn mit <i>Reveal</i> hoch. Du könntest auch <i>Open</i> bzw. <i>Upload</i> wählen oder den URL einsetzen, der zum Volltext bei <i>Wikisource</i> oder dem <i>TextGrid Repository</i> führt, und dann mit <i>Reveal</i> hochladen lassen.</p>
-----------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

The screenshot shows the 'Add Texts' interface in Voyant Tools. At the top, there is a title 'Add Texts' and a search icon. Below the title, there is a text input field containing 'Goethe: Faust I'. Below the input field, there is a list of text elements: 'Zueignung.', '[4]', and '[5] Ihr naht euch wieder, schwankende Gestalten!'. At the bottom of the interface, there are three buttons: 'Open', 'Upload', and 'Reveal' (which is highlighted in blue and has a checkmark icon).

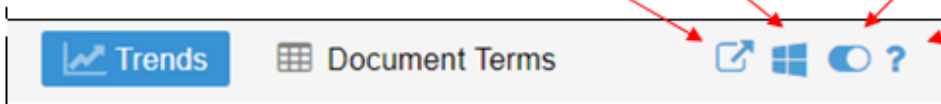
Jetzt öffnet sich das *Voyant*-Interface mit Visualisierungen der statistischen Auswertung deines Texts. Du siehst die fünf Tools *Cirrus*, *Reader*, *Trends*, *Contexts* und *Summary* in je einem Panel:



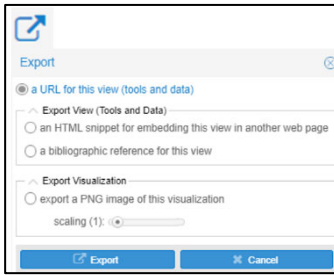
<p>Aufgabe 9</p>	<p>Hovere mit der Maus über die Titelleiste des <i>Voyant</i>-Fensters. Was für Einstellungen können hier vorgenommen werden?</p> <p>Hovere mit der Maus über die Menüleiste der Panels und schau sie dir an. In der rechten oberen Ecke erscheinen die Symbolleisten. Die <i>Voyant</i>-Tools agieren zusammen. Wenn gewisse Optionen mit einem Häkchen vor <i>apply globally</i> gesetzt werden, werden sie von den anderen Tools übernommen.</p>
------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Die Panel Funktionen

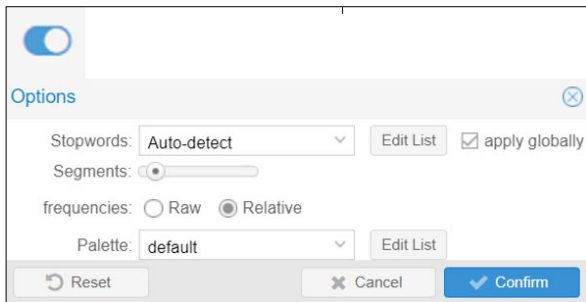
Die wichtigsten Funktionen in den Panels sind *Export*, *Toolwechsel*, *Optionen* und *Quickinfos*.



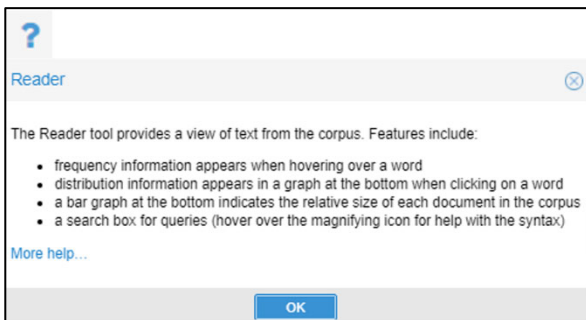
**Export:** Mit *Export a URL for this view (tools and data)* lässt sich eine URL für die gesamte *Voyant*-Sitzung exportieren, mit *Export View (Tools and Data)* ein Einbettungscode oder eine bibliographische Referenz und mit *Export Visualization* ein Bild als SVG- oder PNG-Datei. Die Arbeit auf *Voyant* kann nicht lokal, sondern nur über einen solchen URL gespeichert werden.



**Define Options:** Über *Define Options for this tool* lässt sich die Stoppwortliste bearbeiten. Stoppwörter sind Wörter, die bei der Textanalyse ausgeschlossen werden sollen, weil sie sehr häufig vorkommen und damit den Blick auf die bedeutungstragenden Nomen, Verben und Adjektive verstellen. Stoppwörter sind vor allem gewisse Partikel und Pronomen. Über *Edit list* können sie bearbeitet werden.



**Quickinfos:** Über das *Fragezeichen* kann man sich eine Quickinfo zum Tool anzeigen lassen:



**Toolwechsel:** Die Funktion *Click to choose another tool for this panel location* wird am Schluss dieser Lerneinheit besprochen.

Aufgabe 10	Öffne nun die Stoppwortliste und schau sie dir über <i>Edit list</i> an. Aktiviere über <i>Stopwords</i> die Stoppwortliste <i>German</i> .
Aufgabe 11	Exportiere eine URL, die zu dieser Seite führt und kopiere diese in deine Notizen. So kannst du später daran weiterarbeiten.

## Das Standard-Interface

**Cirrus:** Schau dir nun das Tool *Cirrus* mit der Wortwolke an. Die am grössten dargestellten Wörter kommen im Text am häufigsten vor. Die Farben und die absolute Position der Wörter haben keine Bedeutung. Wenn man mit der Maus über einzelne Wörter fährt, wird die Häufigkeit dieser Wörter

angezeigt. Mit **Scale** wechselt man vom ganzen Korpus zu den Einzeltexten. Im *Cirrus*-Panel kann man auf die Tools *Terms* und *Links* umstellen:

- **Terms** zeigt die zugrunde liegende Wortliste mit der Häufigkeit und einer Timeline der Fundstellen.
- **Links** erstellt ein Netzwerkdiagramm, grün: die Schlüsselwörter; orange: die damit verbundenen Begriffe.

Aufgabe 12	Dokumentiere das ganze Interface mit einem Screenshot in deinen Notizen.
------------	--------------------------------------------------------------------------

**Reader:** Hier siehst du den hochgeladenen Volltext. Wenn man mit der Maus über einzelne Wörter hovert, wird die Häufigkeit dieses Wortes im Text angezeigt. Über die Suchleiste kann man nach Wörtern suchen. Im *Reader*-Panel kann man auf das Tool *TermsBerry* umstellen:

- **TermsBerry** visualisiert mit farblichen Kennzeichnungen, welche anderen Wörter zusammen mit den Schlüsselwörtern (MFWs) vorkommen.

**Trends:** Dieses Panel zeigt das Vorkommen der fünf häufigsten Wörter. Jedes Wort hat eine eigene Farbe. Durch einen Klick auf ein Wort in der Legende wird es ausgeblendet, mit *Reset* kann man wieder alle Wörter anzeigen lassen. Hovert man über einen Knoten, erhält man Informationen zum jeweiligen Wort wie der Häufigkeit, den Titeln des Texts oder den Segmentnummern. Klickt man auf einen Knoten, wird eine Vernetzung zu *Contexts* und *Reader* hergestellt.

Im unteren Bereich kann man über ein Suchfeld Wörter eingeben, die dann im *Reader* hervorgehoben werden. Zeichenketten können in Anführungs- und Schlusszeichen gesetzt werden. Am besten schreibt man alle Wörter klein. Die Syntax erlaubt auch Jokerzeichen wie \* oder ^ oder ~. Der Stern \* steht beispielsweise für eine beliebige Zeichenkette. Über das Fragezeichen findest du dazu weitere Informationen.

Arbeitet man mit mehreren Texten, also mit einem Korpus, werden Abfragen in der #-Spalte die Textnummern vorangestellt, die sich auf die Nummern im *Documents*-Bereich beziehen. Dann werden sie untereinander in einer Art Fundstellenverzeichnis mit der Häufigkeit aufgelistet. *Trends* ist besonders interessant in der Zusammenarbeit mit dem *Context*-Panel. So können Schlüsselwörter in ihrem Kontext untersucht werden. Im *Trends*-Panel kann man auf das Tool *Document Terms* umstellen:

- **Document Terms** zeigt die Dokumentnummer, das Wort, die Häufigkeit und die Verteilung im Dokument.

**Summary:** Unten links wird im Bereich *Summary* eine Zusammenfassung der Analyse angezeigt: die Gesamtanzahl der Wörter, die Anzahl unterschiedlicher Wörter, die häufigsten Wörter (MFW), die durchschnittliche Satzlänge und die lexikalische Dichte des Texts. Wenn man hier auf den Reiter *Phrases* klickt, sieht man die häufigsten Wortkombinationen. *Summary* ist vor allem bei der Arbeit mit einem Korpus interessant.

In der Ansicht *Documents* können über den Link **Modify** Texte dazugeladen oder entfernt werden. So lassen sich Szenen oder andere Segmente vergleichen. Hat man mehrere Texte geladen, kann man hier mit *Shift* und einem *linken Mausklick* die unterschiedlichen Einzeltexte auswählen und so schnell vergleichen! Zudem findet man hier eine Such- und eine Downloadfunktion.

Aufgabe 13	Erstelle von diesem Panel ein PNG-Bild und füge dieses in deine Notizen ein.
------------	------------------------------------------------------------------------------

**Contexts:** Unten findet man das Werkzeug *Contexts*, auch *Keyword in Context (KWIC)*. Dieses Panel präsentiert die Konkordanzen eines Wortes: die häufigsten Stichwörter mit einem Teil des umgebenden Texts, also die Kontextwörter links und rechts neben dem Wort. Über das Plus-Symbol kann eine Zeile erweitert und mehr Kontext eingeblendet werden. Wählt man eine Fundstelle aus, wird sie im *Reader* im ganzen Kontext angezeigt. Im *Context*-Panel kann man auf die Tools *Bubblelines* und *Correlations* umstellen:

- **Bubblelines** zeigt das Textdokument als eine horizontale Linie, auf der die ausgewählten Wörter als unterschiedlich grosse Blasen dargestellt werden.
- **Correlations** gibt an, inwieweit die Wortfrequenzen sich synchron ändern.

*Reader* und *Contexts* sind eher Close-Reading Tools, *Cirrus*, *Trends* und *Summary* hingegen Distant Reading Tools.

Aufgabe 14	<p>Nimm dir nun Zeit, die Tools auszuprobieren. Wechsle innerhalb der Panels auch auf die anderen angebotenen Tools.</p> <p>Was leisten die Tools <i>Bubblelines</i> und <i>TermsBerry</i>?</p>
------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Externe *Voyant* Tools

*Voyant* besteht aus 29 unterschiedlichen Tools. Bis jetzt haben wir nur mit internen Tools gearbeitet. Externe *Voyant* Tools können einfach eingebettet werden durch einen *Klick auf Click to choose another tool for this panel location*.

So kannst du dir deine eigene Toollandschaft zusammenstellen. Die Tools sind in die folgenden Kategorien unterteilt:

- **Corpus Tools:** Tools für die Analyse eines Textkorpus, z. B. *Cirrus*, *Bubblelines*, *Trends*.
- **Document Tools:** Tools für die Analyse einzelner Texte, z. B. *Cirrus*, *Contexts*, *Reader*.
- **Visualizationtools,** z. B. *Cirrus*, *MicroSearch*, *TermsBerry*.
- **Grid Tools:** Tools, die Rastergrafiken und Tabellen erstellen, z. B. *Phrases*, *Documents*, *Topics*.
- **Other Tools,** z. B. *Reader*, *Summary*, *Veliza*.

Aufgabe 15	<p>Teste nun weitere Tools und versuche zu verstehen, was sie leisten. Orientiere dich an den Beschreibungen im <a href="#">Voyant Guide</a> oder an der Tool-Dokumentation: <a href="http://digihum.mcgill.ca/voyant/tools/">http://digihum.mcgill.ca/voyant/tools/</a>.</p> <p>Dokumentiere deine Arbeit in deinen Notizen so, dass du sie später deinen Kollegen kurz erklären kannst.</p> <p>Tausche dich mit zwei KollegInnen über eure Arbeit aus. Ergänze deine Dokumentation.</p>
------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------