

txtkit 1.0 Visual Text Mining Tool

www.txtkit.sw.ofcd.com

Copyright © 2004 Schoenerwissen/OfCD

txtkit ist ein visuelles textmining Werkzeug, das die inhaltliche und maschinelle Analyse von textueller Information visualisiert. Das Interface besteht aus dem txtshell Client und txtvbot. Mit Hilfe von txtshell und einiger weniger Befehle kann Textmaterial durchsucht, gelesen und selektiert werden, während txtvbot die Ereignisse und Inhalte dynamisch aufzeichnet. Die Visualisierung basiert auf *natural language processing*-Verfahren in Verbindung mit den Gebrauchsspuren aller Nutzer. txtshell und txtvbot können separat benutzt werden, aber erst in der alternierenden Wahrnehmung lassen sich visuelle und textuelle Erkenntnisgewinnung kombinieren und begünstigen so die digital gestützte Hypothesenbildung.

Hinweise zur Installation bitte der **readme**-Datei auf der CD entnehmen. Weitere Informationen und Updates befinden sich unter: www.txtkit.sw.ofcd.com

Die Publikation entstand im Projekt "KIT/ Kunst. Informatik. Theorie." (vormals: "Informatik, künstlerische Praxis und Kunsttheorie der digitalen Bildtechnologien", Leitung: Prof. Dr. Hans Ulrich Reck, Prof. Dr. Georg Trogemann), durchgeführt an der Kunsthochschule für Medien Köln. Das Projekt wurde im Rahmen des Programms der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung "Kulturelle Bildung im Medienzeitalter - KuBIM" (Förderkennzeichen A 6681 A) gefördert durch das Bundesministerium für Bildung und Forschung und das Ministerium für Schule, Wissenschaft und Forschung des Landes Nordrhein-Westfalen. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

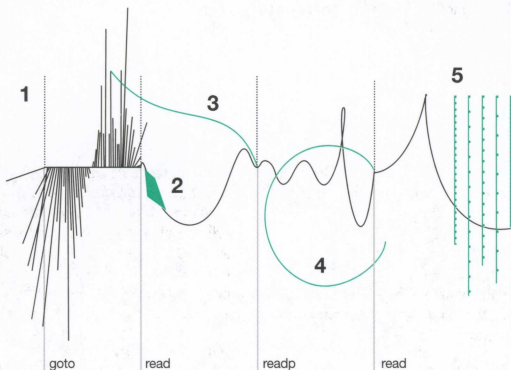
txtvbot zeichnet die Ereignisse und Inhalte der txthell auf und ergänzt sie durch die Eigenschaften der jeweiligen textuellen Information. Die Länge bzw. Form der grafischen Elemente beruht auf statistischen und inhaltlichen Parametern.

1

txtvbot liegt eine zeitliche Ordnung zugrunde: alle Ereignisse werden nacheinander aufgezeichnet. Der Zeitstempel markiert die Eingabe des jeweiligen Befehls.

2

Pinkfarbene Vektoren zeigen die Relevanz eines Schlüsselwortes, Satzes oder Paragraphen an. Je mehr Nutzer ein Schlüsselwort, Satz oder Paragraphen gelesen haben, desto wichtiger müssen sie sein, desto größer ist der entsprechende Vektor. txtkit berechnet den Relevanzwert und zeichnet den Vektor an entsprechender Stelle an.



3

Gelesene Goto-Einträge werden mit den korrespondierenden Sätzen verbunden.

4

Nutzer, die diesen Paragraphen gelesen haben, haben auch gesucht...

5

txtvbot sucht nach verwandten Sätzen.

Selektierte Sätze und Paragraphen werden blau eingefärbt.

Um zwischen Frontal- bzw. Seitenansicht zu wechseln, <v> drücken. Mit Hilfe der Pfeiltasten kann txtvbot räumlich bewegt werden.

Diese Funktionen sind nur aktiviert, wenn txtvbot die aktive Applikation ist! Hierfür mit dem Maus-Cursor auf txtvbot klicken.

txtkit 1.0 Visual Text Mining Tool

www.txtkit.sw.ofcd.com

Copyright © 2004 Schoenerwissen/OfCD

txtshell teilt sich in drei Bereiche: Statuszeile, Ausgabe/GUI und Commandline.

Mit der Eingabe der hier aufgeführten Befehle in die Commandline kann das Textmaterial durchsucht, gelesen und selektiert werden.

Mit Hilfe der **TAB**-Taste kann zwischen Commandline und der Ausgabe/GUI gewechselt werden. Der blinkende Cursor markiert den aktuellen Bereich: Commandline oder GUI.

Bitte auch die Hinweise in der Statuszeile beachten und im GUI-Modus die entsprechenden Pfeiltasten verwenden.

Das Ausführen der <enter> Taste im GUI-Modus startet 'read', 'select' bzw. 'readp', oder 'selectp'.

Alle Befehle können wiederholt und mit unterschiedlichen Argumenten ausgeführt werden.

goto startet eine Suchanfrage in Verbindung mit einem Schlagwort oder einer Phrase. Ist die Anfrage erfolgreich, werden die Ergebnisse aus der txtkit Datenbank abgerufen, nach Relevanzwerten geordnet und gekürzt aufgelistet. Beispiel: 'goto kunst als medientheorie'.

Erst wenn alle Ergebnisse geladen sind, springt der Cursor in die Commandline!

read zeigt die Ergebnissätze ungekürzt an. Hierfür 'read' in Kombination mit der entsprechenden Satznummer eingeben. Beispiel: 'read 6'.

readp gibt den Paragraphen in dem der gekürzte Satz erscheint, aus. Hierzu 'readp' in Kombination mit der entsprechenden Satznummer eingeben. Beispiel: 'readp 8'. Nach Auswahl eines einzelnen Satzes im Paragraphen kann durch das Drücken der Taste <g> im GUI-Modus eine Phrasensuche gestartet werden.

select vermerkt den vollständigen Satz im Pool. Beispiel: 'select 18'.

selectp vermerkt den vollständigen Paragraphen im Pool. Beispiel: 'selectp 18'.

pool zeigt alle selektierten Textstellen an. Mit den '+' und '-' Tasten kann die Reihenfolge der Sätze verändert werden.

savepool (+ Dateiname) speichert den Pool in der txtkit Datenbank, so dass dieser auch von anderen Nutzern gelesen und weiterverwendet werden kann.

getpools listet alle auf dem txtkit Server vorhandenen Pools auf. Anschliessend einzelne Pools zur Ansicht auswählen.

exportpool (+ Dateiname) speichert den Pool als Textdatei auf dem Schreibtisch.

exit beendet txtshell.

Credits

txtkit 1.0 Visual Text Mining Tool

www.txtkit.sw.ofcd.com

Copyright © 2004 Schoenerwissen/OfCD

Konzept: Schoenerwissen/OfCD, Hans Ulrich Reck

Entwicklung, Design und Programmierung: Schoenerwissen/OfCD

Mitarbeit Programmierung: Thomas Chille

Texte: Hans Ulrich Reck

Beratung: Hans Ulrich Reck, Georg Trogemann, Jochen Viehoff, Thomas Chille

Includes & Libraries:

SMYSQL Framework, Serge Cohen

<http://mysql-cocoa.sourceforge.net>

CHARVA, Rob Pitman

<http://www.pitman.co.za/projects/charva>

Cocoa Open Source, Karella Software, LLC

<http://cocoa.karella.com>

NSFont Category for OpenGL, Bryan L. Blackburn

<http://www.withay.com/macosex/opengl/nsfont.html>

Connector/J, MySQL AB

<http://www.mysql.com/products/connector-j>

ThreadWorker, Robert Harder

<http://iharder.sourceforge.net/macosex/threadworker>

This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version. This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details. You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA.