

Für den TXSTEP-Prototypen Oktober 2012 stehen folgende Dateien und Test-Skripte bereit, die zum größten Teil aus Kursmaterial der Tübinger Tustep-Kurse von W. Ott zusammengestellt wurden:

Dateien (Verzeichnis xstep/uebq)

inh	Inhaltsverzeichnis eines Sammelbandes, Tustep-Datei je 1 Satz pro Beitrag, z.B. Autor, Adam: Titel (1-11)
jnh	wie inh, jedoch mit Fehlern: teilweise fehlendes Blank nach Satzzeichen, falsche Reihenfolge der Beiträge
inh.xml	wie inh, jedoch xml-Datei, jedes Element in eigenem Satz. Tags: <inhalt> <beitrag> <a>... <t>...</t> <s>...</s> </beitrag> ... </inhalt>
inhb5.xml	wie inh.xml, jedoch bei mehrfach-Autoren jeder Autor zwischen <a> und
atsf	wie inh, Tustep-Datei, jedoch Bestandteile eines Eintrags vorn durch Kennungen =a= =t= =s= markiert, mit vielen Fehlern (doppelte Kennungen, fehlende Kennungen, falsche Reihenfolge der Kennungen, unbekannte Kennungen)
atsf.sdf	wie atsf, jedoch fdf-Datei
dedale	Tustep-Datei mit französische Wörtern, Reihenfolge der Wörter auf S. 1 entspricht den franz. Sortierregeln (keine vorbereitete Übungsaufgabe)
span	Tustep-Datei mit spanischen Wörtern, zwei Seiten mit identischem Inhalt, Sortierung in Seite 1 wie für die traditionelle Sortierung des Spanischen (keine vorbereitete Übungsaufgabe)
sans	Tustep-Datei mit Anfangs-Buchstaben von Sanskrit-Wörtern in der für die Sortierung relevanten Reihenfolge (keine vorbereitete Übungsaufgabe)
ibam	Schwätzer-Satire des Horaz (sermones I,9; Hexameter, mit Angabe der Versmetrik hinter jedem Vers: 0=spondeus, 1=dactylus; Gedichtnummer=Seitennummer, Versnummer=Zeilennummer)
ibamr	wie ibam, jedoch Gedicht- und Versnummer explizit in [...] vor jedem Vers
ibamr.fdf	wie ibamr, aber fdf-Datei

ibamt.fdf wie ibamr.fdf, aber ohne Metrik

lucas Anfang des Lukas-Evangeliums, griechisch, mit Akzentfehler auf dem 1. Wort (Gravis statt Akut)

lucas.xml wie lucas, aber xml-Datei

schella.x 3 Versionen eines Auszugs aus einer Schelling-Edition
 schellb.x (Tustep-Dateien mit xml-Tags).
 schellc.x Die Datei schella enthält den Editionstext / die Kollationsgrundlage, die Dateien schellb und schellc enthalten zu Demo-Zwecken veränderte Versionen des gleichen Textes.

schella.xml wie schella.x bzw. schellb.x, jedoch xml-Dateien
 schellb.xml

bibl.fdf Bibliographie 1982: Computer im Editionsweisen
 &.ulVorwort, &.ulTeil1, &.ulTeil2
 dort: @a Autor @t Titel @u unselbst. @s selbst.

bibl_see.fdf wie bibl.fdf, aber mit laufend nummerierten Einträgen (z.B. @n 1: @a Autor...) und Verweisen am Eintrags-Ende auf diese Nummern bei Publikationen aus dem gleichen Verlag bzw. der selben Zeitschrift (same publiser / periodical: @m 7 @m @m 50)

persons.xml Personenregister (Seck) mit Referenzen
 biographies.xml zu den Namen zugehörige Kurz-Infos

recker.xml Übungsdatei aus TEI-mail von Ute Recker-Hamm vom 26.4.11;
 Aufgabe: jedes Wort zwischen <w> und </w> setzen

TXSTEP-Skripte (Verzeichnis xstep/scripts):

Datei-Transformation:

xstep_lnr.xml Datei inh.xml: vor den Autornamen dreistellige laufende Nummer einfügen

inh2xml.xml Tustep-datei inh in eine xml-Datei umwandeln nach dem Muster der Datei inh.xml (s.o.).

inh2xml_a.xml wie inh2xml.xml, dabei gleichzeitig Autor, Adam in Adam AUTOR umwandeln

inh2xml_2.xml wie inh2xml.xml, andere Lösung.

inh2xml_2n.xml wie inh2xml_2.xml, aber Autoren nummerieren.

xml2inh.xml Datei inh.xml: Tags entfernen, je Eintrag eine Zeile nach dem Muster der Datei inh (s.o.).

name_pages.xml aus Datei inh eine xml-Datei erstellen, in der nur die Namen der Autoren und die Seitenzahlen genannt sind, ein Element je Autor, nach dem Muster
 <item><name><familyName>Holm</familyName>,
 <givenName>Bart E.</givenName></name>

	<pages>5-26</pages> </item>
vor_famnam.xml	Kopie der Datei inh erstellen, in der die Vornamen der Autoren vor den in Versalien umgewandelten Familiennamen stehen.
snr_rest.xml	Kopie der Datei inh erstellen, in der die Seitenangaben vor den Autorennamen stehen.
word_search.xml	Aus Datei inh die Beiträge, die im Titel eines von mehreren angegebenen Wörtern (hier: die Wörter "in" und "of") enthalten, in die erste von zwei Zieldateien kopieren, die übrigen in die zweite Zieldatei.
word_search_f.xml	wie word_search.xml, mit Fehlern
words_search.xml	Aus Datei inh die Beiträge, die im Titel alle angegebenen Wörter (hier: information und library) enthalten, in die erste von zwei Zieldateien kopieren, die übrigen in die zweite Zieldatei.
satzz_pruf.xml	Aus Datei inh die Einträge, die nach dem/den Satzzeichen kein blank enthalten, in die erste, die übrigen in die zweite Zieldatei kopieren.
atsf_pruf.xml	Kopie der Datei atsf erstellen, in der die fehlerhaften Einträge markiert sind
atsf_pruf_2.xml	Kopie der Datei atsf erstellen, in der die fehlerhaften Einträge durch entsprechende detaillierte Fehlerhinweise markiert sind (fehlende Kategorien; doppelte Kategorien; unbekannte Kategorien; falsche Reihenfolge).
atsf2xml.xml	Die korrekten Einträge aus der Datei atsf.sdf mit xml-Tags versehen in die erste von zwei Zieldateien kopieren, die fehlerhaften Einträge mit detaillierten Fehlerhinweisen (fehlende Kategorien; doppelte Kategorien; unbekannte Kategorien; falsche Reihenfolge) versehen in die zweite Zieldatei kopieren.
atsf2xml_b.xml	Wie agsf2xml.xml; zusätzlich sind Leerzeilen und Zeilen, die mit blanks beginnen, Anfang einer Texteinheit
ibam_silben2.xml	Wörter mit mehr als zwei Silben zählen, Anzahl am Zeilenende ausgeben
bibl_aut1.xml	In der Bibliographie beim zweiten und evtl. weiteren Beiträgen vom gleichen Autor den Namen des Autors ersetzen durch "(id.)"
recker_dol.xml	Wörter mit tags <w>...</w> versehen (Aufgabe Recker in der TEI-Liste vom 26.4.2011, nach dem von Recker gewünschten Ergebnis

recker_do.xml wie recker_dol.xml, jedoch Satzzeichen nicht zum Wort zählen

Vergleichen und Weiterverarbeiten:

ve.xml Wortweises Vergleichen der Dateien schella.x und schellb.x, Erstellen eines Vergleichsprotokolls, Abspeichern der Unterschiede der beiden Dateien in der Syntax von Korrekturanweisungen des Tustep-Programms KAUSFUEHRE in der zu corrections angegebenen Datei; Wortlaut und Umgebung des Lemmas (= vom Unterschied betroffene Wörter der VersionA) und Stellenangabe der Varianten in VersionB mit ausgeben.

ka.xml Erstellen einer Kopie (Name: schellk.x) der Datei schella.x, in der die Korrekturanweisungen der bei ve.xml zu corrections angegebenen Datei ausgeführt sind (die Kopie muss also inhaltlich identisch sein mit der bei ve.xml als versionB abgegebenen datei schellb.x).

cmp.xml Wortweises Vergleichen der Dateien schella.xml und schellb.xml, Erstellen eines Vergleichsprotokolls, Abspeichern der Unterschiede in einer an TEI angelehnten Syntax, die gleichzeitig alle Informationen enthält, die für die Korrekturanweisungen des Tustep-Programms KAUSFUEHRE notwendig sind.

cmp_n.xml Wie cmp.xml, dabei großgeschriebene Umlaute als gleichwertig mit Ae, Oe, Ue behandeln (d.h. Wörter, die sich nur darin unterscheiden, nicht als Unterschiede ausgeben).

vex.xml Wortweises Vergleichen der Dateien schella.xml und schellb.xml, dabei großgeschriebene Umlaute als gleichwertig mit Ae, Oe, Ue behandeln (d.h. Wörter, die sich nur darin unterscheiden, nicht als Unterschiede ausgeben), ebenso die Wörter "als" und "wie" als gleichwertig behandeln (nur als Demo, in der Praxis kaum sinnvoll); Erstellen eines Vergleichsprotokolls, Abspeichern der Unterschiede in der Syntax von Korrekturanweisungen des Tustep-Programms KAUSFUEHRE in der zu corrections angegebenen XML-Datei; Wortlaut und Umgebung des Lemmas (= vom Unterschied betroffene Wörter der VersionA) und Stellenangabe der Varianten in VersionB mit ausgeben. Ausgabe mit minimalem XML-markup.

ve_tuscr.xml wie ve.xml, mit tustep- und tuscript-Einschüben und weiteren richtigen <correct> und unvollständigen <compare> TXSTEP-Anweisungen dahinter. Soll das Zusammenspiel von TXSTEP, TUSTEP und TUSCRIPT zeigen.

vexx.xml wie vex.xml, jedoch Abspeichern der Unterschiede als Lesarten in einer an TEI angelehnten Syntax,

die gleichzeitig alle Informationen enthält,
die für die Korrekturanweisungen des
Tustep-Programms KAUSFUEHRE notwendig sind.

- kax.xml Erstellen einer Kopie (schellk.xml) der Datei schella.xml, in der die Korrekturanweisungen der bei vexx.xml zu corrections angegebenen Datei ausgeführt sind (die Kopie muss also inhaltlich identisch sein mit der bei vexx.xml als versionB abgegebenen datei schellb.xxx, bis auf die Unterschiede, die bei vexx.xml nicht erhoben wurden: Versal-Umlaute vs. Ae,Oe,Ue und "als" vs. "wie").
- precorr.xml Zur halbautomatischen Korrektur nach Doppelerfassung werden in der diff-Datei die Einträge (manuell) markiert, die in der ersten Fassung richtig sind. Vor der (automatischen) Ausführung der Korrektur werden die Einträge in zwei Dateien aufgeteilt: nicht-markierte Korrekturen und markierte. Nur erstere werden (hier nicht mehr gezeigt) für "correct" genutzt.
- vexx_2.xml Vergleichen der Datei schellb.xxx mit der durch kax.xml erstellten Datei schellk.xml; in der zu corrections angegebenen Datei schellk-diff.xml dürfen jetzt nur Unterschiede bezüglich der Versal-Umlaute vs. Ae,Oe,Ue und "als" vs. "wie" nachgewiesen sein.
- transform.xml Die Unterschiede, die durch ve.xml in der Datei schell.diff abgespeichert wurden, als Vorbereitung für einen kritischen Apparat sichten und aufteilen:
Unterschiede, die nur Versalumlaute vs. Ae Oe Ue betreffen, in die Datei schellv
Unterschiede, die nur Satzzeichen betreffen, in die Datei schells
Unterschiede, die nur Orthograpica wie th und t, c und k, y ind i betreffen, in die Datei schello
Andere Varianten in die Datei schella (= Apparat).
- sep_variants.xml Wie transform.xml, Ausgabe in xml-Dateien
- ins.xml Einfügen von Texten anhand von Kürzeln (siehe unten)
- Register-Arbeit, Sortieren:
- autorenliste.xml alphabetische Liste der Autoren aus der Datei inh.xml als xml-Datei, Seitenzahlen als Referenz.
- autreg_inh.xml alphabetische Liste der Autoren + Seitenzahlen aus aus Tustep-Datei inh als Tustep-Datei, xml-tags <eintrag> ... </eintrag> um jede Zeile. Zieldatei nicht überschreiben, sondern neue Liste hinten anhängen.

- inh_alpha.xml Einträge der Datei inh.xml alphabetisch nach dem ersten Autor sortieren; bei mehreren Autoren: Verweiseinträge vom zweiten und dritten Autor auf den ersten erzeugen und einsortieren.
- inhx_stichw.xml Alphabetisches Verzeichnis aller mit t beginnenden Titelstichwörter aus der Datei inh.xml
- rv_greek_s.xml alphabetischer Wortformenindex zur Datei lucas
- rv_greek_x.xml dto., aber zu Datei lucas.xml
- rv_greek_xs.xml dto., mit Satzausgabe
- ibamr_metren.xml Verzeichnis der metrischen Schemata (Folge von Daktylen und Spondeen; Zeichenfolge hinter @ in Datei ibam bzw. ibamr
- ibamr_metren2.xml Verzeichnis der metrischen Schemata (Folge von Daktylen und Spondeen; Zeichenfolge hinter @ in Datei ibam bzw. ibamr), sortiert nach Zahl der Daktylen (= "langsame" Verse nach vorn)
- ibamr_freq.xml zwei Verzeichnisse der in Datei ibamr vorkommenden Wörter:
1. alphabetisch sortiert, mit Referenzen
2. nach Häufigkeiten sortiert, ohne Referenzen
- ibams_freq.xml zwei Verzeichnisse der in Datei ibamr.fdf vorkommenden Wörter:
1. alphabetisch sortiert, mit Referenzen
2. nach Häufigkeiten sortiert, ohne Referenzen (wie ibamr_freq.xml, aber Quelle=fdf-Datei und mit Kommentaren)
- ins.xml Personenregister (mit Referenzen) anreichern durch Kurzinfos zu den Personen
- ed_bibl3.xml Die beiden Teile der Bibliographie bibl.fdf jeweils nach Erscheinungsdatum sortieren.
- ed_bibl3see.xml Die beiden Teile der Bibliographie bibl-see.fdf jeweils nach Erscheinungsdatum sortieren, neu durchnummerieren und Verweise aktualisieren.
- Leerer Rahmen für TXSTEP-Skript
- xstep_rahmen.xml Datei, die als Rahmen für ein TXSTEP-Skript dienen kann