

TextGrid Komponenten

Version 10.12.2006

Arbeitspaket AP2, AP3, AP5

verantwortlicher Partner:

AP2 - Darmstadt / Saphor

AP3 - SUB / DAASI

AP5 - Trier

TextGrid

Modulare Plattform für verteilte und kooperative
wissenschaftliche Textdatenverarbeitung -
ein Community-Grid für die Geisteswissenschaften



Bundesministerium
für Bildung
und Forschung

Projekt: **TextGrid**

Teil des D-Grid Verbundes und der deutschen e-Science Initiative

BMBF Förderkennzeichen: 07TG01A-H

Laufzeit: Februar 2006 - Januar 2009

Dokumentstatus: vorläufig, wächst mit dem Projekt

Verfügbarkeit: öffentlich (detaillierte Spezifikationen werden vorläufig intern geführt)

Autoren: alle TextGrid Partner

1 TextGrid-Tools

Im Folgenden sind eine Reihe von Tools in TextGrid aufgelistet, die in den "TextGrid Szenarien" identifiziert wurden. Diese Liste deckt nicht alle Tools ab, die in TextGrid eingebunden werden könnten, sondern gibt lediglich den vorläufigen Stand wieder, welche Tools in TextGrid eingebunden werden und welche - je nach **Priorität** - in weiterer Folge denkbar sind.

Diese Aufstellung ist nicht nach unterschiedlichen Disziplinen geordnet, da mehrere Tools in verschiedenen Disziplinen angewendet werden. Vielmehr ist die Ordnung (aus einer eher technischen Sicht) nach **Tooltyp**, und weitergehend nach der **Rolle**, die das Tool verwendet, bzw. nach dem Anwendungsbereich (z.B. Annotation, Publikation) strukturiert. In den folgenden Kurzbeschreibungen werden die Tools mit den Arbeitsschritten und mit anderen Tools vernetzt.

1.1 Prioritäten zur Umsetzung

TextGrid ist lebendig und kann wachsen, sollten Kooperationen eine Erweiterung des Leistungsspektrums erlauben. Die TextGrid Tools sind im Folgenden nach ihrer Priorität zu Umsetzung aufgeschlüsselt:

- **Priorität 1 / zentral** - ist ein integraler Bestandteil, war evtl. schon Teil des Antrages und wird auch sofort umgesetzt
- **Priorität 2 / erstrebenswert** - ist nicht absolut essenziell, und wird (hoffentlich) noch im Laufe der weiteren Projektlaufzeit angegangen
- **Priorität 3 / optional** - übersteigt die Ressourcen des Projektes, enthält aber spannende Themen bzw. wichtige Anknüpfungspunkte. Dahingehende zukünftige Aktivitäten und Kooperationen werden verfolgt (dies ist ein besonders dynamischer Bereich, der nur einige Möglichkeiten aufzeigen kann)

1.2 Tooltyp

- *interaktive Tools* haben ein Benutzerinterface und sind auf Benutzereingaben ausgerichtet [Benutzerschnittstelle]
- *Streaming Tools* bestehen aus zwei Teilen: einem im Batchbetrieb steuerbaren "Enactor" und einem GUI-Frontend, mit dem Aufträge (bzw. Konfigurationen) für den Enactor entwickelt werden können. Die GUI-Komponenten sind eng mit den interaktiven Tools integriert. Die Enactors können jeweils sowohl im Rahmen eines Workflows durch den Workflow-Enactor als auch direkt aus ihren jeweiligen GUI-Frontends angestoßen werden. [Service Layer, Benutzerschnittstelle]
- *Basistools* - wie das Recherche-Tool haben Komponenten in allen Schichten [Middleware, Service Layer, Benutzerschnittstelle]
- *Hilfstoools*, werden nur als Bestandteil eines anderen Werkzeugs aufgerufen ohne eine eigene Benutzerschnittstelle anzubieten [Service Layer oder Middleware]

1.3 Übersicht

1.	Recherchetool	5
2.	XML-Editor	5
3.	Workflow-Editor	6
4.	Datei- und Rechtemanagement	6
5.	grafischer Link-Editor	6
6.	Metadaten-Annotation	7
7.	Bibliografietool	7
8.	Tokeniser	8
9.	Sortierer	8
10.	Streaming-Editor	8
11.	Lemmatisierer	9
12.	Kollationierung	9
13.	Publisher für den Druck	10
14.	Publisher für das Web	10
15.	OCR-Erkennung und Nachbearbeitung	10
16.	Katalogsuche	11
17.	Listen-Annotation	11
18.	Kommentierungstool	12
19.	Part-of-Speech Tagging	12
20.	Word Sense Disambiguation	12
21.	Kookkurrenzanalyse	13
22.	Kollokationsanalyse	13
23.	Migration von Textverarbeitungsformaten in TEI	14
24.	Wörterbuch-Recherchetool	14
25.	Wörterbuch Link-Editor	15
26.	statistisches Wörterbuch-Analysetool	15
27.	Wörterbuch-Redaktionssystem	16

1.4 Tools

1. Recherchetool

Priorität 1

Relevant in Arbeitsschritt/en:

- Bestandsaufnahme (Edition)
- Recherche und Analyse (Korpuslinguistik)
- Nutzung

Beschreibung:

Das Recherchetool umfasst Werkzeuge zum Zugang, zur Suche und zum Blättern in den im Grid gespeicherten Texten und Metadaten oder zu Teilmengen davon. Zusätzlich können diverse statistische Analysen durchgeführt werden: In wievielen Texten wird ein bestimmtes Wort verwendet? Wie viele Autoren verwenden die Phrase *xy*?

vernetzt mit:

- Publisher Web (Einbindung der Suchmasken; Navigation)
- XML-Editor (und allen weiteren GUI-Tools)
- Link-Editor
- und anderen

2. XML-Editor

Priorität 1

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung (Edition)
- manuelle Transkription (Digitalisierung)

Beschreibung:

Der XML-Editor bildet eine der zentralen Komponenten des TextGrid-Clients: Er ist ein interaktives Tool, das zur Neuerfassung wie zur nachträglichen (manuellen) Annotation von Texten in XML-basierten Formaten dient. Darüberhinaus wird er mit anderen TextGrid-Tools in der *Rich Client Platform* integriert und kann so etwa für Streaming-Tools als Vorschau und zur Ergebnisdarstellung verwendet werden.

vernetzt mit:

- grafischer Link-Editor

- Metadaten-Annotation
- Workflow-Editor
- Suche
- Streaming-Editoren
- ... (und andere, die aber wiederum über die obigen angesprochen werden)
- direkte Einbindung der Wörterbuch-Suche (Doppelklick auf ein Wort)

3. Workflow-Editor

Priorität 1

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung (Edition)

Beschreibung:

Der Workflow-Editor ist ein interaktiver Bestandteil der Benutzerumgebung und erlaubt es, die Automatisierung von Arbeitsabläufen (Workflows) zu definieren (orchestrieren). Neben den Servicekomponenten der Streaming-Tools können dabei auch interaktive Prozesse in einen Workflow eingebunden werden: In diesem Falle wird der Bearbeiter etwa per E-Mail oder IRC benachrichtigt, dass nun ein interaktiver Prozess erreicht ist, und es wird eine einfache Möglichkeit angeboten, die Benutzerschnittstelle entsprechend dieser aktuellen Aufgabe zu konfigurieren.

vernetzt mit:

- die Streaming-Komponenten wie den **Streaming-Editor**, **Lemmatisierung**, **Kollationierung**, und **Sortierung** können über den Workflow-Editor verknüpft werden.

4. Datei- und Rechtemanagement

Priorität 1

Beschreibung:

Dieses Tool sind ggf. zwei: eines zum Verwalten der Projektdaten (Verknüpfung mit der Metadaten-Annotation), und eines zum Erstellung der Projektgruppen (Virtual Organisations) und Rechteverteilung.

5. grafischer Link-Editor

Priorität 1

Relevant in Arbeitsschritt/en:

- manuelle Transkription (Digitalisierung)
- Auszeichnung und Anreicherung (Edition)

- Publikation (Edition) [evtl. nicht in der ersten Version]

Beschreibung:

Als separates Tool erlaubt der **grafische Link-Editor** die Verknüpfung von Bildsegmenten und der Transkription (d.h. ein Bildsegment ist ein Wort). Wichtig ist hier die Zuordnung auf verschiedensten Ebenen: grafisch - Zuordnung Segment und transkribiertes Wort, annotationstechnisch - Einbettung in entsprechendes Markup.

6. Metadaten-Annotation

Priorität 1

Relevant in Arbeitsschritt/en:

- Bestandsaufnahme (Edition)

Beschreibung:

Eintragen von stark strukturierten Daten insbesondere den Metadaten wie Autor, Titel usw. Struktur der Datenfelder und Validierungsschema sollen frei definierbar sein. Konversion zwischen verschiedensten Quellen und Export sollen möglich sein (XML-Daten nach TEI, Tiff-Header).

vernetzt mit:

- Bibliografiertool
- Recherche

7. Bibliografiertool

Priorität 1

Relevant in Arbeitsschritt/en:

- Bestandsaufnahme (Edition)

Beschreibung:

Ein Modul für die Verwaltung bibliographischer Daten, wie sie nicht nur in textwissenschaftlichen Projekten auftauchen. Es wird hochverteilte und kollaborative Bibliographien unterstützen.

Daten können in verschiedenen Formaten (TEI, DocBook, TeX, Word, OpenOffice, TUSTEP etc.) und Formatierungskonventionen ausgeliefert werden, bzw. auch mit den Dokumenten verknüpft werden (siehe BibTeX, Endnote).

vernetzt mit:

- (bezieht Daten aus) Katalogsuche

- (liefert Daten an) Metadaten-Annotation
- (liefert Daten / Referenzen an) XML-Editor
- Publisher Print, Publisher Web

8. Tokeniser

Priorität 1

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung

Beschreibung:

Zerlegt Texte in Wörter nach den jeweils sprachüblichen Regeln und dem Verständnis von Wort und Wortgrenze.

vernetzt mit:

- Lemmatisierer und andere Streaming-Tools

9. Sortierer

Priorität 1

Relevant in Arbeitsschritt/en:

- Publikation (für Register)
- Lexikografie (bei der Extraktion von sortierten Wortlisten aus Primärtexten)

Beschreibung:

Anordnung von gegebenen Zeichenketten gemäß kulturellen und fachlichen Erwartungen. Transformation ungeordneter Zeichenketten in geordnete. Konventionen vom Nutzer vorgebar. Bekannte nationale und europäische Standards, etwa DIN 5007, NFZ44-001 (AFNOR), TK 34.1 (SIS) und ENV 13710 (CEN) werden direkt unterstützt.

Bestimmte Wörter können durch Angabe einer Stoppwortliste aus der zu sortierenden Liste herausgefiltert (sort), bzw. auch mehrfach vorkommende Wörter gestrichen werden (uniq).

vernetzt mit:

- Publisher Web, Publisher Print

10. Streaming-Editor

Priorität 1

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung (Edition)

Beschreibung:

Transformationen von Dateien aufgrund von Regeln, z.B. automatisierte Anreicherung potentiell unstrukturierter Texte mit XML-Strukturen. Die Eingabe muss nicht XML sein, sondern kann ein beliebiger Datenstream sein (etwa OCR-Rohdaten, reiner Text), Ausgabe wird üblicherweise, muss aber nicht XML sein. Das System wird allerdings auf die Ausgabe von XML-Daten optimiert sein. Module werden Streams aus ausgewählten Anwendungsformaten unterstützen.

vernetzt mit:

- Listen-Annotation

11. Lemmatisierer

Priorität 1

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung

Beschreibung:

Automatische Reduktion von flektierten Wortformen auf die Grundform (Lemmatisierung) mit automatischer Grundlemmatisierung und manueller Korrektur / Überarbeitung. Kompositaanalyse für zusammengesetzte Wörter.

vernetzt mit:

- Tokeniser (wird vorausgesetzt)
- Part-of-Speech-Tagging (die Algorithmen POS Tagging und Lemmatisierung könnten sich gegenseitig ergänzen)
- (Wörterbuch-) Recherchetool (zu suchende Wörter könnten als Art der Query-Expansion vor der Suche lemmatisiert werden)

12. Kollationierung

Priorität 1

Relevant in Arbeitsschritt/en:

- Erfassung genetischer Informationen (Edition)

Beschreibung:

Zwei bis beliebig viele TEI-kodierte Dokumente werden verglichen und ihre Unterschiede werden nach TEI kodiert und notiert.

13. Publisher für den Druck

Priorität (wird in einzelnen Modulen definiert)

Relevant in Arbeitsschritt/en:

- Publikation (Edition)

Beschreibung:

In diesem Arbeitsmodul werden TEI-Daten für den Druck vorbereitet. Das Textsatzsystem kann mit kritischen Apparaten und mathematischen Formeln umgehen. Wo möglich wird es im Batch-Modus arbeiten; wo angebracht ist die Komponente interaktiv zur typografischen Umsetzung von Tags und zur Verfeinerung der Typografie (etwa Hinweis auf übervolle Zeilen und Möglichkeit, Trennkennstellen in den Text einzubringen).

vernetzt mit:

- Sortieren (für Register etc.)
- Bibliografie

14. Publisher für das Web

Priorität (wird in einzelnen Modulen definiert)

Relevant in Arbeitsschritt/en:

- Publikation (Edition)
- Nutzung

Beschreibung:

Ziel ist die Aufbereitung von TEI-Daten für die wissenschaftlich adäquate Publikation im WWW.

vernetzt mit:

- Recherche
- Sortieren
- Bibliografie

15. OCR-Erkennung und Nachbearbeitung

Priorität 1

Relevant in Arbeitsschritt/en:

- Transkription (Digitalisierung)

Beschreibung:

Prototypische Einbindung eines OCR-Moduls. Es soll kein Business Model mit Accounting/Billing Funktionalität geschaffen werden, sondern lediglich die FineReader API als Grid Service in TextGrid eingebunden werden, und prototypisch einige Dokumente damit verarbeitet werden.

16. Katalogsuche

Priorität 3

Relevant in Arbeitsschritt/en:

- Bestandsaufnahme (Edition)

Beschreibung:

Zur Unterstützung des [Projektleiters](#) in der Quellensuche können externe Kataloge an TextGrid angebunden werden (z.B. ZvDD, Kalliope, siehe "externe Services" in der Wiki). So eine Meta-Suche über die bibliografischen Metadaten in externen Katalogen könnte mit der TextGrid-internen Suche ([Retrieval](#)) verknüpft sein, und damit auch auf Ressourcen verweisen, die sich tatsächlich in TextGrid befinden. Eine saubere Trennung der Suchinterfaces bzw. deutliche Kennzeichnung des Suchraums ist dabei aber notwendig. Zugriff auf die [Katalogsuche](#) wie die das [Retrieval](#) besteht für den [Projektleiter](#) wie auch den [Benutzer](#).

vernetzt mit:

- Bibliografietool

17. Listen-Annotation

Priorität 2

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung (Edition)

Beschreibung:

Auf Basis von vorgefertigten Listen (z.B. Plain Text jede Zeile ein Wort) können Texte ausgezeichnet werden. Hierbei wird jedes Wort, das in dem Text vorkommt, mit einem definierten Tag annotiert. Mögliche Anwendungsfälle sind Eigennamenerkennung (Named Entity Recognition) oder Terminologieextraktion.

vernetzt mit:

- evtl. mit dem Streaming-Editor - überlappen von der Funktionalität her oder könnten über ein Benutzerinterface bedient werden

18. Kommentierungstool

Priorität 3

Relevant in Arbeitsschritt/en:

- Auszeichnung - Erschließungshilfen (Edition)
- manuelle Transkription (Digitalisierung)

Beschreibung:

Mit einem **Kommentierungstool** können **Projektleiter** und **Bearbeiter**, und ggf. auch authentifizierte (d.h. eingeloggte) **Benutzer** vorübergehend Notizen, Lesezeichen und dergleichen auf bestimmte Textteile legen. Der Austausch von Notizen und Lesezeichen zwischen unterschiedlichen Benutzern wäre u.U. ein hervorragendes Mittel zur Kollaboration. Ob die Zielgruppe (d.h. Textwissenschaftler) so ein kollaboratives Tool annehmen würden und was ihre Anforderungen daran wäre, müsste vorher ausgelotet werden.

vernetzt mit:

- XML-Editor
- Publisher Web

19. Part-of-Speech Tagging

Priorität 3

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung (Edition)
- Erschließungshilfe - Wörterbücher

Beschreibung:

Automatische Wortartenklassifizierung für flektierte und unflektierte Wortformen.

vernetzt mit:

- Tokeniser (Voraussetzung)
- Lemmatisierer (siehe dort)

20. Word Sense Disambiguation

Priorität 3

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung (Edition)

Beschreibung:

Mittels Word Sense Disambiguation (Sense Tagging) wird der Sinn eines Wortes aus seinem Kontext heraus festgestellt, wenn das Wort für sich verschiedene Bedeutungen haben könnte. Es gibt vielfältige Methoden zur Word Sense Disambiguation, die auch mit **Part-of-Speech Tagging** und syntaktischen Analysen interagieren können. Analysetools und der Kookkurrenzanalyse könnten die Ressourcen in TextGrid nutzen, um die Zuverlässigkeit von Word Sense Disambiguation zu erforschen.

vernetzt mit:

- Part-of-Speech Tagging

21. Kookkurrenzanalyse

Priorität 3

Beschreibung:

Automatische Anzeige der Kontexte, in denen ein bestimmtes Wort vorkommt. Interessant für die Erschließung der Wortlesart: Welche assoziierten Wörter treten mit diesem Wort auf? In welchen syntaktischen Konstruktionen wird das Wort gebraucht?

vernetzt mit:

- statistisches Wörterbuch-Analysetool - die Algorithmen des statistischen Wörterbuch-Analysetools und der Kookkurrenzanalyse könnten in einigen Bereichen überlappen. Vielleicht kann man das Wörterbuch-Analysetool soweit generalisieren, dass es auch Kookkurrenzanalyse in Primärtexten vornehmen kann

22. Kollokationsanalyse

Priorität 3

Beschreibung:

Tool zur automatischen Erkennung von festen Wortgefügen (z.B. "Rotes Kreuz"). Die Bedeutungserschließung ist nicht-kompositionell, d.h. die Lesart entspricht nicht der Summe der Teile. Interessant für die Erkennung von Mehrwortausdrücken.¹

vernetzt mit:

- Kookkurrenzanalyse (evtl. überlappende Algorithmen)

¹ Hier gibt es Anknüpfungspunkte an das Kovalevskaja-Projekt von Natalia Filatkina zur historischen Phraseologie; sie gehört zu unseren Kooperationspartnern.

- Lemmatisierer (evtl. zur Erkennung von Komposita)

23. Migration von Textverarbeitungsformaten in TEI

Priorität 3

Relevant in Arbeitsschritt/en:

- manuelle Transkription (Digitalisierung)

Beschreibung:

Vorhandene, in Textverarbeitungssoftware geschriebene Texte werden ins TEI-Format konvertiert, um sie in TextGrid einzubringen.

Der Originaltext wird von seiner Formatierungsinformation getrennt und in ein XML-Format überführt. Auf dieser Ebene werden also Formatierungsinformationen (Paragraph, Quotes, etc.) explizit gekennzeichnet. Dazu gehört die Rückübersetzung der Escape-Sequenzen in Umlaute.

vernetzt mit:

- XML-Editor
- Streaming-Editor

24. Wörterbuch-Recherchetool

Priorität 1

Relevant in Arbeitsschritt/en:

- Auszeichnung und Anreicherung (Edition)
- Erschließungshilfe - Wörterbücher
- Auszeichnung - Erschließungshilfen : Namensregister, Werkregister, Chronik, Glossar, Wörterbücher

Beschreibung:

Das Wörterbuch-Recherchetool bietet einen Zugang zu den Wörterbuch-Inhalten über eine dezidierte Suchmaske, die eine ausdrucksseitige und eine inhaltliche Suche bietet (bzw. eine Eingrenzung anhand der Strukturinformation ermöglicht; vgl. Mhd. Wörterbuchverbund). Besondere Anforderungen stellt aufgrund der Heterogenität und unterschiedlichen Auszeichnungstiefe der vorliegenden Wörterbücher dabei die Volltextsuche in bestimmten Kategorien. So sind beispielsweise Belegzitate, Eigennamen und Bedeutungsangaben nicht in allen Wörterbüchern ausgezeichnet; will man sich nicht auf den kleinsten gemeinsamen Nenner beschränken, grenzt man also entweder einige Wörterbücher aus oder erhält unscharfe Ergebnisse (dies wird der Regelfall sein).

vernetzt mit:

- XML-Editor
- Publisher Web

25. Wörterbuch Link-Editor

Priorität 2

Relevant in Arbeitsschritt/en:

- Erschließungshilfe - Wörterbücher

Beschreibung:

Mit dem Wörterbuch Link-Editor kann ein **Bearbeiter** zwischen den einzelnen Lemmata Verknüpfungen erstellen, z.B. für die **onomasiologische Suche**. Er kann die Ausgabe der **statistischen Wörterbuch Analyse** und entsprechend codierte im Druckwerk explizit angeführte Verweise direkt einlesen. Diese Funktionalität ist essenziell für die Erstellung einer Metalemmaliste, bzw für die Erstellung einer Metalemma-Ontologie.

Der Wörterbuch Link-Editor könnte in ein **Wörterbuch-Redaktionssystem** bzw. - Verwaltungstool integriert werden, um die manuelle Einrichtung von Querverweisen zu erleichtern.

vernetzt mit:

- Wörterbuch-Redaktionstool
- statistische Wörterbuch Analyse

26. statistisches Wörterbuch-Analysetool

Priorität 2

Relevant in Arbeitsschritt/en:

- Erschließungshilfe - Wörterbücher

Beschreibung:

Das statistische Analysetool richtet mittels automatischer statistischer Verfahren Verweise innerhalb eines Wörterbuchs zwischen verwandten Lemmata ein. Dazu werden (wie unten beschrieben) die Wortvektoren einzelner Einträge untereinander verglichen und statistisch ausgewertet. Die Ergebnisse dieses automatischen Tools werden direkt in den **Wörterbuch Link-Editor** eingeführt, um dort manuell bereinigt und weiterverarbeitet zu werden.

(Dieses Verfahren könnte auf Textkorpora angewandt eine **Kookkurrenzanalyse** ermöglichen. Hierbei hätte jedes Dokument einen Wortvektor, und der Vektorraum wäre ein ganzes Korpus. Durch Änderung der Grundeinheiten für dieses Tool könnten somit jeweils andere Analysen mit nahezu identischen statistischen Algorithmen durchgeführt werden: es können jeweils alle Werke eines Autors aus einer bestimmten Schaffensperiode mit den Werken aus

einer anderen Periode verglichen werden, um die Veränderung des Wortschatzes nicht von Werk zu Werk sondern zwischen Schaffensperioden aufzuzeigen. Ebenso sind Vergleiche zwischen Autoren oder zwischen verschiedenen Epochen denkbar. - Diese Generalisierung des "statistische Wörterbuch-Analysetool" zu einem "statistischen Analysetool" sollte geprüft werden.)

vernetzt mit:

- Wörterbuch-Linkeditor

27. Wörterbuch-Redaktionssystem

Priorität 3

Relevant in Arbeitsschritt/en:

- Wörterbücher
- Erschließungshilfe - Wörterbücher

Beschreibung:

Eine Arbeitsumgebung, die die Arbeit des klassischen Lexikografen effizienter gestaltet bzw. neue Möglichkeiten eröffnet und von der Zusammenstellung und Analyse des Korpus bis zum fertigen Artikel und dessen Aufbereitung für den Druck bzw. die online-Präsentation sämtliche Arbeitsschritte unterstützt. Vgl. entsprechenden Subcommunity-Antrag.]

Das Terminologietool ermöglicht einem **Bearbeiter** die Verwaltung von ein- oder mehrsprachigen Wörterbucheinträgen und damit die Erstellung von neuen Wörterbüchern aus existierenden Korpora. Der **Wörterbuch Link-Editor** könnte eine Komponente eines Wörterbuch-Redaktionssystems sein.

Basisfunktionalitäten: Editieren, Löschen und Ergänzen von Lexikoneinträgen

Weitere Funktionalitäten: Erstellung von Terminologielisten, Glossaren, Import und Export von externen Wörterbüchern, Zusammenführen von Wörterbüchern (Mergen), Konsistenzprüfungen, automatisches Nachschlagen im Lexikon für Wörter, die im Text markiert sind.

vernetzt mit:

- XML-Editor (für die Erfassung des Texts)
- Wörterbuch-Linkeditor (wird eingebettet)
- Recherche