
TETRIS Workshop

„Sprachtechnologie im Einsatz“

**Terminologie — Workflow — Sprachkontrolle
in der Technischen Dokumentation**

Saarbrücken, 28. 02. 2002

Handout zu den Arbeitsgruppen

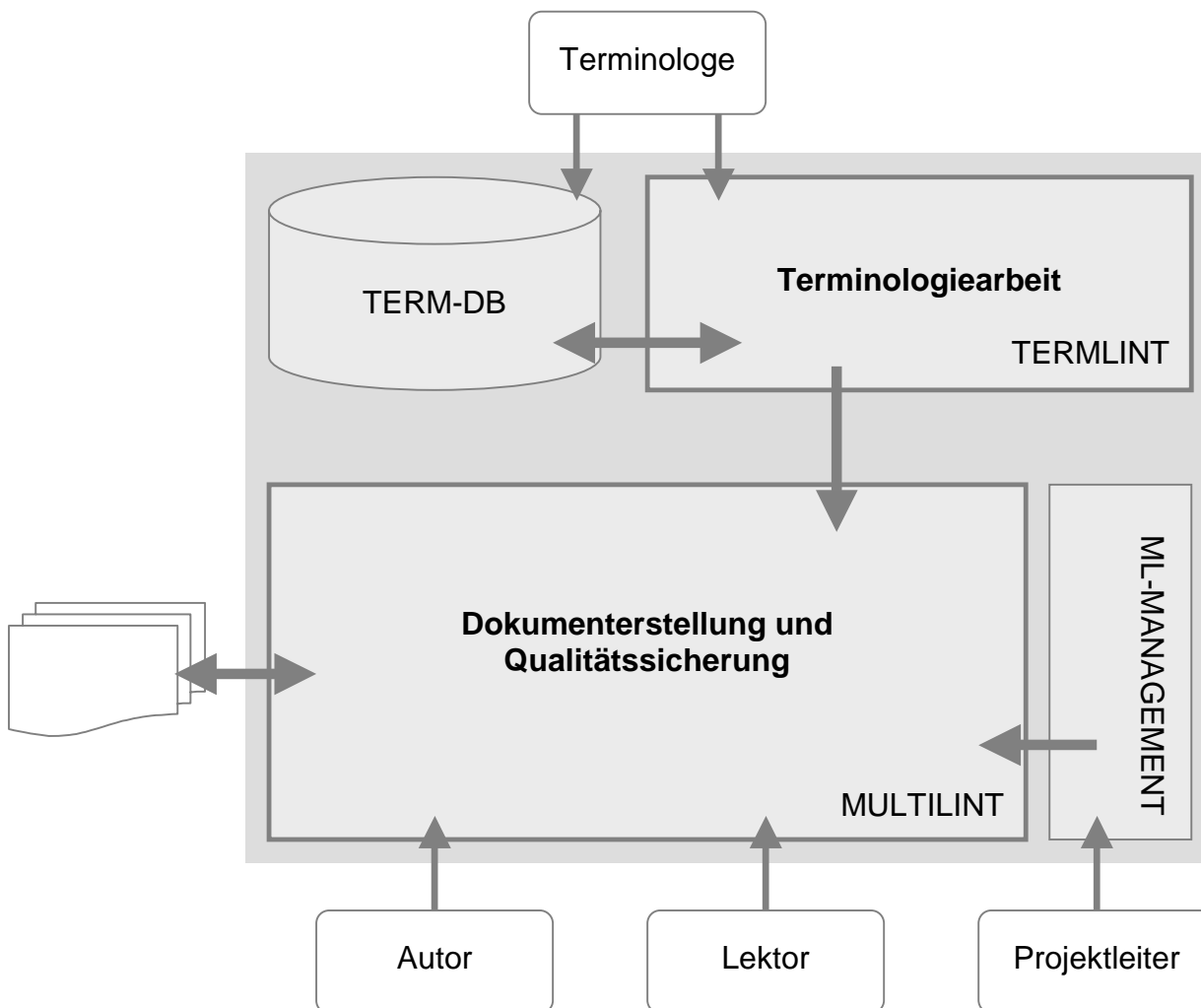
Terminologiarbeit und Texterstellung/Lektorat

TETRIS Sprachtechnologie

Dem Bereich der technischen Dokumentation kommt durch die zunehmende Produktvielfalt und die kürzer werdenden Produktionszyklen eine immer größere Bedeutung im Wirtschaftskreislauf zu. Deshalb sind Systeme gefordert, die den gesamten Ablauf der Dokumenterstellung intelligent unterstützen.

Innerhalb dieses Prozesses der Dokumenterstellung hat die ausgangssprachliche Textqualität eine herausragende Rolle: Je besser, je verständlicher, je eindeutiger, je höher die Qualität eines ausgangssprachlichen Textes ist, umso effizienter lässt er sich in allen folgenden Schritten weiterverarbeiten. Dies führt zu geringerem Redigier- und Postitionsaufwand, kürzeren Erstellungszyklen bei Updates, besseren Retrievalergebnissen, geringerem Übersetzungsaufwand, etc. All diese Teilbereiche der technischen Dokumentation können erheblich optimiert werden, indem linguistische Werkzeuge zum Einsatz kommen, wie sie in TETRIS entwickelt wurden.

Das folgende Diagramm schematisiert die in den Arbeitsgruppen vorgestellten Kernkomponenten der TETRIS-Sprachtechnologie TERMLINT und MULTILINT in ihrer Einbettung in den Dokumentationsprozess.



Terminologiearbeit mit TERMLINT und TEDDY

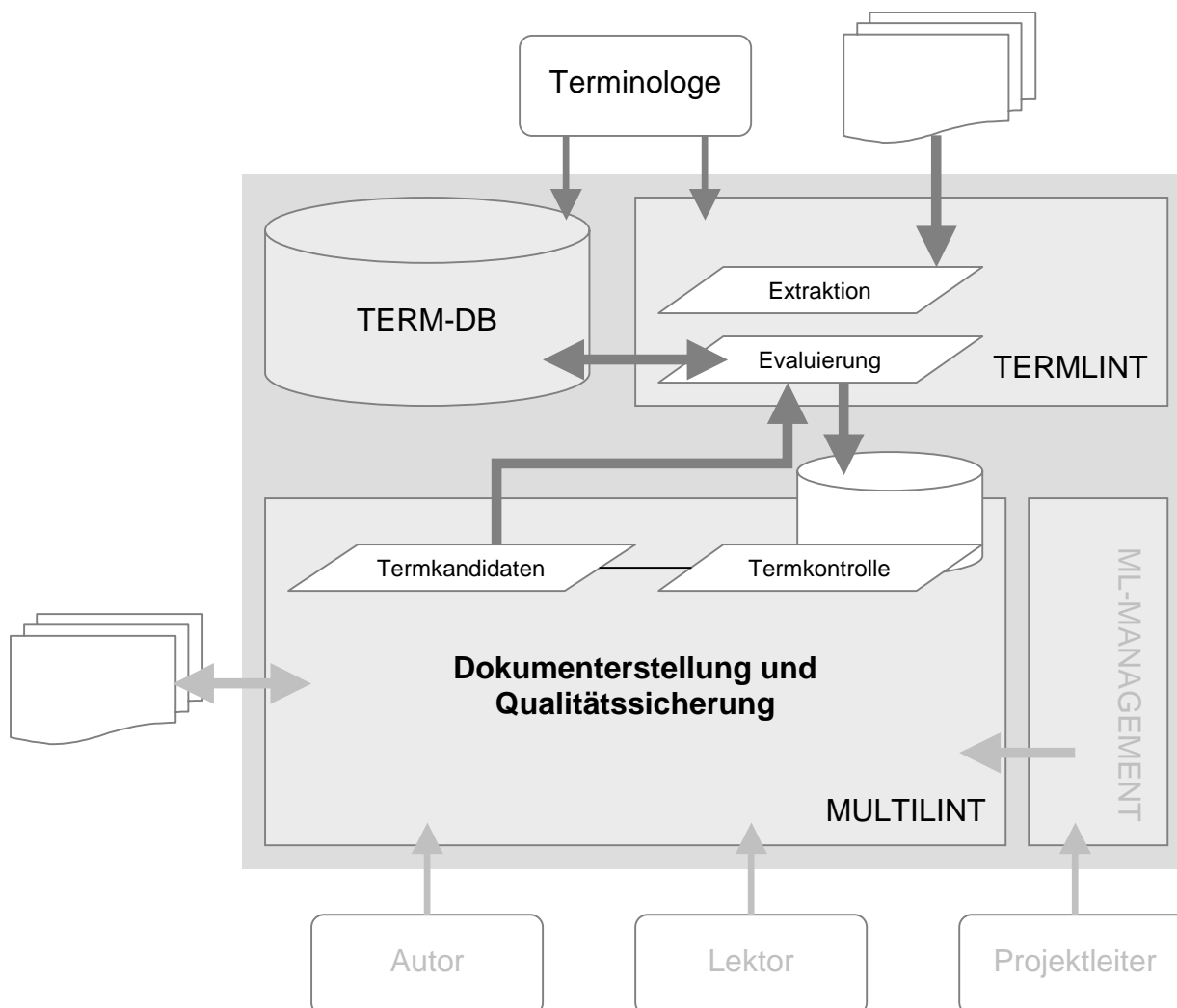
TERMLINT ist ein Werkzeug, mit dem ein Terminologe eine bestehende Terminologie evaluieren kann, wie sie etwa in einer externen Term-DB erfasst ist. Die verfügbaren Evaluierungsfunktionen decken folgende sprachliche Bereiche ab:

- Rechtschreibung
- Grammatik
- Termbildung
- Konsistenz

Eine auf dieser Grundlage bereinigte Terminologie kann in der Folge für die MULTILINT-Termkontrolle hinterlegt werden (siehe S. 9).

Weiterhin stellt TERMLINT eine Funktionalität zur Extraktion von Terminologie aus Dokumentbeständen zur Verfügung, sei es, um eine Terminologie erstmalig zu erstellen, sei es, um eine bestehende Terminologie zu erweitern.

MULTILINT verfügt seinerseits über einen Mechanismus zur Termkandidatenerkennung, mit dem eventuelle Lücken in der bestehenden Terminologie erkannt und im Hintergrund registriert werden (siehe S. 10). Diese Termkandidaten können dann in den Evaluierungszyklus eingehen.



Die TERMLINT-Funktionen im Einzelnen

Terminologieevaluierung:

Rechtschreibkontrolle

Zeigt Verstöße gegen allgemeingültige Rechtschreibregeln an. Drei Rechtschreibeinstellungen werden unterstützt: *Neue Rechtschreibung allgemein*, *progressive Schreibung*, *konservative Schreibung*.

Fehlertyp	Beispiele
Falsche Schreibung	Grundmenue, Innenraumstandart
Unbekanntes Wort	Leuchdiode
Falsche Großschreibung	Dauerhaft
Falsche Kleinschreibung	maximalfrequenz
Alte Rechtschreibung	aneinanderliegen
Progressive Schreibung	Potenzial
Konservative Schreibung	Potential
Falsche Fuge	Kolbenrückzugfeder

Grammatikkontrolle

Zeigt Verstöße gegen allgemeingültige Grammatikregeln sowie flektierte Terme an.

Fehlertyp	Beispiele
Falsche Beugung	automatisches Bestückung
Falscher Kasus der Präpositionsergänzung	gegen dem Motor
Plural-markierte Form	Montagearbeiten
Kasus-markierte Form	Motors
Nicht-infinitivische Verbform	aktiviert
Fehlender Ergänzungsbindestrich	Ein und Ausgabe

Termbildungskontrolle

Zeigt Verstöße gegen Termbildungsrichtlinien an.

Fehlertyp	Beispiele
5- oder mehrteiliges Kompositum ohne Bindestrich	Heißfilmluftmassensensor
2- oder 3-teiliges Kompositum mit Bindestrich	Gravur-Nummer
Infinitivnominalisierung	Aktivieren
Koordinierte Wortgruppe	Anfahr- und Bremsnickausgleich
Kompositumsfähige Wortgruppe	Abschalten des Geräts
Nicht-syntaktische Wortgruppe	Abdeckung Windlauf
	Baureihe klein
Terminologisch fragliche Wortgruppe	Gas geben
	bei Ausfederung
	Kraftstoffanzeige mit Leuchte im Reservebereich
	Momenteinleitung mittels

Konsistenzkontrolle

Zeigt mögliche Inkonsistenzen innerhalb einer Terminologie an. Dabei kann der Benutzer auswählen, ob er einen Termbestand auf interne Konsistenz oder auf Konsistenz mit einer bereits bestehenden MULTILINT-Terminologie prüfen will.

Fehlertyp	Beispiele
Schreibvarianten	Rückführungsventil ↔ Rückführungs-Ventil
Ableitungsvarianten	Rückführventil ↔ Rückführungsventil
Synonymvarianten	Schutzring ↔ Abdeckscheibe ↔ Abdeckring
Bildungsvarianten (Wortgruppe → Kompositum)	Rohr für Wasserablauf ↔ Wasserablaufrohr
Bildungsvarianten (Kompositum → Wortgruppe)	Türabdichtung ↔ Abdichtung für Tür
Reduktionsvarianten	Abblendlichtrelais ↔ Abblend-Relais

Eine Variantenerkennung, die Synonymbeziehungen zwischen Termbestandteilen berücksichtigt, wird durch die Definition von Synonymklassen erreicht, die jedoch hochgradig benutzer- und fachgebietsspezifisch sind (z. B. *schützen* - *abdecken*). Verfahren zur Gewinnung und Pflege von Synonymklassen werden in TETRIS erarbeitet.

Terminologieextraktion

Extrahiert mögliche terminologische Einträge aus Dokumentbeständen. Die Ergebnisse der Termextraktion können ihrerseits mit TERMLINT evaluiert werden.

Für die Auswahl messstoffberührender Materialien hinsichtlich ihrer Korrosionsbeständigkeit im Prozess ist der Anwender verantwortlich. Der Hersteller übernimmt keine Haftung! Bei Schweißarbeiten an der Rohrleitung darf die Erdung des Schweißgerätes nicht über das Meßgerät erfolgen. Der Installateur hat dafür Sorge zu tragen, dass das Messsystem gemäß den elektrischen Anschlussplänen korrekt angeschlossen ist. Der Messumformer ist zu erden, außer bei galvanisch getrennter Hilfsenergie! Beachten Sie grundsätzlich die in Ihrem Land geltenden Vorschriften bezüglich Öffnen und Reparieren von elektrischen Geräten. Die Geräte sind nach dem Stand der Technik betriebssicher gebaut und geprüft und haben das Werk in sicherheitstechnisch einwandfreiem Zustand verlassen. Die Geräte berücksichtigen die einschlägigen Normen und Vorschriften nach EN 61010 „Sicherheitsbestimmungen für elektrische Mess-, Steuer-, Regel- und Laborgeräte“. Wenn sie unsachgemäß oder nicht bestimmungsgemäß eingesetzt werden, können jedoch Gefahren von ihnen ausgehen. Achten Sie deshalb in dieser Betriebsanleitung konsequent auf Sicherheitshinweise, die mit den folgenden Symbolen gekennzeichnet sind: „Warnung“ deutet auf Aktivitäten oder Vorgänge, die - wenn sie nicht ordnungsgemäß durchgeführt werden - zu Verletzungen von Personen oder zu einem Sicherheitsrisiko führen können. Beachten Sie die Arbeitsanweisungen genau und gehen Sie mit Sorgfalt vor.

Die Konfiguration des Stromausganges als „aktiv“ oder „passiv“ erfolgt über verschiedene Steckbrücken auf der I/O-Platine.

Stromschlaggefahr! Offenliegende Bauteile mit berührungsgefährlicher Spannung. Vergewissern Sie sich, dass die Hilfsenergie ausgeschaltet ist, bevor Sie die Elektronikraumabdeckung entfernen. Hilfsenergie ausschalten.

Terminologiebereinigung mit TEDDY

TEDDY, der "kleine Terminologie-Editor", wird bei BMW im Bereich Service zur Bereinigung von ca. 32.000 Termen aus Altbeständen eingesetzt. Ergänzt um aktuelle Terme und Konzepte bildet diese Terminologie die Basis für Übersetzung, Online-Abfragen und MULTILINT.

Teddy unterstützt die Arbeit der Terminologen insbesondere durch die tolerante Suche, das Aufzeigen verwandter Terme sowie durch Prüfungen, die die Einhaltung festgelegter Regeln zur Attribuierung auf Term- und Konzeptebene überwachen.

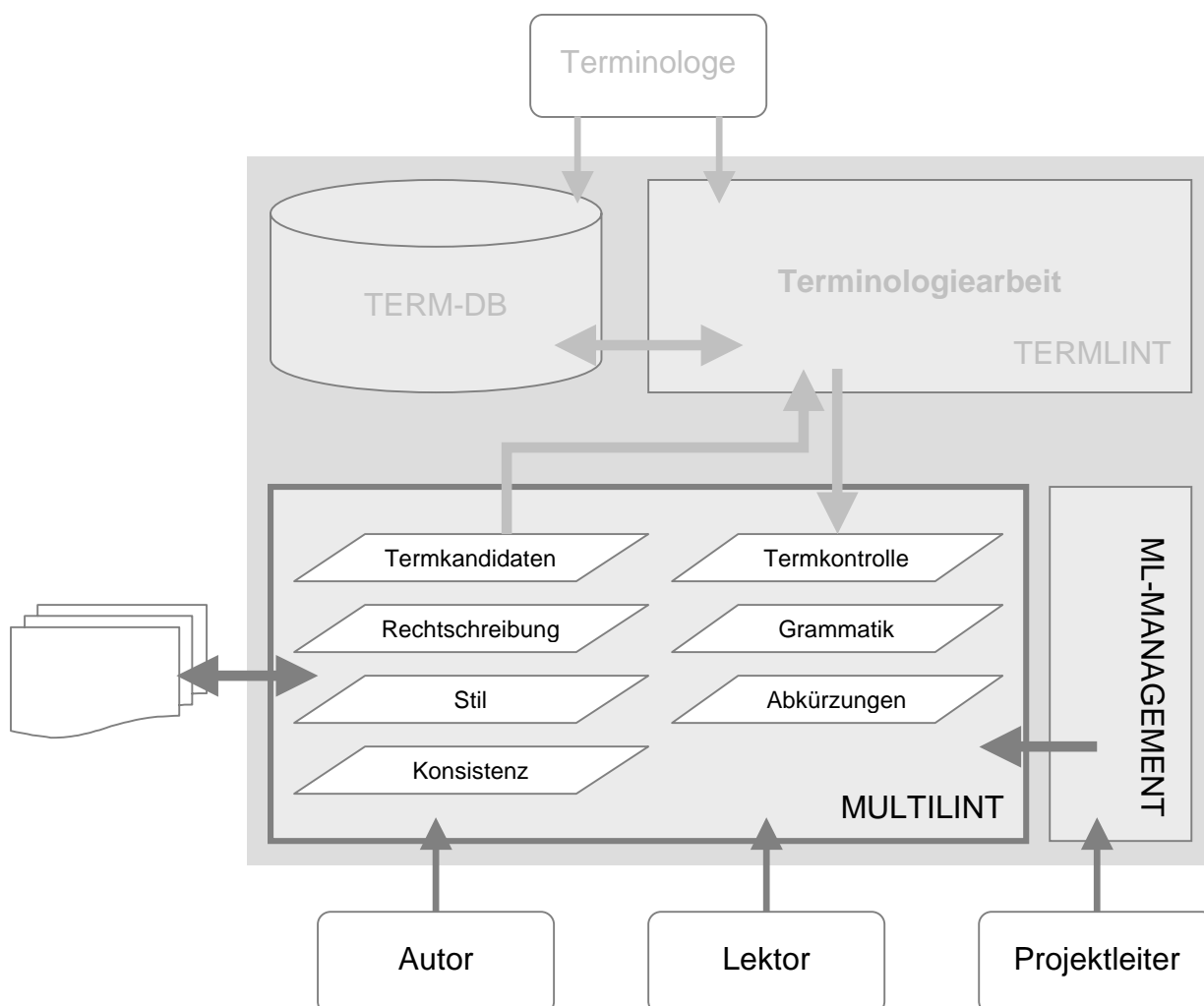
Texterstellung mit MULTILINT

MULTILINT ist ein Werkzeug für sprachliche Textkontrolle, das den Autor bei der Texterstellung sowie den Lektor bei der Qualitätssicherung unterstützt. Die in MULTILINT verfügbaren Kontrollfunktionen decken verschiedene sprachliche Bereiche ab. Folgende Kontrollfunktionen können mit MULTILINT ausgeführt werden:

- Rechtschreibkontrolle
- Grammatikkontrolle
- Stilkontrolle
- Termkontrolle
- Abkürzungskontrolle
- Konsistenzkontrolle

Weiterhin verfügt MULTILINT über einen Mechanismus zur Erkennung von Termkandidaten, der im aktuellen Dokument mögliche Benennungen für eine Erweiterung der bestehenden hinterlegten Terminologieliste identifiziert.

Projektleiter können die hinterlegten linguistischen Wissensquellen mit der ML-MANAGEMENT-Funktionalität projektspezifisch konfigurieren und parametrisieren.

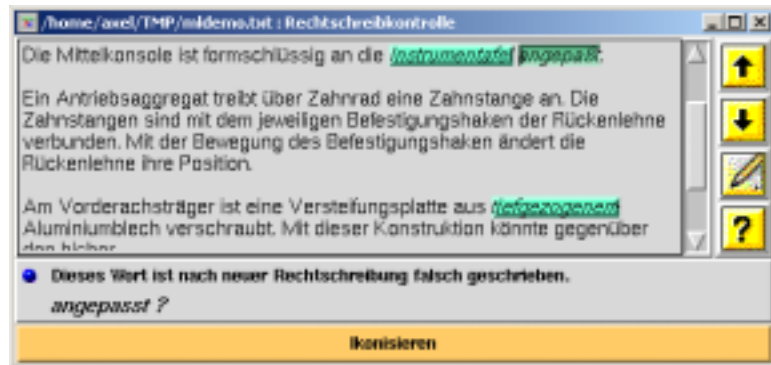


Die MULTILINT-Funktionalitäten im Einzelnen:

Rechtschreibkontrolle

Überprüft Dokumente auf orthografische Korrektheit:

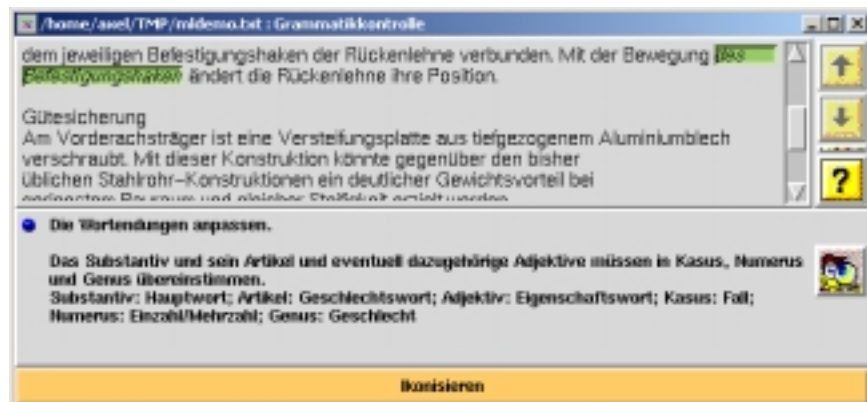
MULTILINT- Rechtschreibkontrolle



Grammatikkontrolle

Überprüft Dokumente auf grammatische Korrektheit.

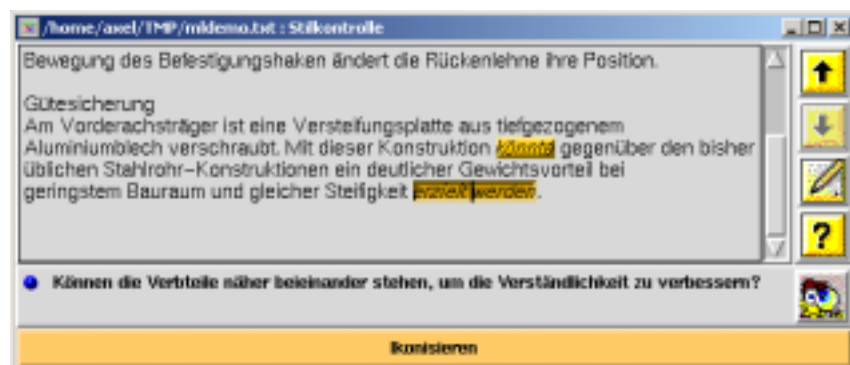
MULTILINT- Grammatikkontrolle



Stilkontrolle

Überprüft die stilistische Korrektheit in Bezug auf Eindeutigkeit und Verständlichkeit.

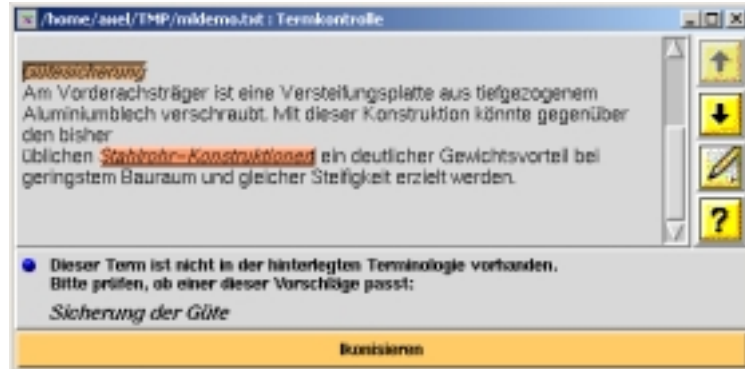
MULTILINT- Stilkontrolle



Termkontrolle

Identifiziert nicht autorisierte Varianten von autorisierten Termen.

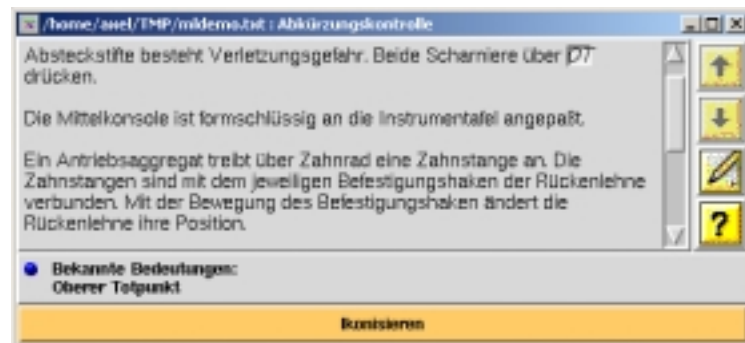
MULTILINT-
Termkontrolle



Abkürzungskontrolle

Überprüft die korrekte Verwendung von Akronymen und Abkürzungen.

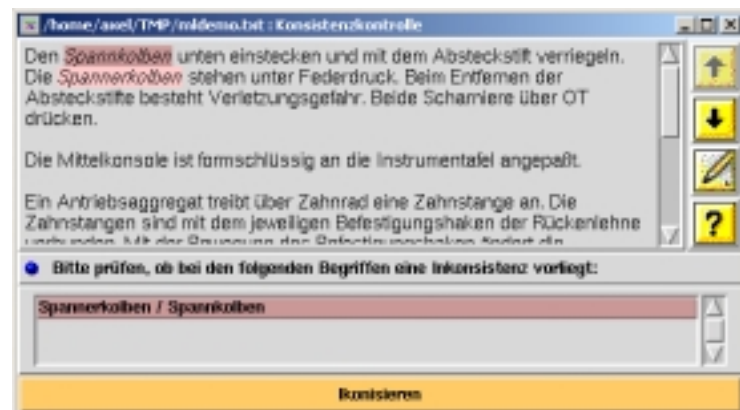
MULTILINT-
Abkürzungskontrolle



Konsistenzkontrolle

Überprüft die konsistente Verwendung von Benennungen innerhalb eines Dokuments, und zwar **unabhängig von der hinterlegten Terminologie**.

MULTILINT-
Konsistenzkontrolle



Termkandidatenerkennung

Beim Prüfen von Dokumenten werden eventuelle Lücken in der bestehenden Terminologie erkannt und im Hintergrund registriert.

MULTILINT-
Termkandidaten-
erkennung

Termkandidat:
maschinenausführung

Kontext:

Rechnerkarte Steckplatz REK2 (Hardware abhängig von **Maschinenausführung**)

Dateiname:
/home/chris/scripts/DE_UTKL.S0170006002000.gli

Autor:
chris

Datum:
3.12.2001 13:45:56

Termkandidaten:

maschinenausführung**** verschickt am 3.12.2001 13:47:40

MULTILINT-Management

Mithilfe der ML-MANAGEMENT-Funktionalität können die in MULTILINT hinterlegten linguistischen Wissensquellen projektspezifisch konfiguriert und parametrisiert werden.

Neben der Hinterlegung projektspezifischer Terminologie können Module zur Prüfung einzelner Phänomene im Bereich der Stil- und Grammatikkontrolle gemäß den Anforderungen bestimmter Projekte aktiviert bzw. deaktiviert werden.

Darüber hinaus können für Module, die z. B. Informationsdichte oder Satzlänge betreffen, die entsprechenden Schwellenwerte benutzerdefiniert eingestellt werden.