

# Informationsaufbereitung und -erschließung im Rahmen des Learning Content Management

## Beitrag von sprach- und texttechnologischen Verfahren zur Bestandssicherung der erstellten Lehrmaterialien

**Abstract:** Im Rahmen des Förderprogramms *Neue Medien in der Bildung* haben zahlreiche akademische Institutionen in knapp 600 Teilprojekten hochwertige Lehrmaterialien geschaffen, digitalisiert und multimedial aufbereitet. Sie stellen diese zur Begleitung der Präsenzlehre oder für die virtuelle Lehre in einer Vielzahl von Fachbereichen zur Verfügung. Für das Management dieser Lehrmaterialien, auch *Content* genannt, sowie zur Verwaltung der mit diesen Materialien organisierten Lehrveranstaltungen kommen Learning Management bzw. Learning Content Management Systeme zum Einsatz.

In diesem Beitrag soll aufgezeigt werden, wie der Einsatz von sprach- und texttechnologischen Verfahren das Management von Lehrmaterialien zugleich vereinfachen und effektiver gestalten kann. Die Nutznießer der dargestellten Hilfsmittel sind sowohl AutorInnen der Lehrmaterialien als auch deren Bereitsteller und NutzerInnen.

### 1. Die Ausgangslage und Motivation

Gegenstand der Förderung im Programm "Neue Medien in der Bildung" des BMBF (Laufzeit 2001 bis 2003) war und ist die Einführung multimedialer und virtualisierter Lehr- und Lernformen an den Hochschulen und insbesondere die Einführung solcher Studienelemente in den Normalbetrieb der Hochschule (s. Ausschreibungstext des Förderprogramms; [http://www.gmd.de/PT-NMB/Programm/ Programm.html](http://www.gmd.de/PT-NMB/Programm/Programm.html)). In knapp 600 Teilprojekten wurden hierfür geeignete Lehrmaterialien - "Learning Objects" - erstellt und in Lehrveranstaltungen eingesetzt. Zum Teil konstituieren diese Materialien und Lehrveranstaltungen ganze Studiengänge.

Bei der Verwaltung der Lehrmaterialien und der Lehrveranstaltungen wurden bisher eine ganze Reihe von "Learning Management" und "Learning Content Management" Systemen (zur Unterscheidung s. Hettrich / Koroleva 2003) eingesetzt. Einer kleinen Zahl häufig verwendeter Systeme steht dabei eine größere Zahl von Nischenlösungen gegenüber. Einige Learning (Content) Management Systeme wurden als Open Source entwickelt und werden von der E-Learning Community an den Hochschulen weiterentwickelt (s. Projekt "Medienquadrat", <http://141.54.159.52/m2demo/m2/extern/index.php>; und "Campussource", <http://www.campussource.de/>) Aus dieser Heterogenität der im Hochschulbereich für das E-Learning verwendeten Learning (Content) Management Systeme müssen, will man die erzielten Ergebnisse sichern und nachhaltig verwenden, folgende Schlüsse gezogen werden:

1. Lehrinhalte, im folgenden: Lernobjekte genannt, müssen in eine Form gebracht werden, in der sie unabhängig von konkreten Systemen verwaltet oder zumindest zwischen Systemen ausgetauscht werden können. Ein system- und plattformunabhängiges, möglichst weit standardisiertes Annotationsformat für Lernobjekte ist ein adäquates Mittel, die gebotene Unabhängigkeit und Austauschbarkeit zu gewährleisten.
2. Es müssen Metadatensätze angelegt werden, die die Inhalte dieser Lernobjekte beschreiben und diese so der Suche zugänglich und sie in anderen Kontexten wiederverwendbar machen.
3. Es erhöht den Nutzwert insbesondere von Hypertexten, also Lehrmaterialien, die nicht zwingend linear konsumiert werden, wenn die Textbausteine so miteinander vernetzt werden, dass die LeserInnen an jeder Stelle auf die für das Verständnis dieser Stellen notwendige Vorwissen zurückgreifen können. Dies setzt eine dichte, inhaltlich motivierte Vernetzung zwischen den Bausteinen der Lehrtexte, aber auch zwischen Lehrtexten und externen Ressourcen, wie z.B. Glossaren, voraus.

Sowohl für die Strukturierung von Lernobjekten als auch für deren Beschreibung mit Metadaten gibt es etablierte (Quasi-)Standards. XML (extensible Markup Language, <http://www.w3c.org/TR/2000/REC-xml-20001006>) als abstrakte Syntax im ersten Fall, LOM (Learning Objects Metadata; <http://ltsc.ieee.org>) oder SCORM (Sharable Courseware Object Reference Model; <http://www.rhassociates.com/scorm.htm>) im anderen Fall (vgl. Lobin et al. 2003). Auch die Verwendung von avancierteren Metadatensprachen wie EML (Educational Modelling Language, <http://eml.ou.nl/introduction/explanation.htm>) kommt möglicherweise als Alternative in Frage.

Die Vorteile der Verwendung von Standards auf diesen beiden Ebenen liegen auf der Hand:

1. Mit XML und in Übereinstimmung mit einem offenen Dokumentenmodell strukturierte Lernobjekte sind nicht an eine proprietäre Plattform gebunden, sondern zwischen den Plattformen austauschbar. Dies gilt auch für Plattformen und Systeme, die in der Zukunft entwickelt werden. Die klare Trennung von System und Daten sichert den Bestand der Daten, die im Rahmen des Förderprogramms als das eigentliche Gut geschaffen wurden.
2. Die Orientierung an einen Metadatenstandard erlaubt die Übernahme von Metadaten-Sätzen in Standard-Kataloge und erhöht somit die Sichtbarkeit und Benutzbarkeit dieser Objekte.

Besonders die Erstellung von Metadaten ist jedoch eine nicht zu unterschätzende Belastung für die Autoren von Lernobjekten. Ähnliches lässt sich für die inhaltsbezogene Vernetzung der Lernobjekte feststellen. Je granularer das System von Lernobjekten und je aussagekräftiger die Metadaten - beides sind erstrebenswerte Zielparame-ter - desto höher ist der Aufwand bei der Vernetzung der Lernobjekte und bei der Erstellung von Metadaten.

Um Möglichkeiten der Unterstützung dieser beiden Strategien einer nachhaltigen Erstellung und Nutzung von Lehrmaterialien soll es in der folgenden Skizze gehen. Wir werden darstellen, welchen Beitrag sprach- und texttechnologische Methoden und Werkzeuge bei der *Unterstützung* dieser Arbeitsgänge, zum Nutzen von Autoren, Distributoren und Nutzern der Lehrmaterialien, leisten können.

Für die Realisierung der in diesem Text skizzierten Softwaremodule soll ILIAS als Referenzplattform dienen. ILIAS scheint uns als Referenzplattform für die Softwareentwicklung besonders geeignet: Erstens gehört dieses LMS zu den weiter verbreiteten Systemen, nicht nur im universitären Bereich. Eine Begleitstudie zum Förderprogramm hat ergeben, dass von 50 Projekten, die den Einsatz einer Lernplattform erwogen, sieben Projekte ILIAS favorisierten (*kevi*h Begleitstudie). Es ist zweitens als "Open Source" für Funktionserweiterungen besonders gut geeignet. Mit einer Schnittstelle für Anwendungsprogramme können zudem zukünftige Anforderungen der Benutzer in der Entwicklung einfach und schnell umgesetzt werden. Drittens steht durch die hohe Nutzerzahl in ILIAS eine große Zahl und Bandbreite von Lernobjekten zur Verfügung. Dies ist eine unabdingbare Voraussetzung für Training, Test und Evaluation der im folgenden beschriebenen texttechnologischen Verfahren. Selbstverständlich wird sich der Anwendungsbereich der Softwaremodule nicht auf ILIAS beschränken. Die Softwarearchitektur muss berücksichtigen, dass die Module auch in andere LMS integriert werden können sollten.

Im folgenden Abschnitt werden wir die Infrastruktur für die geplanten sprach- und texttechnologischen Softwaremodule skizzieren. In Abschnitt drei stellen wir die Funktionalität der Softwaremodule für *AutorInnen* von Lehrmodulen vor und in Abschnitt vier die Module, die eine nachhaltige Nutzung der Lehrmaterialien unterstützen. Abschnitt fünf gibt eine Zusammenfassung und einen Ausblick.

## **2. Die Software-Infrastruktur: Agenten**

Die im Folgenden beschriebenen Softwarewerkzeuge sollen AutorInnen, Distributoren und BenutzerInnen von Lernobjekten unterstützen. Sie sollen diesen Akteuren keinesfalls die Kontrolle über die Arbeitsschritte abnehmen und selbstständig Daten erzeugen. Die Aufgabe eines Software-Agenten ist es, auf Anforderung der Benutzer und auf Grund der existierenden Daten weitere Daten zu erzeugen, die anschließend durch die AnwenderIn zu prüfen und zu ergänzen oder zu

korrigieren sind. Softwareagenten nehmen also den Personen, die sie unterstützen, die Arbeit nicht gänzlich ab, sondern sie erleichtern diese. Die Software-Agenten werden entweder automatisch aktiviert, etwa wenn ein Arbeitsschritt abgeschlossen ist, oder sie werden durch den Benutzer aufgerufen. Sie verrichten ihre Arbeit im Hintergrund und treten für die AnwenderInnen nur durch ihre Ergebnisse in Erscheinung.

### **3. Texttechnologische Werkzeuge für AutorInnen und Bereitsteller von Lernobjekten: Textaufbereitende Agenten**

#### **Vorbemerkung**

Lernobjekte können entweder innerhalb des Learning Content Management Systems erstellt werden oder außerhalb, etwa durch ein Authoring Werkzeug. ILIAS erlaubt beides: Lernobjekte, die in einem der ILIAS Document Type Definition für Lernobjekte entsprechenden Format entwickelt wurden, können in ILIAS importiert und dort mit Attributen für das Layout der einzelnen Textteile versehen werden. Lernobjekte können aber auch mit Hilfe eines integrierten Editors direkt in ILIAS erstellt werden.

Es erscheint daher sinnvoll, die im Folgenden beschriebenen Agenten unabhängig vom Erstellungsprozess eines Lernobjekts, also erst nach dessen Fertigstellung, zu aktivieren. Dies entspricht auch den Prinzipien der modularen Entwicklung dieser Agenten, was wiederum deren Wiederverwendbarkeit in anderen Kontexte erhöht.

Da die Erstellung und die texttechnologische Aufbereitung der Texte zwei getrennte Prozesse sind, kann letztere auch die Bereitsteller von Lernobjekten unterstützen, wenn sie nicht die AutorInnen dieser Lernmodule sind. In diesem Falle ist die hier skizzierte Unterstützung sogar noch wertvoller, da die Kenntnis der Texte beim Bereitsteller weitaus geringer sein dürften als bei den AutorInnen.

#### **Ermittlung von Schlüsselwörtern**

Zentral für das spätere Auffinden eines Lernobjekts sind die ein Lernobjekt inhaltlich beschreibenden Schlüsselwörter in den Metadaten. Ein Softwareagent kann AutorInnen beim Ermitteln von Schlüsselwörtern durch eine quantitative und qualitative Analyse des Lernobjekts unterstützen.

Der einfachste Fall ist, dass Schlüsselwörter bereits im Text markiert sind, etwa zum Zwecke der späteren Hervorhebung. Dieser Fall muss hier nicht mehr diskutiert werden, ist aber bei der Erstellung eines Agenten zu berücksichtigen.

Wenn diese Informationen nicht gegeben sind, muss das Programm auf quantitative und qualitative Charakteristika von typischen Schlüsselwörtern zurückgreifen:

- Schlüsselwörter sollen ein Lernobjekt, im Kontrast zu anderen Lernobjekten, charakterisieren. Ein Schlüsselwort taucht deshalb typischerweise in dem Lernmodul, das es charakterisiert, öfter auf als in anderen Lernmodulen. Dies lässt sich mit Hilfe von Referenz-Lernmodulen, die zum Vergleich herangezogen werden, und durch die Bildung von Schwellenwerten und der Etablierung einer numerischen Ordnung der Kandidatenliste modellieren (vgl. Matsuo / Ishizuka 2003).
- Schlüsselwörter tauchen typischerweise an bestimmten Stellen in einem Text auf: in den Überschriften, in der ersten Passage eines Textes nach einer Überschrift, in der Zusammenfassung. Das Vorkommen der Schlüsselwortkandidaten an diesen Stellen geht in die Gewichtung dieser Kandidaten mit ein.
- Die Ergebnisse der Extraktion verbessern sich deutlich, wenn man die zu bearbeitenden Texte vorher linguistisch annotiert (vgl. Hulth 2003). Hierzu benötigt man die für eine präzise linguistische Annotation notwendigen Werkzeuge und Daten.

Der Softwareagent liefert eine Liste von Schlüsselwortkandidaten für ein Lernobjekt oder für eine Gruppe von Lernobjekten höherer Ordnung, z.B. für einen ganzen Kurs. Die Schlüsselwörter

werden dabei auf ihre Grundform reduziert, wofür es verlässliche sprachtechnologische Verfahren gibt. Die ausgewählten Schlüsselwörter können in die Metadaten übernommen werden.

### **Ermittlung von Kandidaten für ein Glossar**

Glossare sind wichtige Sekundärindizes zu Texten. Sie sind quasi textspezifische kleine Lexika, die die definitorische Entschlüsselung von zentralen Termen des Textes, auf den sie sich beziehen, leisten. Grundlage eines Glossars sind die Definitionen oder Bedeutungsparaphrasen, die in den Texten selber gegeben werden. Da nicht davon ausgegangen werden kann, dass komplexe, als Hypertext organisierte Lernobjekte ganz und in einer festen Reihenfolge rezipiert werden, empfiehlt es sich, für jeden ins Glossar aufgenommenen Term das erste Vorkommen in jeder Texteinheit einen Link auf diesen Glossareintrag zu erstellen.

Für das Erkennen von definitorischen Kontexten gibt es mittlerweile zuverlässige regelbasierte Verfahren. Diese arbeiten auf der Grundlage der Erkennung typischer sprachlicher Muster für Definitionen (vgl. Beisswenger / Lenz / Storrer 2002; Klavans / Muresan 2001).

### **Erstellung von Linkangeboten**

Ein komplexes Lernobjekt oder gar ein Kurs lebt von den Bezügen, die zwischen seinen Teilen hergestellt werden und die als alternative Rezeptionspfade aufgefasst und modelliert werden können. Auf der Ebene digitaler Hypertexte können diese Bezüge durch Hyperlinks hergestellt werden. Auch dieser Prozess kann durch einen Softwareagenten unterstützt werden, der auf Grund einer normalisierten und strukturierten lexikalischen Repräsentation des Textes Quell- und Zielstellen für verschiedene Typen von Hyperlinks ermittelt (vgl. Pustejovsky et al. 1997, eine Anwendung wird in Beisswenger / Lenz / Storrer 2002 beschrieben).

### **Extraktion von Schlüsselsätzen für eine Textzusammenfassung**

Eine weiteres wichtiges Mittel für die Erschließung von Lernmodulen ist das Abstract, welches ein Surrogat des Textes darstellt. Es dient u.a. dazu, die Relevanz eines Lernmoduls für den eigenen Lernzweck oder den Lernzweck der Zielgruppe abzuschätzen. Die Techniken des automatischen Zusammenfassens sind allerdings nicht dazu geeignet, ein solches Surrogat automatisch zu erstellen. Dafür ist die Zahl der Parameter, die das Verhältnis von Surrogat und Volltext bestimmen, zu groß (für einen Überblick siehe Mani 2001).

Ein Verfahren für die Bildung eines Textsurrogats, welches in Reichweite der Texttechnologie auf dem heutigen Stand liegt, ist die Extraktion der relevantesten Sätze des Volltextes. Diese werden ohne weitere Bearbeitung den AnwenderInnen vorgelegt. Diese können nach Maßgabe ihrer Zielgruppe, der verfolgten Zwecke etc. hieraus eine angemessene Zusammenfassung erstellen.

### **Zusammenfassung**

Ziel der *textaufbereitenden Agenten* ist es, AutorInnen und Bereitstellern bei der Erstellung von Metadaten und sekundären Daten (Schlüsselwörter, Glossar, Links, Abstract) zu unterstützen. Die beschriebenen Verfahren und daraus resultierenden Agenten werden auf Lernobjekte angewendet, deren Erstellung abgeschlossen ist. Sie können daher auch auf Lernobjekte angewendet werden, die extern mit einem Autorentool erstellt wurden. Sie können außerdem mehrfach auf ein Lernobjekt angewendet werden, etwa wenn dieses grundlegend überarbeitet wurde.

## **4. Texttechnologische Anwendungen für die Benutzer eines Learning Content Management Systems: Texterschließende Agenten**

Im Folgenden betrachten wir die Interessen von zwei Gruppen von Benutzern:

- DozentInnen, die für eine bestimmte Zielgruppe ein Lehrangebot erstellen wollen und dafür auf vorhandene Lernobjekte bzw. Kurse zurückgreifen wollen.

- Lernende, die ausgehend von einem konkreten Lernbedürfnis Lernobjekte suchen, die dieses Lernbedürfnis durch Selbststudium oder Unterricht befriedigen helfen können.

Der wichtigste Schlüssel für den Zugriff auf Lernobjekte sind die Metadaten. Den Suchenden steht darüber hinaus in der Regel auch der Volltext zur Verfügung, der mit Hilfe einer Suchmaschine erschlossen werden kann. Mittlerweile dürfte aber jeder schon mal die Erfahrung gemacht haben, dass die stichwortbezogene Suche über eine Suchmaschine durch die geringe Präzision der oft großen Trefferzahlen unbefriedigend ist. Vor allem die Relevanzbewertung der angezeigten Dokumente lässt oft zu wünschen übrig (niedrige Präzision). Zudem kann man nicht sicher sein, dass für die Anfrage wichtige Dokumente übersehen wurden (niedriger Recall). Eine Möglichkeit, die Ergebnisse der Suche zu verbessern, ist es, mit Hilfe einer semantisch organisierten lexikalischen Ressource den Suchraum zu erweitern bzw. einzuengen:

1. Durch die Hinzunahme von Synonymen und Oberbegriffen kann der Suchraum erweitert werden;
2. Durch den Ausschluss von Kohyponymen kann der Suchraum verkleinert werden.

Für die lexikalische Erweiterung der Suchanfrage ist ein digitales Wortnetz des Deutschen notwendig. Dieses liegt mit GermaNet (<http://www.sfs.uni-tuebingen.de/lsd>) für den zentralen Teil der deutschen Allgemeinsprache und mit WordNet (<http://www.cogsci.princeton.edu/~wn/>) für den zentralen Teil der englischen Allgemeinsprache vor. Strategien für die Erweiterung eines Thesaurus aus fachsprachlichen Texten werden in Hahn et al. (1995) dargestellt.

Auch hier gilt wieder, dass der Software-Agent den AnwenderInnen Kandidaten für Suchterme vorlegt, die diese bestätigen, ergänzen oder verwerfen können.

## **Benutzergerechte Auswahl von Lernobjekten**

Ein erst jüngst entwickeltes Verfahren im Information Retrieval die "Latente Semantische Analyse" (vgl. Lanadauer / Foltz / Laham 1998). Die für die Dokumentenaufbereitung im Information Retrieval üblichen Dokument-Term-Matrizen werden durch Singulärwertzerlegung auf wenige, aber aussagekräftige Zellen reduziert. Durch diese Reduktion entstehen einerseits starke Verbindungen zwischen Clustern von Termen, welche auf eine große semantische Nähe hindeuten, und andererseits starke Ähnlichkeitsbeziehungen zwischen den analysierten Dokumenten. Letztere lassen sich für die benutzergerechte Auswahl von Lernobjekten nutzen. Die potenziellen BenutzerInnen eines Kurses schreiben vor Beginn dieses Kurses einen Essay. Dieser wird als Referenztext mit allen Lernobjekten verglichen und ein jeweils paarweiser Ähnlichkeitswert ermittelt. Wird ein Schwellenwert überschritten, dann kann davon ausgegangen werden, dass die Inhalte des Lernobjekts bereits hinreichend gut beherrscht werden (vgl. hierzu Rehder et al. 1998). Diese Lerneinheit kann übersprungen bzw. als Ergänzung angeboten werden. Ähnliche Möglichkeiten ergeben sich bei der Lernerfolgskontrolle am Ende eines Seminars oder Seminarabschnitts.

## **5. Zusammenfassung**

Im Kontext eines eines Learning Content Management Systems, aber unabhängig von einem bestimmten System, können auf der Basis bewährter sprach- und texttechnologischer Verfahren Software-Agenten entwickelt werden. Diese Software-Agenten unterstützen Autoren bei der Beschreibung und Klassifizierung ihrer Lernobjekte, Distributoren bei der nachträglichen Aufbereitung dieser Lernobjekte, DozentInnen bei der zielgruppengerechten Auswahl und LernerInnen bei der Wahl ihres individuellen Lernpfades für den selbstgesteuerten Wissenserwerb. Der Einsatz sprach- und texttechnologischer Werkzeuge macht textuelle Lernobjekte transparenter, leichter austauschbar und auffindbar und verbessert die Möglichkeiten zu deren Rekombinierung in neuen Kontexten. Dies ist ein wichtiger Schritt zur Bestandssicherung der im Rahmen des Förderprogramms entstandenen Projekten und für deren Verbreitung im Alltagsbetrieb der Hochschulen.

## 6. Literatur

- Beißwenger, Michael / Eva Anna Lenz / Angelika Storrer (2002): Generierung von Linkangeboten zur Rekonstruktion terminologiebedingter Wissensvoraussetzungen. In: Stephan Busemann (Hrsg.): KONVENS 2002. 6. Konferenz zur Verarbeitung natürlicher Sprache. Proceedings, Saarbrücken, 30.09.-02.10.2002. Saarbrücken 2002 (DFKI Document D-02-01), 187-191.
- Hahn, Udo / Manfred Klenner / Klemens Schnattinger (1995): Learning from Texts: A Terminological Metareasoning Perspective. LIF Report 1/95. URL: <http://www.coling.uni-freiburg.de/publications/papers/ijcai95.ps.gz>
- Hettrich, Alexander / Natascha Koroleva (2003): Marktstudie Learning Management Systeme (LMS) und Learning Content Management Systeme. Fokus deutscher Markt. Schriften des Fraunhofer Instituts für Arbeitswirtschaft und Organisation, Stuttgart.
- Hulth, Anette (2003): Improved Automatic Keyword Extraction Given More Linguistic Knowledge. Proc. EMNLP 2003. URL: <http://acl.ldc.upenn.edu/acl2003/emnlp/pdfs/Hulth.pdf>
- Jennings, N.R. (2002): On agent-based software engineering. In: Artificial Intelligence 117 (2002) 2, S. 277-296
- Klavans, Judith L. / Smaranda Muresan (2000): DEFINDER: Rule-Based Methods for the Extraction of Medical Terminology and their Associated Definitions fom On-line Text. URL: <http://www.amia.org/pubs/symposia/D200895.PDF>
- Klavans, Judith L. / Smaranda Muresan (2001): Evaluation of the DEFINDER System for Fully Automatic Glossary Construction. Proceedings of the AMIA Symposium 2001. Washington DC, USA. URL: <http://persival.cs.columbia.edu/papers/klavans-amia01-cr.pdf> (DEFINDER)
- Landauer, Thomas K. / Foltz, Peter W. / Darrell Laham (1998): An Introduction to Latent Semantic Analysis. In: Discourse Processes 25 (1998), S. 259 - 284. URL: <http://lsa.Colorado.EDU/papers/dp1.LSAintro.pdf>
- Lobin, Henning / Maik Stührenberg / Gerog Rehm (2003): eLearning und offene Standards. Zum Einstatz XML-strukturierter Lernobjekte. In: Sprache und Datenverarbeitung 2003, Heft 1.
- Mani, Inderjeet (2001): Automatic Summarization. Amsterdam 2001
- Matsuo, Yutaka / Mitsuru Ishizuka (2003): Keyword Extraction from a Single Document using Word Co-occurrence Statistical Informatation. URL: [www.miv.t.u-tokyo.ac.jp/papers/matsuoFLAIRS03.pdf](http://www.miv.t.u-tokyo.ac.jp/papers/matsuoFLAIRS03.pdf)
- Pustejovsky / Boguraev / Verhagen / Buitelaar / Johnston (1997): Semantic Indexing and Typed Hyperlinking. In: Proc AAAI '97. URL: <http://www.cs.brandeis.edu/~llc/publications/aaai97.ps>
- Rehder, B. / M.E. Schreiner / B.W. Wolfe / D. Laham / T.K. Landauer / W. Kintsch (1998): Using Latent Semantic Analysis to assess knowledge. In: Discourse Processes 25 (1998) 337-354. URL: <http://lsa.Colorado.EDU/papers/dp3.rehder.ps>

## Kontakt

Dr. Lothar Lemnitzer  
Project MiLCA (FKZ 01NM167)  
Seminar für Sprachwissenschaft, Abt. Computerlinguistik  
Wilhelmstr. 19  
72 074 Tübingen  
E-Mail: [lothar@sfs.uni-tuebingen.de](mailto:lothar@sfs.uni-tuebingen.de)