

Der Einsatz der automatischen Indexierungssoftware AUTINDEX im Zentrum für Psychologische Information und Dokumentation (ZPID)

Abstract

Implementierung, Funktionsweise und erste Ergebnisse der Evaluation des automatischen Indexierungssystems AUTINDEX im Rahmen der Dokumentation psychologischer Literatur und Medien in der Datenbank PSYINDEX werden beschrieben. Auf der Grundlage eines mittels einer etwa 23.500 Begriffe umfassenden Indikatorenliste erweiterten Thesaurus generiert AUTINDEX aus Abstracts, Titeln und Autorenschlagworten der Dokumente Deskriptorvorschläge zur Unterstützung der inhaltlichen Erschließung durch manuelle Indexierer. Der Aufbau der Indikatorenliste sowie die technische und methodische Integration von AUTINDEX in den Dokumentationsablauf werden ausführlich dargestellt. Im Anschluss wird über eine erste Überprüfung von AUTINDEX berichtet. Insgesamt 63 Dokumente aus PSYINDEX, deren manuelle Indexierung auf der Grundlage des aktuellen APA-Thesaurus 2006 ohne Unterstützung durch die automatische Indexierung erfolgt war, wurden nachträglich mit AUTINDEX indiziert. Ein Häufigkeitsvergleich zeigt, dass AUTINDEX im Schnitt acht bis neun Deskriptorvorschläge pro Dokument generierte, während in der manuellen Indexierung im Schnitt sechs bis sieben Deskriptoren pro Dokument vergeben wurden. Bezogen auf die Anzahl der von AUTINDEX generierten Deskriptorvorschläge ergab sich eine Übereinstimmung von 35,36 Prozent, das heißt im Durchschnitt waren etwa drei der insgesamt acht bis neun Vorschläge von AUTINDEX in der manuellen Indexierung vorhanden. Bezogen auf die Anzahl der von den manuellen Indexierern vergebenen Deskriptoren bedeutet das, dass 46,7 Prozent dieser Deskriptoren auch von AUTINDEX generiert wurden. Da die Übereinstimmung zwischen AUTINDEX und der manuellen Indexierung maßgeblich von der Gesamtzahl der vorgeschlagenen bzw. vergebenen Deskriptoren abhängt, wurden in Anlehnung an R. S. Hooper und L. Rolling zwei Koeffizienten der Indexierungsübereinstimmung berechnet, in denen die Gesamtzahl der Deskriptorvorschläge bzw. Deskriptorvergaben berücksichtigt werden. Dabei ergaben sich Konsistenzen von 25,22 Prozent nach Hooper und 40,24 Prozent nach Rolling. Zur Einordnung dieser Werte werden zwei Vergleichsmaßstäbe angeführt: (1) Bei einer früheren Überprüfung der Konsistenz zwischen manueller Indexierung und einer AUTINDEX-Version ohne Thesauruserweiterung durch Indikatoren wurden Werte von 14,33 Prozent nach Hooper und 25,06 Prozent nach Rolling ermittelt. (2) In einer von J. Beling auf der Basis von 37 PSYINDEX-Dubletten durchgeführten Bestimmung der Konsistenz zwischen zwei manuellen Indexierern wurden Werte von 42,58 Prozent nach Hooper und 56,99 Prozent nach Rolling festgestellt. Abschließend wird die Aussagekraft der durchgeführten Untersuchung diskutiert, und Möglichkeiten der Optimierung von AUTINDEX werden benannt.

Überblick

Das vom "Institut der Gesellschaft zur Förderung der Angewandten Informationsforschung e.V." (IAI) entwickelte Softwarepaket AUTINDEX (**A**utomatic **I**ndexing) zur automatischen Extraktion von Schlagworten aus deutschen und englischen Texten wird im ZPID zur Unterstützung des Indexierungsprozesses eingesetzt: AUTINDEX generiert aus den Titeln, Zusammenfassungen und Autorenschlagworten eines Dokumentes Deskriptorvorschläge, die dem Humanindexierer zur Auswahl angezeigt und von ihm - sofern er sie als zum Dokumentinhalt passend bewertet - übernommen werden.

AUTINDEX beruht auf einer komplexen aber robusten natürlichsprachigen Analyse, dessen Kernstück eine morpholinguistische Textanalyse darstellt. Dazu gehören die Identifikation von Wortformen, eine morphologische Analyse, und eine flache syntaktische Analyse. Das Ergebnis ist eine disambiguierte syntaktische Repräsentation, die unter anderem auch mehrwortige Einheiten mit ihren syntaktischen Varianten darstellt. Die Berechnung von Deskriptoren beruht auf einer statistischen Funktion, die im Gegensatz zu anderen Verfahren nicht allein auf der Frequenz von Wörtern, sondern auf der Frequenz von semantischen Klassen beruht. Sie wird über die lexikalisch-semantische Information, die mit den Wörtern kodiert ist, ermittelt. Für die Hauptkategorien (Grundformen von Verben, Nomen, Adjektive in einem Text) werden diese semantischen Klassen zusammengestellt, wobei nicht nur Einzelwörter, sondern auch die Bestandteile von Wortzusammensetzungen, etwa die von deutschen Komposita, ausgewertet werden. Die Wörter und Phrasen, die zu den häufigsten semantischen Klassen gehören, werden als Deskriptorkandidaten gesammelt und anschließend mit den Begriffen aus einem benutzerspezifischen Thesaurus abgeglichen. Sofern die Deskriptorkandidaten mit dem kontrollierten Vokabular des Thesaurus übereinstimmen und einen definierten Schwellenwert überschreiten, werden sie als Deskriptoren ausgegeben und können vom Humanindexierer validiert werden.

Der Nutzen einer solchen Vor-Indexierung eines Dokumentes durch AUTINDEX wird vor allem in der Verringerung des Arbeitsaufwandes und der Zeitersparnis gegenüber einem manuellen oder datenbankgestützten Nachschlagen im Thesaurus sowie in einer Steigerung der Konsistenz der Indexierung gesehen. Im Unterschied zu internationalen Konkurrenzprodukten ist AUTINDEX in besonderer Weise in der Lage, mit deutschsprachigen Texten umzugehen.

Wie die Konkurrenzprodukte muss AUTINDEX zur Entfaltung seines Potenzials an das zu bearbeitende Vokabular angepasst werden. Eine erste Erprobung der Rohversion von AUTINDEX im ZPID im Jahr 2004, bei der die Titel und Abstracts von 63 Dokumenten mit dem hierarchischen APA-Thesaurus als kontrolliertem Vokabular abgeglichen wurden, erbrachte denn auch noch unbefriedigende Ergebnisse: Zum

einen schlug AUTINDEX nur relativ wenige Deskriptoren vor, die auch von einem manuellen Indexierer vergeben worden waren, und zum anderen wurden nahezu alle Deskriptorenvorschläge von AUTINDEX, die über die manuelle Indexierung hinausgingen, als irrelevant erachtet. Im Folgenden wird deutlich gemacht, welche Maßnahmen ergriffen wurden, um die Leistungsfähigkeit des Systems zu erhöhen. Anschließend wird der gegenwärtige Stand zur Implementation von AUTINDEX im ZPID zusammengefasst und abschließend wird über eine erneute Evaluation der überarbeiteten Version von AUTINDEX berichtet.

Aufbau der Indikatorenliste

Verbesserungen von AUTINDEX zielen darauf ab, dass AUTINDEX

- (a) dem menschlichen Indexierer möglichst viele Deskriptoren liefert, die er für die Indexierung eines Dokuments benötigt (Erhöhung des Anteils der von AUTINDEX vorgeschlagenen Deskriptoren am Gesamt der von Indexierer vergebenen Deskriptoren), und zugleich
- (b) möglichst wenige Vorschläge macht, die unbrauchbar sind (Erhöhung des Anteils der vom Indexierer benötigten Deskriptoren am Gesamt der von AUTINDEX vorgeschlagenen Deskriptoren).

Als eine wesentliche Maßnahme zur Erreichung dieser Ziele wurde eine Überarbeitung des APA-Thesaurus vorgenommen. Die deutsche Übersetzung dieses Thesaurus, die so genannten PSYINDEX Terms, beinhalten in der Version von 2006 insgesamt 5488 Deskriptoren in englischer Sprache bzw. deutscher Übersetzung. Den deutschen Übersetzungen sind 1195 weitere deutsche Synonyme zugeordnet.

Dieses kontrollierte Vokabular wurde durch die Einführung so genannter "Indikatoren" deutlich erweitert. Die Indikatoren stellen Begriffe dar, die in enger Beziehung zu den Deskriptoren stehen, ohne notwendigerweise als direkte Synonyme zu gelten. Folgendes Beispiel verdeutlicht den Unterschied zwischen den verschiedenen Begriffsebenen: Das englische APA-Thesaurus-Schlagwort "Acalculia" wird in der deutschen Version der PSYINDEX-Terms mit "Rechenschwäche" übersetzt. Als Synonym verweist der ebenfalls in den PSYINDEX-Terms aufgeführte Begriff "Rechenunfähigkeit" auf diesen Deskriptor. Ergänzend wurden nun zu diesen Begriffen folgende Indikatoren formuliert: Akalkulie, Dyskalkulie, Dyskalkulia, Rechenstörung, mathematische Lernschwierigkeiten, rechenschwach, Probleme im Rechnen, Rechenprobleme, verzögerter Rechenerwerb. Trifft AUTINDEX im Dokument auf einen Begriff, der in der Indikatorliste (bestehend aus dem Deskriptor, seinen Synonymen und den zusätzlichen Indikatoren) enthalten ist, wird - vorausgesetzt es werden bestimmte Gewichte und Schwellenwerte erreicht - der entsprechende Deskriptor vorgeschlagen.

Zur Generierung der Indikatoren wurde ergänzend zu Fachwörterbüchern in der Regel die Datenbank PSYINDEX herangezogen: Die in PSYINDEX mit einem Deskriptor verschlagworteten Dokumente wurden gesucht, und ihre Titel und Abstracts wurden nach passenden Indikatorbegriffen für den Deskriptor durchgesehen. Insgesamt wurden so zusätzlich zu den 5488 Deskriptoren und 1195 Synonymen weitere 16792 Indikatoren ermittelt, so dass die Indikatorliste insgesamt (Thesaurusdeskriptoren, Synonyme, Indikatoren) über 23400 Begriffe umfasst.

Um die Anzahl irrelevanter Deskriptorvorschläge von AUTINDEX zu verringern, bietet die Indikatorenliste die Möglichkeit, den Vorschlag eines Deskriptors davon abhängig zu machen, in welchem Feld der entsprechende Indikatorbegriff identifiziert wurde. Unterschieden wird zwischen den Feldern Titel, Abstract sowie Autorschlagworte, aus denen die Begriffe in der Regel mit der Indikatorliste abgeglichen werden. Eine Einschränkung auf beispielsweise den Titel und/oder die Autorschlagworte bedeutet, dass ein Deskriptor nur dann vorgeschlagen wird, wenn der Indikatorbegriff in diesen Feldern auftaucht. Wird er dagegen im Abstract identifiziert, unterlässt AUTINDEX hier den entsprechenden Vorschlag. Durch dieses Vorgehen können auch Deskriptoren, die ja selbst Bestandteil der Indikatorliste sind, gekennzeichnet werden, die AUTINDEX ignorieren soll. In diesem Fall werden Titel, Abstract und Autorschlagworte aus der Analyse ausgeschlossen, das heißt, es erfolgt kein Abgleich zwischen den in diesen Feldern enthaltenen Begriffen und der Indikatorliste. Ausgeschlossen wurden vor allem Deskriptoren, die nicht nur zur Fachsprache gehören, sondern auch eine alltagssprachliche Bedeutung haben. So wurde etwa der Deskriptor "Analyse" eliminiert, weil Formulierungen wie "...wird analysiert" sehr häufig in den Dokumenten auftauchen.

Integration von AUTINDEX in die Datenbank PSYINDEX

Workflow der automatischen Indexierung

Die Ermittlung der Deskriptorvorschläge mit AUTINDEX läuft zur Zeit im Batch-Betrieb, d.h. alle Dokumente, die die notwendigen Bedingungen für eine Verarbeitung erfüllen, werden nächtlich mit AUTINDEX verarbeitet und die Deskriptorvorschläge werden in die Datenbank PSYINDEX zurückgespielt. Damit ein PSYINDEX-Datensatz verarbeitet werden kann, müssen mindestens ein Abstract (deutsch oder englisch) und der Titel jeweils mit Angabe der Sprache vorliegen. Von AUTINDEX bearbeitete Datensätze werden markiert und in späteren Durchläufen nicht berücksichtigt, es sei denn, dass der zuständige Auswerter dies explizit möchte, beispielsweise nach Veränderung des Abstracts.

Kommunikation AUTINDEX – PSYINDEX (STAR Datenbank)

Die Kommunikation zwischen AUTINDEX und der STAR Datenbank PSYINDEX läuft vollständig über die XML-Schnittstelle von STAR. Abbildung 1 zeigt den Ablauf schematisch.

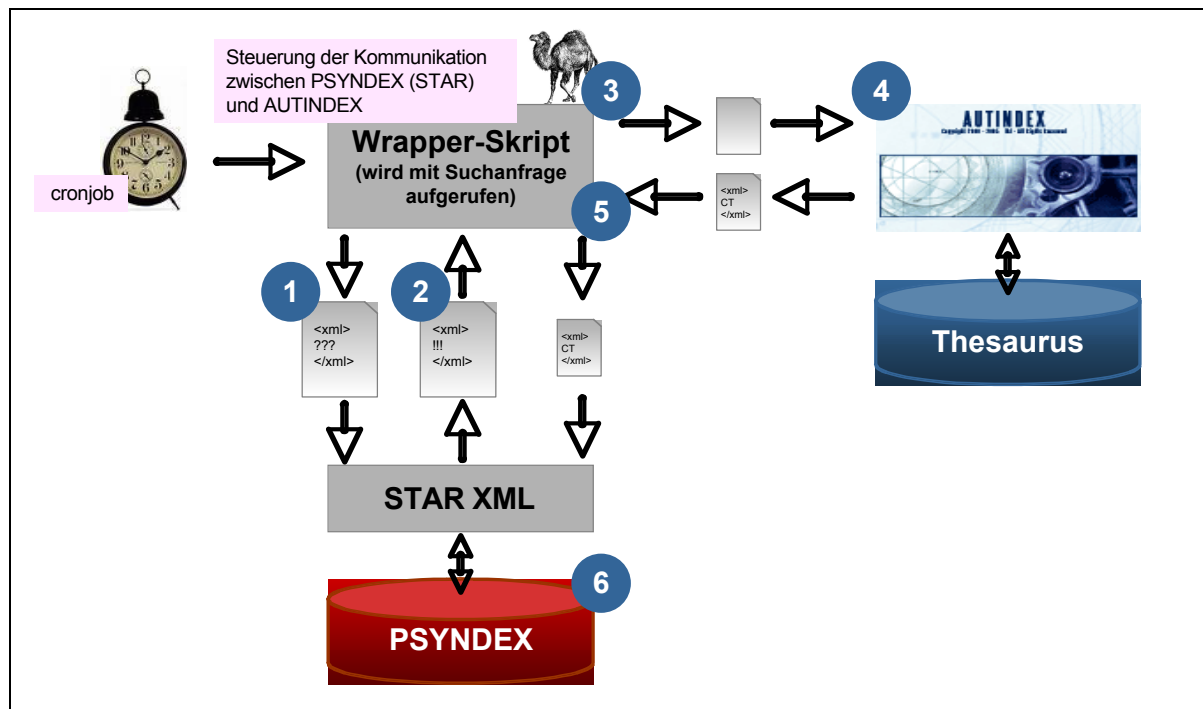


Abbildung 1: Kommunikation PSYINDEX - AUTINDEX

1. Ein durch einen so genannten cronjob angestoßenes Perl-Skript sendet eine XML-Anfrage mit einer Suche an STAR XML.
2. Als Antwort sendet STAR XML ein XML-Dokument, das alle Datensätze enthält, die den Bedingungen der Suchformulierung entsprechen¹. In Tabelle 1 sind die Felder aufgeführt, die im STAR XML Report aus PSYINDEX enthalten sind.
3. Das Perl-Skript verarbeitet das XML-Antwortdokument und splittet es für jeden Datensatz in eine einzelne Datei auf. Jede der Dateien wird anschließend an das AUTINDEX-Steuerungsskript übergeben.
4. AUTINDEX parst das übergebene XML-Dokument und ermittelt mithilfe des eingebundenen Thesaurus entsprechende Schlagworte. Als Ergebnis liefert AUTINDEX ein Dokument im STAR XML-Format zurück.
5. Das „Wrapper“-Skript ergänzt das von AUTINDEX gelieferte Dokument um eine Markierung („Dokument ist bearbeitet“) und sendet dieses an STAR XML.
6. Das Dokument wurde in PSYINDEX um die automatisch erzeugten Schlagworte ergänzt.

¹ Die Anzahl der zurück gelieferten Datensätze ist durch STAR XML auf 500 begrenzt.

Die automatisch erzeugten Schlagworte werden nicht in das Feld CT geschrieben, das üblicherweise diese Daten enthält, sondern in eine Hilfsfeld CTAI.

Feldbezeichnung	Beschreibung
DFK	Eindeutige Dokumentennummer in PSYINDEX
TI	Titel
TIL	Titelsprache (GERMAN oder ENGLISH)
TIU	Untertitel
TIUL	Untertitelsprache (GERMAN oder ENGLISH)
ABH	Hauptabstract
ABLH	Sprache des Hauptabstracts (GERMAN oder ENGLISH)
ABN	Nebenabstract
ABLN	Sprache des Nebenabstract (GERMAN oder ENGLISH)
UTE	Autorenschlagwörter ENGLISH
UTG	Autorenschlagwörter GERMAN

Tabelle 1: In AUTINDEX benutzte PSYINDEX-Felder

Komponenten

Die Indexierung mit AUTINDEX wird über ein vom IAI geliefertes Perl-Skript gesteuert, das als Eingabeparameter ein XML-File eines PSYINDEX Datensatzes erwartet (Felder siehe Tabelle 1), eine Ausgabedatei (diese wird im STAR XML Format geschrieben) und eine Konfigurationsdatei mit den entsprechenden Gewichtungen. Das selbst entwickelte „Wrapper“-Skript (vgl. Abbildung 1) steuert die Kommunikation zwischen PSYINDEX (STAR XML) und dem AUTINDEX Steuerungsskript.

Thesaurus

AUTINDEX benötigt für die Verarbeitung der Dokumente einen speziell erweiterten Thesaurus, der neben PSYINDEX Terms auch eine zu jedem Term passende Reihe von Indikatoren enthält (siehe Seite 3). Um den in einer STAR-Datenbank abgelegten erweiterten Thesaurus mit AUTINDEX nutzen zu können, wird dieser als Report aus STAR in eine Textdatei geschrieben und anschließend mit einem vom IAI gelieferten Skript für AUTINDEX kompiliert. Dieser Vorgang muss bei jeder Änderung am Thesaurus durchgeführt werden. Die Einbindung eines neuen Thesaurus geschieht über das Setzen eines symbolischen Links.

Performance

Die Verarbeitung eines Dokuments benötigt im Schnitt eine Minute, die Kompilation eines Thesaurus ungefähr 90 Minuten. Die Daten wurden gemessen auf einer SUN Ultra 10 Workstation mit UltraSPARC Iii (440 MHz). Mit modernerer Hardware lassen sich insbesondere die Zeiten für die Indexierung verringern. Dann ist auch eine interaktive Benutzung von AUTINDEX möglich.

Gewichte und Schwellenwerte

Bei AUTINDEX wird die Entscheidung darüber, ob ein ermittelter Begriff zum Vorschlag eines Deskriptors führt, über die Vergabe von Gewichten und die Festlegung eines Schwellenwerts geregelt. Trifft AUTINDEX im Text auf Begriffe oder Begriffsvarianten, die in der Indikatorenliste abgelegt sind, wird der zu dem entsprechenden Indikator passende Deskriptor gewichtet. Das mehrmalige Auftauchen eines Begriffs im Text erhöht das Gewicht des zugewiesenen Deskriptors in additiver Form.

Die Höhe des Gewichts, die beim erstmaligen Auftauchen eines Begriffs dem zugehörigen Deskriptor zugeteilt wird, kann vom Administrator des Systems festgelegt werden. Dabei kann danach differenziert werden, in welchem deutsch- oder englischsprachigen Feld (Titel/Untertitel; Abstract/Nebenabstract; Autorenschlagworte) die Begriffe auftauchen. Damit ein Begriff zum Vorschlag eines Deskriptors führt, muss die Summe der Gewichte, die ein Deskriptor auf sich vereint, einen Schwellenwert (defaultoutputlimit) erreichen. Die Höhe des Schwellenwerts und die Gewichtungen von Treffern in einzelnen Feldern kann über eine Konfigurationsdatei festgelegt werden.

Beim Abgleich der Begriffe aus den Feldern Abstract, Titel, Autorenschlagworte mit den Begriffen aus der Indikatorenliste kann zwischen einer strengen und einer weniger strengen Einstellung gewählt werden. Zur Zeit ist die weniger strenge Einstellung aktiviert, das bedeutet, dass ein Deskriptorvorschlag auch dann erfolgt, wenn der gefundene Begriff und der Begriff aus der Indikatorenliste sprachlich nicht genau übereinstimmen (Beispiel: im Abstract wurde "beruflichen Mobilität" gefunden, in der Indikatorenliste steht "berufliche Mobilität"). Die weniger strenge Einstellung führt einerseits dazu, dass ein Deskriptor auch dann vorgeschlagen wird, wenn im Text nur eine Variante des Indikatorbegriffs identifiziert wird, birgt jedoch andererseits auch das Risiko unsinniger Vorschläge (Beispiel: in einem Dokument wurde der Begriff "erbracht" als Variante des Indikatorbegriffs "Erbrechen" erkannt und führte zum Vorschlag des Deskriptors "Vomiting").

Eine weitere Einstellmöglichkeit ergibt sich darüber, dass die ermittelten Deskriptoren unterschiedlich gewichtet werden können in Abhängigkeit davon, wie genau die Begriffe im Text mit den Indikatorbegriffen übereinstimmen (kterm legt das Gewicht fest, das ein gefundener Begriff erhält, der exakt mit dem Indikatorbegriff identisch ist, DEFTERM legt das Gewicht fest, das ein Begriff bekommen soll, der als Variante eines Indikatorbegriffs erkannt wurde).

In der zurzeit benutzten Konfiguration sind die Einstellungen von AUTINDEX wie folgt optimiert:

- ▶ Bei deutschsprachigen Originaldokumenten erfolgt ein Abgleich der deutschen Titel, deutschen Abstracts und deutschen Autorenschlagworte mit der Indikatorenliste. Die englischen Nebenabstracts werden gegenwärtig nicht zur Generierung von Deskriptorvorschlägen verwendet.
- ▶ Bei englischsprachigen Originaldokumenten erfolgt ein Abgleich der englischen Titel, englischen Abstracts und englischen Autorenschlagworte mit dem Thesaurus, der keine englischsprachigen Indikatoren enthält. Entsprechend ist es nicht verwunderlich, dass AUTINDEX im englischsprachigen Bereich gegenwärtig deutlich schlechtere Ergebnisse liefert als im deutschsprachigen Bereich.

Sowohl für englische als auch für deutsche Dokumente werden die Beziehungen zwischen den Thesaurusbegriffen für die Gewichtung nicht berücksichtigt, da sich in Test dadurch keine Qualitätsverbesserung ergeben hat.

Die folgenden Tabellen enthalten die aktuelle Konfiguration der Gewichte für die einzelnen Felder, die von AUTINDEX benutzt werden. Momentan beträgt das Gewicht, das von einem Deskriptor erreicht werden muss, damit es zum Vorschlag kommt, 1 (Schwellenwert).

Fundstelle	Art des gefundenen Begriffs			Zusätzlich vergebene Gewichte		
	Indi	Thesaurus Deskriptor	Thesaurus-synonym	BT	NT	RT
Titel deu	1	0	0	0	0	0
U-titel deu	1	0	0	0	0	0
Haupt-AB deu	1	0	0	0	0	0
Neben-AB eng	0	0	0	0	0	0
Autor-Begriffe deu	1	0	0	0	0	0
Autor-Begriffe Eng	1	1	1	0	0	0

Tabelle 2: Gewichtungen für deutschsprachige Originale

Fundstelle	Art des gefundenen Begriffs			Zusätzlich vergebene Gewichte		
	Indi	Thesaurus Deskriptor	ThesaurusSynonym	BT	NT	RT
Titel eng	0	1	1	0	0	0
U-titel eng	0	1	1	0	0	0
Haupt-AB eng	0	1	1	0	0	0
Neben-AB deu	1	0	0	0	0	0
Autor-Begriffe deu	1	0	0	0	0	0
Autor-Begriffe Eng	1	1	1	0	0	0

Tabelle 3: Gewichtungen für englischsprachige Originale

Präsentation der Ergebnisse von AUTINDEX im STAR-Client

Wie oben beschrieben, werden die Ergebnisse von AUTINDEX zu den entsprechenden Datensätzen in PSYINDEX zurückgespielt. Die STAR-Client Erfassungsmasken wurden um entsprechende Felder erweitert, die diese Vorschläge enthalten. Durch die Markierung von Vorschlägen in Checkboxen kann der Auswerter diese einfach zu den Controlled Terms des Datensatzes hinzufügen (siehe Abbildung 2).

The screenshot shows the STAR Client interface with the following elements:

- Top Bar:** Buttons for 'Weiter', 'Abbrechen', 'Eintrag Eintr.', 'Speichern', and 'Hilfe'.
- Navigation:** 'Öffnen', 'Zurück', 'Inhaltliche Erfassung', 'Eintrag Löschen'.
- Menu:** 1 Abstract, 2 Inhalt, 3 Controlled Term, 4 Titelübers., 5 Nebenabstr., 6 Tests, 7 Fehler.
- Main Content:**
 - Document ID: 186368: Hell, Benedikt, 2006, 58-78, Verwendung und Einschätzung von Verfahren der internen Personalauswahl und Personalentwicklung im 10 Jahres-Vergleich
 - Table with columns: 'Controlled Term (engl oder germ)', 'Controlled Term (Übersetzung)', 'Gewichtet?'.
 - Table of 'AUTINDEX-CT-Vorschläge' with columns for suggestions and checkboxes.
 - Text area for 'Uncontrolled Terms (engl)' and 'Uncontrolled Terms (germ)'.
 - Footer: 'AUTINDEX durchführen bzw. wiederholen' (nein/ja) and 'Inhaltliche Erfassung unvollständig'.

Abbildung 2: AUTINDEX Vorschläge in der STAR Client Maske

Bisher noch nicht realisiert ist die Integration der AUTINDEX-Vorschläge in die Web-Erfassungsmaske. Dies wird Anfang 2007 umgesetzt werden.

Evaluation der aktuellen AUTINDEX-Version (November 2006)

Zur Überprüfung der Qualität des automatischen Indexierungssystems AUTINDEX wurde ein Abgleich zwischen den vom System generierten und den von Humanindexierern vergebenen Deskriptoren vorgenommen: Je mehr Übereinstimmungen zwischen den automatisch generierten und intellektuell vergebenen Deskriptoren erzielt wird, umso höher ist die Qualität der automatischen Indexierung zu bewerten, wobei selbstverständlich darauf zu achten ist, dass automatische und intellektuelle Indexierung unabhängig voneinander erfolgen. Dabei wird implizit vorausgesetzt, dass die intellektuelle Indexierung eine angemessene Beschreibung des Dokumentinhalts durch die vergebenen Deskriptoren darstellt und als Vergleichsmaßstab dienen kann. Eine korrekte Indizierung eines Dokumentes durch den menschlichen Indexierer ist somit eine notwendige Vorbedingung für die Qualitätsbewertung der automatischen Indexierung mittels eines Konsistenzvergleichs. Nur wenn diese Bedingung erfüllt ist, kann man davon ausgehen, dass Übereinstimmungen zwischen automatischer und intellektueller Indexierung auch eine hohe Indexierungsqualität widerspiegeln und in diesem Sinne die Konsistenz als Maß für die Güte einer automatischen Indexierung heranziehen. Gleichzeitig sprechen diese Überlegungen dafür, die Deskriptoren, die bei der automatischen bzw. intellektuellen Indexierung abweichend voneinander vergeben wurden, qualitativ dahingehend zu bewerten, inwieweit sie die Inhalte eines Dokumentes wiedergeben.

Vorgehen

Insgesamt 63 Dokumente aus PSYINDEX, deren manuelle Indexierung auf der Grundlage des aktuellen Thesaurus 2006 ohne Unterstützung durch die automatische Indexierung erfolgt war, wurden ausgewertet. Alle Dokumente wurden nachträglich mit AUTINDEX bearbeitet, so dass ein Vergleich der manuellen und der automatischen Indexierung möglich wurde.

Die Auswertung wurde von AG und MG vorgenommen. Bei der qualitativen Bewertung der manuell vergebenen bzw. von AUTINDEX vorgeschlagenen Deskriptoren nach einem Kategoriensystem wurde eine gemeinsame Entscheidung getroffen.

Ermittelt wurden:

- (1) Anzahl der vom Indexierer vergebenen Deskriptoren (manuelle Indexierung); dabei wurden Haupt- und Nebendeskriptoren unterschieden.

- (2) Anzahl der von AUTINDEX vorgeschlagenen Deskriptoren (automatische Indexierung)
- (3) Anzahl der übereinstimmenden manuell und automatisch vergebenen Deskriptoren
- (4) Anzahl der manuell und automatisch vergebenen Deskriptoren, die als gleichwertig betrachtet wurden (gleichwertig): Manuelle und automatische Indexierung stimmen bei diesen Deskriptoren zwar nicht sprachlich, wohl aber konzeptuell überein.
- (5) Aus der Differenz "manuelle Indexierung - Übereinstimmung" bzw. "automatische Indexierung - Übereinstimmung" ergibt sich, wie viele zusätzliche Deskriptoren manuell bzw. automatisch vergeben wurden (zusätzliche manuelle bzw. zusätzliche automatisch)
- (6) Die zusätzlich automatisch bzw. manuell vergebenen Deskriptoren wurden qualitativ nach folgenden Kriterien kategorisiert:

obligatorisch: die Deskriptoren spiegeln zentrale Inhalte des Dokumentes wider; für eine angemessene Indexierung des Dokumentes ist ihre Berücksichtigung unerlässlich

passend: die vorgeschlagenen bzw. vergebenen Deskriptoren sind zur Beschreibung der Inhalte des Dokumentes gut geeignet

irrelevant: der vorgeschlagene bzw. vergebene Begriff ist zwar aus dem Inhalt des Abstracts erklärbar, ist aber für die Verschlagwortung der sinntragenden Konzepte des Dokumentes nicht geeignet, da er allenfalls nebensächliche, meist aber sogar unbedeutende Inhalte abdeckt.

falsch: der vorgeschlagene bzw. vergebene Begriff passt nicht zum Inhalt des Dokuments.

Für die zusätzlich von der automatischen Indexierung vorgeschlagenen Deskriptoren wurden darüber hinaus folgende Kategorien berücksichtigt:

zu breit: die vorgeschlagenen Deskriptoren geben zwar im Grundsatz die im Dokument behandelte Thematik wider, allerdings auf einem zu allgemeinen Niveau. Die Inhalte des Dokuments könnten durch spezifischere Deskriptoren treffender beschrieben werden.

zu eng: die vorgeschlagenen Deskriptoren sind zu spezifisch, die Thematik wird im Dokument auf einem allgemeineren Niveau behandelt.

verwandt: zwischen den vorgeschlagenen Deskriptoren und den eigentlich treffenden Deskriptoren besteht eine inhaltliche Verbindung

Die Einordnung in die Kategorien "zu breit", "zu eng" und "verwandt" erfolgte nur dann, wenn AUTINDEX keinen treffenderen Deskriptor vorgeschlagen hatte. In diesem Fall weisen die vorgeschlagenen Deskriptoren in eine richtige Richtung, auch wenn der Vorschlag bezüglich des Spezifitätsniveaus nicht genau zutreffend ist. Hatte AUTINDEX einen treffenden Vorschlag unterbreitet, wurden die relativ zu diesem Vorschlag "zu breiten", "zu engen" und "verwandten" Vorschläge als irrelevant eingestuft.

Ergebnisse

Insgesamt wurden in den 63 ausgewerteten Dokumenten bei der manuellen Indexierung 415 Deskriptoren vergeben, wovon 187 als Haupt- und 228 als Nebendesriptoren zugewiesen wurden. Im Durchschnitt wurden somit bei einem Dokument sechs bis sieben Deskriptoren vergeben.

Von AUTINDEX wurden für die Dokumente insgesamt 549 Deskriptoren vorgeschlagen, d. h. im Durchschnitt generierte AUTINDEX für ein Dokument acht bis neun Deskriptoren.

Übereinstimmung

Zwischen der manuellen und der automatischen Indexierung gab es 179 übereinstimmende Deskriptoren, wobei 105 manuell vergebene Haupt- und 74 manuell vergebene Nebendesriptoren mit den AUTINDEX-Vorschlägen identisch waren. Zusätzlich wurden 15 Deskriptorenpaare trotz Abweichungen in der sprachlichen Formulierung als konzeptuell gleichwertig erachtet. Addiert man die sprachlich korrekt übereinstimmenden und die als gleichwertig angesehenen Deskriptoren, sind es 194 Deskriptoren, die sowohl bei der manuellen als auch bei der automatischen Indexierung angeführt wurden. Im Durchschnitt werden damit ungefähr drei Deskriptoren, die der manuelle Indexierer vergeben hat, auch von AUTINDEX erkannt.

Setzt man die Übereinstimmungen in Beziehung zur Anzahl der manuell vergebenen bzw. von AUTINDEX generierten Deskriptoren, zeigt sich, dass 46,7 Prozent der vom manuellen Indexierer vergebenen Deskriptoren auch von AUTINDEX vorgeschlagen wurden ($\frac{x}{3,08} = 100/6,59$). Gleichzeitig wurden 35,36 Prozent der von AUTINDEX vorgeschlagenen Deskriptoren auch vom manuellen Indexierer vergeben ($\frac{x}{3,08} = 100/8,71$).

Bezogen auf die ausgewertete Stichprobe bedeutet dies, dass ungefähr 35 Prozent der von AUTINDEX vorgeschlagenen Begriffe vom manuellen Indexierer übernommen werden können, da in diesem Bereich die automatisch generierten Vorschläge mit der Praxis des manuellen Indexierens übereinstimmen; diese stellen etwa 46 Prozent der von ihm gesamt vergebenen Deskriptoren dar. Die unterschiedlichen Prozentwerte verdeutlichen die Abhängigkeit von der Anzahl der jeweils manuell vergebenen bzw. vorgeschlagenen Deskriptoren, die wiederum Einfluss auf die Zahl der Übereinstimmungen hat: die Wahrscheinlichkeit einer Übereinstimmung zwischen einem manuell vergebenen und einem automatisch generierten Deskriptor steigt mit der Gesamtanzahl der vorgeschlagenen/vergebenen Deskriptoren (Beispiel: Nennen der Indexierer und AUTINDEX bei einer Gesamtmenge von 5495 Deskriptoren jeweils nur drei Deskriptoren, ist die Wahrscheinlichkeit, dass es zwischen diesen De-

skriptoren Übereinstimmungen gibt, niedriger als wenn jeweils 1000 Deskriptoren genannt würden).

	Anzahl der Deskriptoren bei insgesamt 63 Dokumenten	Durchschnitt
Manuelle Indexierung-Gesamt	415	6-7 (6,59)
Manuelle Indexierung-Hauptdeskriptoren	187	3 (2,97)
Manuelle Indexierung-Nebendeskriptoren	228	3-4 (3,62)
Automatische Indexierung	549	8-9 (8,71)
Übereinstimmungen-Gesamt	179	2,84
Übereinstimmungen-Hauptdeskriptoren	105	1,67
Übereinstimmungen-Nebendeskriptoren	74	1,17
Gleichwertig	15	
Summe Übereinstimmungen-Gesamt + Gleichwertig	194	3,08
Zusätzliche manuelle	221	3-4 (3,51)
Zusätzliche automatische	355	5-6 (5,6)
Bewertung der zusätzlich automatisch vergebenen:		
Obligatorisch	9	0,14
Passend	38	0,6
Zu breit (broader)	44	0,7
Zu eng (narrower)	1	0,015
Verwandt (related)	20	0,32
Irrelevant	174	2,76
Falsch	69	1,1
Bewertung der zusätzlich manuell vergebenen:		
Obligatorisch	56	0,89
Passend	117	1,86
Irrelevant	40	0,63
Falsch	8	0,13

Tabelle 4: Häufigkeit, Übereinstimmung und Bewertung der Vorschläge von AUTINDEX und der manuellen Indexierung

Um die Abhängigkeit der Übereinstimmungszahl von der Gesamtanzahl der angeführten Deskriptoren zu berücksichtigen wurde in Anlehnung an Formeln zur Berechnung der Indexierungskonsistenz die Übereinstimmung zwischen AUTINDEX und der manuellen Indexierung ermittelt. Der Index von Hooper (1965) berücksichtigt die Anzahl der Übereinstimmungen dividiert durch die Summe aus (a) der Anzahl der Übereinstimmungen, (b) der Anzahl der ausschließlich manuell vergebenen Deskriptoren und (c) der Anzahl der ausschließlich von AUTINDEX vorgeschlagenen Deskriptoren. An dieser Formel wurde kritisiert, dass sie Übereinstimmungen und Abweichungen nicht gleich gewichtet: Um Konsistenz zu erreichen, muss ein Deskriptor von jedem Indexierer angeführt werden - insgesamt muss der Begriff also zweimal vergeben werden; dagegen reicht die einmalige Vergabe eines Deskriptors, der nur in einem Indexat zugewiesen wurde, aus, um die Konsistenz zu beeinträchtigen. In der Formel von Rolling (1981) wurde diese Untergewichtung ausgeglichen, indem im Zähler die Zahl der Übereinstimmungen verdoppelt wird, um dem Umstand der zweimaligen Zuweisung des Deskriptors gerecht zu werden. Im Nenner der Rolling-Formel wird die Gesamtzahl der von beiden Indexierern vergebenen Deskriptoren angeführt. Abgesehen von dem unwahrscheinlichen Fall, dass überhaupt keine oder eine vollständige Konsistenz erreicht wird, führt die Verdoppelung der Übereinstimmungen in der Rolling-Formel dazu, dass ihre Konsistenzwerte höher sind als in dem Index von Hooper. Nach der Formel von R. S. Hooper ergibt sich für die untersuchte Stichprobe von 63 Dokumenten eine Konsistenz zwischen manueller und automatischer Indexierung von 25,22 Prozent. Die Berechnung der Indexierungskonsistenz nach der Formel von L. Rolling führt zu einem Konsistenzwert von 40,24 Prozent.

In einer getrennten Berechnung der Konsistenz zwischen automatischer und manueller Indexierung für die Haupt- bzw. der Nebendeskriptoren wurden nur die sprachlich übereinstimmenden Deskriptoren berücksichtigt, da die als gleichwertig erachteten Deskriptoren nicht nach Haupt- und Nebendeskriptor differenziert werden konnten. Für die Hauptdeskriptoren ergaben sich Konsistenzwerte von 16,64 (Hooper) bzw. 28,53 (Rolling), für die Nebendeskriptoren errechneten sich Werte von 10,53 (Hooper) bzw. 19,05 (Rolling). Die im Vergleich zur Gesamtauswertung niedrigeren Werte erklären sich aus der fehlenden Unterscheidung von Haupt- und Nebendeskriptoren bei der automatischen Indexierung: da dadurch immer die Gesamtzahl der Deskriptorenvorschläge berücksichtigt werden musste, liegt die Zahl der von AUTINDEX zusätzlich vorgeschlagenen Deskriptoren relativ hoch, was die Konsistenzwerte mindert. Allerdings weisen die für die Hauptdeskriptoren im Vergleich zu den Nebendeskriptoren höheren Werte darauf hin, dass AUTINDEX insbesondere die für ein Dokument zentralen Deskriptoren vorschlägt.

Der Vergleich der Werte der Gesamtauswertung mit einer früheren Überprüfung der Konsistenz zwischen der manuellen Indexierung und einer Version von AUTINDEX, in der nur die ursprünglichen Thesaurusbegriffe und -synonyme, aber noch keine

Indikatoren berücksichtigt worden waren, zeigt die durch die Hinzunahme der Indikatoren erzielte Verbesserung von AUTINDEX. In dieser Untersuchung wurden ebenfalls 63 Dokumente ausgewertet, für die die manuellen Indexierer 453 Deskriptoren vergeben und AUTINDEX (Version ohne Indikatoren) 369 Vorschläge generiert hatten. Bei einer Übereinstimmung von 94 Deskriptoren und 9 als gleichwertig erachteten Deskriptoren ergaben sich für diese Stichprobe lediglich die Konsistenzwerte 14,33 nach Hooper und 25,06 nach Rolling.

Trotz dieser Verbesserung der automatischen Indexierung zeigt sich, dass die Konsistenz zwischen automatischer und manueller Indexierung niedrigerer ist, als die Übereinstimmung zwischen zwei manuellen Indexierern. Eine von J. Beling durchgeführte Berechnung der Übereinstimmung zwischen manuellen Indexierern bei 37 Dubletten aus PSYINDEX ergab Konsistenzwerte von 42,58 nach Hooper und 56,99 nach Rolling, die damit höher liegen als die Konsistenzwerte zwischen der automatischen und manuellen Indexierung.

Bewertung der von den Indexierern bzw. AUTINDEX zusätzlich vergebenen Deskriptoren

Die zwischen der manuellen Indexierung und AUTINDEX übereinstimmenden Deskriptoren wurden nicht weiter qualitativ bewertet. Die Tatsache der Übereinstimmung wurde als hinreichender Beleg für die Angemessenheit dieser Deskriptoren zur Beschreibung des Dokumentinhalts angesehen. Daher beziehen sich die folgenden Ausführungen nur auf die Deskriptoren, die von den Indexierern bzw. von der automatischen Indexierung zusätzlich vergeben bzw. vorgeschlagen wurden.

Ausschließlich manuell vergebene Deskriptoren

Insgesamt wurden von den manuellen Indexierern zusätzlich 221 Deskriptoren vergeben, die hinsichtlich ihrer Relevanz in verschiedene Kategorien eingestuft wurden. Zur Auswertung wurden die Kategorien "obligatorisch" (56 Deskriptoren) und "passend" (117 Deskriptoren) bzw. die Kategorien "irrelevant" (40 Deskriptoren) und "falsch" (8 Deskriptoren) zusammengefasst. Damit zeigt sich, dass 78,35 Prozent der ausschließlich manuell vergebenen Deskriptoren als relevant und 21,65 Prozent als irrelevant eingestuft wurden ($\frac{x}{2,75} = \frac{100}{3,51}$ bzw. $\frac{x}{0,76} = \frac{100}{3,51}$). Dieses Ergebnis belegt die Unverzichtbarkeit der manuellen Indexierung, da einerseits die relevanten Deskriptoren von AUTINDEX nicht erkannt wurden und andererseits nur ungefähr ein Fünftel der zusätzlich manuell vergebenen Deskriptoren irrelevant waren.

Ausschließlich automatisch vorgeschlagene Deskriptoren

Von der automatischen Indexierung wurden zusätzlich 355 Deskriptoren vorgeschlagen. Zur Auswertung ihrer Relevanz wurden die in die Kategorien "obligatorisch" (9 Deskriptoren) und "passend" (38 Deskriptoren) bzw. die Kategorien "irrelevant" (174 Deskriptoren) und "falsch" (69 Deskriptoren) eingestuften Deskriptoren

zusammengefasst. Demnach wurden 13,2 Prozent der von AUTINDEX zusätzlich vergebenen Deskriptoren als relevant eingestuft ($x/0,74=100/5,6$), d. h. AUTINDEX weist den Indexierer in einem geringen Umfang auf bedeutsame Deskriptoren hin, die er bei seiner Auswertung nicht berücksichtigt hat. Mit 68,5 Prozent wurde allerdings der überwiegende Teil der ausschließlich automatisch vergebenen Deskriptoren als irrelevant oder falsch eingestuft. Die restlichen Deskriptoren verteilen sich auf die Kategorien "zu breit" (44 Deskriptoren), "zu eng" (1 Deskriptor) und verwandt (20 Deskriptoren) und machen insgesamt 18,3 Prozent der von AUTINDEX zusätzlich vergebenen Schlagworte aus. Obwohl die Deskriptoren in diesen Kategorien aufgrund des unzureichenden Spezifitätsniveaus nicht als passend für die Verschlagwortung des Dokumentinhaltes bewertet wurden, geben sie dem Indexierer Hinweise auf Deskriptorenbereiche, in denen er passende Schlagworte finden kann.

Implikationen und Schlussfolgerungen

Fasst man die Befunde zusammen, so ergibt sich auf der Basis unserer Auswertung der 63 Dokumente für den Indexierer folgende Situation:

- Er bekommt im Durchschnitt von AUTINDEX acht bis neun Deskriptoren vorgeschlagen
- Von diesen Vorschlägen wird er ungefähr drei Deskriptoren übernehmen können, da diese Vorschläge mit seiner manuellen Indexierungspraxis übereinstimmen
- Er muss zusätzlich drei bis vier eigene Deskriptoren vergeben, die AUTINDEX nicht erkannt hat
- Bei fast jedem Dokument schlägt AUTINDEX im Schnitt einen brauchbaren Deskriptor vor, den der Indexierer bei der rein manuellen Bearbeitung nicht vergeben hätte.
- In dem von AUTINDEX vorgeschlagenen Indexat wird ein weiterer Deskriptoren enthalten sein, der - obwohl er aufgrund der mangelnden Präzision nicht direkt zur Verschlagwortung verwendet werden kann - als Hinweis auf passende Deskriptoren genutzt werden kann.

Insgesamt kann somit nahezu die Hälfte der von AUTINDEX vorgeschlagenen Deskriptoren in die manuelle Indexierung einfließen. Verbesserungen gegenüber der ersten Evaluation von AUTINDEX haben sich vor allem in zwei Punkten ergeben: (1) Die Zahl der Übereinstimmungen zwischen den automatisch vorgeschlagenen und manuell vergebenen Deskriptoren ist gestiegen, (2) AUTINDEX generiert nach der Anreicherung mit den Indikatoren mehr über die manuelle Indexierung hinausgehende Vorschläge, die als passend bewertet wurden.

Bei der Interpretation dieser Befunde ist zu bedenken, dass die Zahl der ausgewerteten Dokumente relativ klein ist. Ebenso muss die hier nicht näher ausgewertete Vari-

anz bezüglich der Anzahl der Vorschläge bei den Dokumenten (0 bis 25 Vorschläge) berücksichtigt werden. In anderen Stichproben können sowohl die Zahl der von AUTINDEX generierten Vorschläge als auch der Übereinstimmungen durchaus anders ausfallen. Für 2007 ist deshalb eine computerunterstützte Überprüfung der Übereinstimmungen an einer größeren Stichprobe von etwa 1000 Dokumenten beabsichtigt.

Eine positive Bewertung von AUTINDEX könnte dadurch relativiert werden, dass die AUTINDEX-Vorschläge in unserer Auswertung auf der Basis der von den Indexierern überarbeiteten Abstracts erfolgten. Möglicherweise werden durch diese Überarbeitung die Abstracts mit Begriffen angereichert, die zu mehr relevanten Schlagwortvorschlägen führen, als dies bei einer Verarbeitung der "Rohversion" des Abstracts durch AUTINDEX der Fall ist.

Ebenso handelte es sich bei den ausgewerteten Dokumenten ausschließlich um Zeitschriftenbeiträge. AUTINDEX wird bei Zeitschriftenbeiträgen und Sammelwerksbeiträgen in der Regel bessere Ergebnisse liefern als bei Büchern, weil dort die Abstracts wegen ihrer Länge bzw. durch Wiedergabe der kompletten Inhaltsverzeichnisse sehr viele sinntragende Begriffe enthalten und somit auch sehr viele Deskriptoren vorgeschlagen werden, die in der manuellen Indexierung eher nicht berücksichtigt werden.

Konsistenzmindernd wirkt sich die Zuordnung eines Indikators zu mehreren Deskriptoren aus. Diese Zuordnung kommt vor, wenn die beiden Deskriptoren inhaltlich ähnlich sind, oder wenn der Begriff je nach Kontext mehrere inhaltliche Bedeutungen (etwa Ball im Sinne von Fußball oder Tanzveranstaltung) hat: ein Indikator besitzt in diesem Fall keine Trennschärfe in Bezug auf ähnliche Deskriptoren bzw. ist nicht eindeutig einem Gegenstandsbereich zugeordnet. Bei der Erstellung der Indikatorenliste wurde in solchen Fälle in der Regel entschieden, dem Indikator mehrere passende Deskriptoren zuzuweisen. Für die automatische Indexierung bedeutet dies allerdings, dass beim Auftreten des Indikatorbegriffs im Dokument und Überschreiten des Schwellenwertes sämtliche dem Indikator zugeordneten Deskriptoren vorgeschlagen werden. Da von diesen aber zumeist nur einer zum Inhalt des Dokumentes passt (der dann im günstigsten Fall auch von manuellen Indexierer vergeben wurde), erhöht sich die Zahl der ausschließlich automatisch vergebenen, aber irrelevanten Deskriptoren, was zu einer verringerten Konsistenz zwischen automatischer und manueller Indexierung führt.

Um die Leistung von AUTINDEX aufrecht zu erhalten und zu verbessern, ist eine nachhaltige Überarbeitung der Indikatorenliste Voraussetzung: die getroffene Zuordnung der Indikatoren zu den Deskriptoren bedarf der ständigen Überprüfung, neue Deskriptoren müssen eingearbeitet werden, es ist zu kontrollieren, bei welchen sehr häufig vorgeschlagenen Deskriptoren die Auftretenshäufigkeit durch Einschränkungen der Indikatoren auf die Felder TI und UT verringert werden soll, die

bestehende Liste der Indikatoren kann ergänzt werden. Die Ergänzung der Liste könnte durch die Humanindexierer erfolgen, indem diese während der praktischen Dokumentationsstätigkeit die zu einem Deskriptor passenden Indikatoren in das Datenerfassungsprogramm eingeben.

Weitere Optimierungsmöglichkeiten zeigen sich hinsichtlich der Bearbeitung englischsprachiger Haupt- und Nebenabstracts. Bei englischsprachigen Dokumenten wird AUTINDEX zur Zeit zur Indexierung von englischsprachigen Titeln, Hauptabstracts und Autorenschlagworten eingesetzt. Da es aber keine Indikatorenliste in englischer Sprache gibt, werden bei englischen Abstracts die Vorschläge nur aufgrund eines Abgleichs der Begriffe im Text mit dem APA-Thesaurus generiert, was die Qualität der Vorschläge beeinträchtigt. Da ebenso ein Ausschluss von Deskriptoren und zugehörigen Thesaurussynonymen nicht möglich ist, werden relativ viele irrelevante Deskriptoren vorgeschlagen. Die Vorschläge von AUTINDEX sind deshalb bei englischsprachigen Originalarbeiten eher schlechter als im deutschsprachigen Bereich. Aus diesem Grund wurde bisher darauf verzichtet, in deutschsprachigen Dokumenten vorhandene englischsprachige Nebenabstracts bei der Erstellung der Deskriptorvorschläge zu berücksichtigen. Durch eine Überarbeitung des englischsprachigen Thesaurus könnte diese potentielle Ressource zur Verbesserung der automatischen Indexierung erschlossen werden. Da eine mit dem bisherigen Vorgehen vergleichbare Zusammenstellung englischer Indikatoren allerdings zu aufwendig ist, muss sich diese Überarbeitung darauf beschränken, durch eine Art Stoppwortliste die Deskriptoren, die häufig vorgeschlagen werden, obwohl sie irrelevant sind, auszuschließen. Inwiefern diese Maßnahme zu einer Verbesserung der Leistung von AUTINDEX führt, ist durch weitere Testdurchgänge zu überprüfen.

Bisher werden die an einem Tag formal erfassten Dokumente, bei denen ein Abstract vorhanden ist, über Nacht von AUTINDEX bearbeitet. Ebenso kann ein Humanindexierer AUTINDEX veranlassen, die automatische Indexierung zu wiederholen. Allerdings muss er dann bis zum nächsten Tag auf die überarbeiteten Deskriptorvorschläge warten. Günstiger wäre hier eine Einstellung, die eine direkte Bearbeitung eines Abstracts durch AUTINDEX ermöglicht. Dadurch könnten für stark von der "Rohversion" abweichende oder ganz neu eingegebene Abstracts ohne längere Zeitverzögerung Deskriptorenvorschläge generiert werden. Ebenso wäre ein solches Vorgehen eine wichtige Voraussetzung, um AUTINDEX auch den außerhalb der ZPID tätigen Mitarbeitern (Fremdauswerter) zur Verfügung zu stellen, sofern diese ihre Auswertungen über ein Web-Formular tätigen. Daher ist für das nächste Jahr die Entwicklung eines entsprechenden Web-Interfaces geplant.

Insgesamt wird das Vorhaben, AUTINDEX in der Praxis einzusetzen, durch den Befund gestärkt. Das Verfahren erweist sich keineswegs als geeignet, eine hochwertige intellektuelle Indexierung zu ersetzen (das war auch nicht erwartet worden), es wird aber für gut genug erachtet, die Indexierungsarbeit zu unterstützen und damit Aufwände zu reduzieren und die Qualität zu erhöhen. Dabei scheint AUTINDEX durch-

aus Vorschläge zu generieren, die bei einer rein manuellen Indexierung übersehen worden wären. Möglicherweise hat AUTINDEX auch eine normierende Wirkung und trägt dadurch zur Konsistenz der Indexierung bei. Allerdings ist auch zu bedenken, dass AUTINDEX insbesondere wenig erfahrene Indexierer dazu verleiten könnte, die Vorschlagliste ungeprüft und unkritisch zu übernehmen, sich etwa mit einem vorgeschlagenen "Broader-Term" zufrieden zu geben, anstatt durch eigene Thesaurusarbeit passendere "Narrower-Terms" zu ermitteln oder auf die Vergabe von weiteren sinnvollen, aber von AUTINDEX nicht vorgeschlagenen Deskriptoren zu verzichten. Derartige Auswirkungen müssen in zukünftigen Evaluationen von AUTINDEX überprüft werden.

Literatur

- Beling, J. (2006). Inter-Indexer-Konsistenz in PSYNDEX. Auswertung in Anlehnung an eine Untersuchung der APA bei PsycINFO. *Unveröffentlichtes Manuskript.*
- Gerards, A. & Gerards, M. (2004). Erste Auswertung AUTINDEX (06.07.2004). *Unveröffentlichtes Manuskript.*
- Hooper, R. S. (1965). Indexer consistency tests: origin, measurement, results, and utilization. Bethesda, MD: IBM Corporation.
- Rolling, L. (1981). Indexing consistency, quality, and efficiency. *Information Processing & Management*, 17 (2), 69-76.