

*Johannes Akkach (Heidelberg)*

**Die Generierung einer terminologischen Datenbank für  
Übersetzer und Mediziner**



**Editors:**

Viktorija Bilić  
Anja Holderbaum  
Anne Kimmes  
Joachim Kornelius  
John Stewart  
Christoph Stoll

**Publisher:**

Wissenschaftlicher Verlag Trier

*Johannes Akkach (Heidelberg)*

## **Die Generierung einer terminologischen Datenbank für Übersetzer und Mediziner**

### **Abstract:**

Zur Recherche von Terminologie in der Medizin steht heute eine große Zahl an Wörterbüchern im Internet und als Printausgabe zur Verfügung. Viele dieser Ressourcen sind jedoch Bestandteil eines größeren allgemeinsprachlichen Wörterbuchs und vor allem auf die dekodierende Verwendung ausgerichtet. Für den idiomatisch korrekten Sprachgebrauch in der fremden Sprache und für die Bedeutungserschließung bei der vorbereitenden Terminologierecherche ist die Mikrostruktur solcher Wörterbücher in der Regel nicht ausführlich genug.

Ein online verfügbares, für den enkodierenden Gebrauch konzipiertes medizinisches Wörterbuch der englischen und deutschen Sprache kann sowohl Übersetzern als auch Mediziner durch seine ausführliche terminologische Eintragsstruktur einen erheblichen Mehrwert bei der Textproduktion bieten. Kontextbeispiele, Definitionen und fachsprachliche Kollokationen unterstützen Übersetzer im digitalen Übersetzungsprozess beim Verfassen fremdsprachlicher Texte.

There is a wealth of resources for researching medical terminology available to professional translators on the Internet and in print. Many of these are, however, part of a more comprehensive, general-purpose dictionary with no particular focus on technical language. The scope of such resources is often limited to the decoding of a source language term and not intended to facilitate text production in the target language.

Therefore, an online dictionary with extended medical terminology would be a valuable resource both for health professionals and for technical translators. With a standardized and comprehensive terminological structure, such resources can help users to produce complex yet precise wordings in their non-native target language. Examples of useful information for text production include definitions, context sentences and collocations for each term.

As part of the research project *Teaching Medical Translation* at the University of Heidelberg, approximately 78 bachelor students in the class of 2011 extracted medical terminology from case studies provided by the Indiana University School of Medicine, USA, using a predefined template. The present contribution deals with the processing of the collected data to create an online database which will power a web-based dictionary.

### **Keywords:**

Terminology Database; Medical Translation; Online Dictionary; Technical Translation

Terminologiedatenbank; Medizinische Übersetzung; Online-Wörterbuch; Fachübersetzung

**Hinweis:**

In der vorliegenden Arbeit wurde aus Gründen der besseren Lesbarkeit weitgehend auf eine geschlechtsneutrale Formulierung verzichtet. Es sind jedoch immer beide Geschlechter im Sinne der Gleichbehandlung angesprochen.

**Inhalt:**

1	Das Forschungsprojekt Teaching Medical Translation.....	3
2	Die terminologische Eintragsstruktur .....	4
2.1	Sprachdaten .....	4
2.2	Begriffsdaten .....	5
2.3	Kollokationen .....	6
2.4	Verwaltungsdaten .....	7
3	Eine einheitliche Vorlage zur Terminologieerfassung mit Microsoft Excel .....	7
4	Die Terminologiedatenbank.....	10
4.1	Datenstruktur .....	10
4.2	Das Suchformular auf der Website .....	11
4.3	PHP als verbindendes Element .....	12
5	Möglichkeiten zur Erweiterung.....	13
5.1	Auswertung der Morphologiedaten .....	13
5.2	Auswertung der Kollokationsdaten.....	13
5.3	Export von Terminologie .....	14
6	Zusammenfassung und Ausblick .....	14
7	Literaturverzeichnis.....	15
8	Abbildungsverzeichnis.....	16
9	Tabellenverzeichnis.....	16
10	Anhang .....	17

## 1 Das Forschungsprojekt Teaching Medical Translation

Teaching Medical Translation (TMT) ist eine Kooperation der Englischen Abteilung des Seminars für Übersetzen und Dolmetschen (SUED) an der Universität Heidelberg mit der Indiana University School of Medicine (IUSM) in den USA. Sie steht unter der Leitung von Prof. Dr. Joachim Kornelius (Leiter der Englischen Abteilung des SUED), in Zusammenarbeit mit Prof. Dr. Patrick Bankston (Assistant Dean Director der IUSM Northwest in Gary, Indiana). Im Projekt TMT entstanden in den vergangenen Jahren mehrere, in der wissenschaftlichen Reihe Lighthouse Unlimited veröffentlichte, terminologische Abschlussarbeiten von Studierenden sowie zwei Dissertationen (Orbán 2008 und Hüging 2011).

Im Frühjahr 2011 wurden am Seminar für Übersetzen und Dolmetschen die Abschlussarbeiten im Fach B.A. Übersetzungswissenschaft in das Projekt TMT integriert. Bei diesem Projekt stand eine möglichst standardisierte und einheitliche Erfassung der medizinischen Fachtermini durch alle beteiligten Studierenden im Vordergrund. Als Textkorpus dienten Fallstudien der Indiana University School of Medicine, die den dortigen Studierenden im Rahmen des Curriculums *Problem-Based Learning* zur selbstständigen Bearbeitung vorgelegt worden waren.

Die Studierenden wurden in einer Lehrveranstaltung (Kimmes und Kornelius 2011) auf das Projekt vorbereitet. In der Übung wurden die Grundlagen der Terminologielehre wiederholt und vertieft und es wurde das Konzept des Problem-Based Learning eingeführt.

Für die praktische terminologische Arbeit erhielten die Studierenden auf der E-Learning-Plattform *Moodle* eine speziell entwickelte Vorlage als Microsoft-Excel-Datei. In dieser Vorlage können die terminologischen Daten nach einer vorgegebenen Struktur intuitiv eingegeben werden. Für technische Probleme und formale Fragen zum Erfassungsprozess stand ein Diskussionsforum für alle Kursteilnehmer zur Verfügung. Aufbau und Funktionen der MS-Excel-Vorlage sind in Kapitel 3 beschrieben.

## 2 Die terminologische Eintragsstruktur

Bei der Gestaltung der terminologischen Einträge wurde auf die von Nicole Keller in ihrer Dissertation entwickelte Systematik zurückgegriffen (Keller 2006). Die Mikrostruktur nach Keller mit Eintragsfeldern für Sprachdaten, Begriffsdaten, Kollokationen und Verwaltungsdaten diente als Grundlage für eine Weiterentwicklung in die der Datenbank zugrunde liegende Struktur (Abbildung 1). Die gezeigten Eintragsfelder werden jeweils vollständig für die deutsche und für die englische Sprache gespeichert.

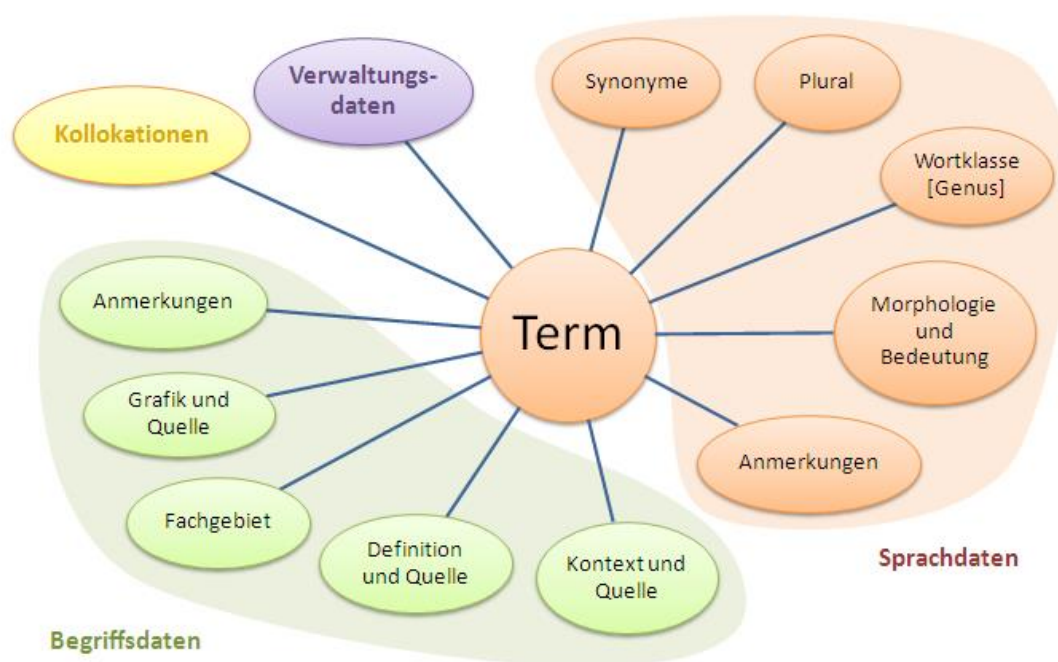


Abbildung 1: Die Eintragsstruktur der Datenbank auf Grundlage von N. Keller (2006: 66)

### 2.1 Sprachdaten

In der Eintragsstruktur nach Keller sind unter der Kategorie *Sprachdaten* die zu einem Term gehörenden Eintragsfelder gruppiert, die mit der Benennung in Zusammenhang stehen und von der Begriffsebene unabhängig sind. Zu diesen Eintragsfeldern gehören neben der Benennung als zentralem Element die Synonyme, Antonyme und die orthographischen Varianten, außerdem grammatische Angaben, eine kurze morphologische Beschreibung und eventuelle Anmerkungen (Keller 2006: 61f).

Die Sprachdaten wurden vor allem in Bezug auf die Morphologie erweitert, um mit etymologischen Angaben die Bedeutungserschließung des stark spezialisierten medizinischen Fachvokabulars zu erleichtern. Im Gegenzug wurden orthographische Varianten wie Kurzformen, Langformen und regional spezifische Schreibweisen in das Eintragsfeld *Anmerkungen* ausgelagert und auf eine generelle Recherche und Erfassung von Antonymen verzichtet.

## 2.2 Begriffsdaten

Die Begriffsdaten umfassen bei Keller zuerst die zentrale Definition zu einem Term und Kontextbeispiele, jeweils mit Quellenangabe. Ergänzt werden diese Eintragsfelder durch die Angabe eines Fachgebiets, das in Klassifikation und Teilbestand untergliedert ist. Eine Grafik, Analogieverweise und spezielle begriffsbezogene Anmerkungen vervollständigen diese Kategorie.

Zur eindeutigen Identifikation eines Terms und zur Bedeutungsklä rung bei Polysemie werden Grafiken mit ihrer Quellenangabe im terminologischen Eintrag direkt eingebunden, wobei die Größe gegebenenfalls automatisch bis zu einem Mindestwert reduziert wird. Es werden nur solche Grafiken direkt angezeigt, die über eine entsprechende Nutzungslizenz freigegeben sind; anderenfalls erscheint nur ein *Hyperlink* auf die jeweilige Webseite als Quelle.

Definitionen und Kontextbeispiele werden unter Wahrung der wissenschaftlichen Zitierweise direkt aus der Quelle übernommen. Sofern eine Quellenangabe eine Internetadresse enthält, wird diese in einen Hyperlink umgewandelt, über den die Quelle in einem neuen Browserfenster geöffnet werden kann.

Für die Klassifikation der medizinischen Fachtermini wurde in MS-Excel ein Eintragsfeld als *Pickliste* angelegt, in dem das medizinische Fachgebiet auf Englisch aus einer Liste ausgewählt wird. Die Nomenklatur entspricht nahezu vollständig der Systematik der U.S. National Library of Medicine, einer Fachbibliothek des US-amerikanischen Gesundheitsministeriums.

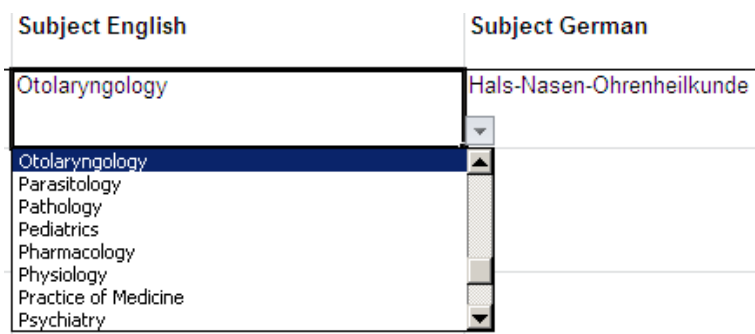


Abbildung 2: Auswahl des Fachgebiets in MS-Excel

### 2.3 Kollokationen

Kollokationen werden in sieben Kategorien pro Sprache hinterlegt. Die ersten sechs Kategorien entsprechen der Typologie nach Franz Josef Hausmann (1989: 1010), in der die Lexeme einer Kollokation als Basis und Kollokator unterschieden werden. Eine zusätzliche siebte Kategorie dient der Erfassung von *grammatical collocations*, sog. Konstruktionen. In diesem Eintragsfeld werden Wortgruppen gespeichert, die durch ihre gebräuchliche Verwendung in der medizinischen Fachsprache nützliche Zusatzinformationen zu einem Term liefern können. Hierzu zählen vor allem die *grammatical collocations* im Sinne der Beschreibung von Benson, Benson und Ilson (1986: ix), beispielsweise englische Ausdrücke mit einem *phrasal verb*. Eine Übersicht der Typologie ist in der nachstehenden Tabelle mit hervorgehobener Basis dargestellt.

Kollokationstyp
a) <b>Substantiv</b> + Adjektiv
b) <b>Substantiv</b> (Subjekt) + Verb
c) Verb + <b>Substantiv</b> (Objekt)
d) <b>Verb</b> + Adverb
e) <b>Adjektiv</b> + Adverb
f) Substantiv (+ Präposition) + Substantiv
g) Grammatische Kollokationen

Tabelle 1: Die erweiterte Kollokationstypologie auf der Basis von Hausmann

## 2.4 Verwaltungsdaten

Bei den Verwaltungsdaten wurde die Eintragsstruktur nach N. Keller soweit reduziert, dass nur eine eindeutige Term-Identifikationsnummer, das Eintragsdatum für Deutsch und Englisch und der Name des Eintragsverfassers gespeichert werden. Ein nahezu vollständiger Eintrag, wie er auf der Website angezeigt wird, ist beispielhaft zum Term *Hämoglobin* im Anhang dargestellt.

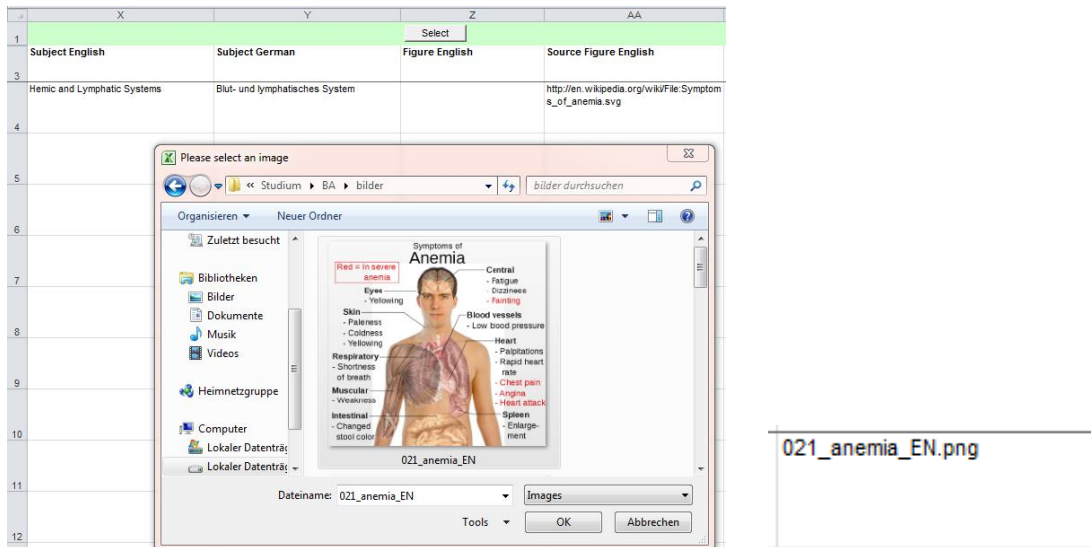
## 3 Eine einheitliche Vorlage zur Terminologieerfassung mit Microsoft Excel

Aufgrund der großen Zahl der beteiligten Studierenden im Fach B.A.

Übersetzungswissenschaft erschien es sinnvoll, eine möglichst einfache und einheitliche Eingabemethode für die aus den Textkorpora extrahierten Fachtermini zu entwickeln. Um eine möglichst intuitive und übersichtliche Bedienung zu ermöglichen, wurde dafür eine Vorlage im Tabellenkalkulationsprogramm *Microsoft Excel* entwickelt. Die über das E-Learning-System der Universität Heidelberg (Kimmes und Kornelius 2011) bereitgestellte Datei enthält eine definierte Eintragsstruktur. Die Eintragsfelder entsprechen den Spalten der Tabelle und dienen zum großen Teil der manuellen Eingabe von Text. Einige Eintragsfelder sind mit speziellen Funktionen ausgestattet.

Im Zuge der Recherche ist die Einbettung von Bildern zu einem Term sinnvoll. Zum einfachen Speichern von Grafiken verfügt die MS-Excel-Vorlage über einen Auswahldialog, in dem eine Datei auf der lokalen Festplatte gewählt werden kann. Eine solche Datei kann etwa eine aus dem Internet heruntergeladene Illustration sein, die über eine entsprechende Lizenz für diese Nutzung freigegeben wurde. Der Dateiname der Grafik wird als Text in einem Eintragsfeld des Terms hinterlegt. Sobald die Datei zusammen mit den Eintragsdaten hochgeladen wurde, entsteht dadurch in der Datenbank eine Zuordnung zu dem zugehörigen Term. Der Browser kann dann zu einem Term den zugehörigen Dateinamen auslesen und die entsprechende Grafik inkl. Quellenangabe anzeigen.



Abbildung 3: Auswahl einer lokal gespeicherten Grafik<sup>1</sup>

Alternativ zum Auswahldialog können Bilder auch direkt von Mediendatenbanken eingebunden werden, die ihre Inhalte unter einer entsprechenden Lizenz freigeben (beispielsweise Wikimedia Commons). Dazu kopiert man einen HTML-Codeausschnitt, der auf der Quellwebseite bereitgestellt wird, in das Eintragsfeld. In diesem Fall werden keine Dateien hochgeladen. Stattdessen stellt der Browser die Zuordnung her, indem die Grafik direkt von der Quellwebseite eingebunden wird.

Im Hinblick auf die morphologischen Informationen in der Terminologiedatenbank wurden die von N. Keller beschriebenen Angaben ausgeweitet. Zur Analyse der Morpheme dient die von Chabner (2009: 2ff.) entwickelte Klassifikation in vier Morphemtypen: *prefixes*, *roots*, *combining vowels* und *suffixes*. Jeder Typ erfüllt eine bestimmte Funktion innerhalb des Terms und kann speziell in der medizinischen Fachsprache bei vielen Termini in identischer Ausprägung vorkommen.

Die Morphologie jedes Terms wird in zwei getrennten Eintragsfeldern eingegeben. Dafür wurde eine Notation mit Klammern entwickelt. Jedes Morphem wird gemäß seinem Typ automatisch von einem Klammerpaar umschlossen.

<sup>1</sup> Quelle der gezeigten Grafik: [http://commons.wikimedia.org/wiki/File%3ASymptoms\\_of\\_anemia.svg](http://commons.wikimedia.org/wiki/File%3ASymptoms_of_anemia.svg), By Mikael Häggström [Public domain], via Wikimedia Commons, Zugriff am 14.09.2011

Typ	Notation	
prefixes	spitze Klammern	<>
roots	runde Klammern	()
combining vowels	eckige Klammern	[]
suffixes	geschweifte Klammern	{}

Tabelle 2: Die Klammern zur Notation der Morphemtypen nach Chabner (2009)

(gastr)	[o]	(enter)	[o]	{logy}
stomach	comb. vowel	intestines	comb. vowel	the study of

Tabelle 3: Beispiel einer morphologischen Notation

Dieses Verfahren ermöglicht später die maschinelle Auswertung jedes Terms nach seinen Morphemen und deren Typen (siehe auch Abschnitt 5.1). Zusätzlich zur Notation mit Klammern werden die Bedeutungen der Morpheme nochmals separat aufgeführt. Beide Eintragsfelder werden nach einem standardisierten Ablauf über ein Formular ausgefüllt. Es können bis zu zehn Morpheme pro Term gemäß Chabners Klassifikation erfasst werden.

Abbildung 4: Eingabe der morphologischen Daten zum Term *anemia*

## 4 Die Terminologiedatenbank

Bei der Entwicklung einer Terminologiedatenbank waren Datensicherheit, Flexibilität, Kostenfreiheit und die Unabhängigkeit von der verwendeten Plattform ausschlaggebend. Zur Umsetzung wurde das weit verbreitete und bewährte Datenbankverwaltungssystem MySQL gewählt. MySQL ermöglicht es, eine webbasierte Datenbank anzulegen, aus der mittels einer Skriptsprache Datensätze abgefragt werden.

Als Skriptsprache wird aufgrund seiner weiten Verbreitung (W<sup>3</sup>Techs, „Usage of server-side programming languages for websites“) PHP verwendet. Die Plattform für die Ausführung des PHP-Quellcodes ist der am häufigsten verwendete (W<sup>3</sup>Techs, „Usage of web servers for websites“) Webserver *Apache* unter dem Betriebssystem Linux.

### 4.1 Datenstruktur

Die Datenbank entspricht dem etablierten relationalen Datenbankmodell auf Basis der Standard-Abfragesprache SQL (Structured Query Language). In der Datenbank gibt es zwei Haupttabellen. Sie enthalten einen großen Teil der terminologischen Eintragsstruktur. Einige Daten wurden in Hilfstabellen ausgelagert, um Redundanzen weitgehend zu vermeiden und die Datenspeicherung flexibler zu gestalten. Sie sind mit den Haupttabellen über Beziehungen verbunden. Unter anderem umfassen diese Hilfstabellen:

- eine Liste der Kollokationen,
- eine Liste der Grafiken mit Quellenangaben,
- eine Liste der morphologischen Daten,
- eine Liste der Fachgebiete,
- eine Liste der Autoren aller Einträge.

Wenn ein Nutzer einen Term abfragt, werden alle zugehörigen Informationen aus den Tabellen wieder zusammengeführt und gemeinsam angezeigt, sodass der terminologische Eintrag in einer einzigen Tabelle erscheint (siehe Abbildung 7 im Anhang). Um Probleme bei der Darstellung von Sonderzeichen zu vermeiden, ist die Datenbank für die Kodierung UTF-8 ausgelegt. Dabei ist für nahezu jedes in menschlichen Sprachen vorkommende

Schriftzeichen eine Darstellung vorgesehen. So ist sowohl die Darstellung von Sonderzeichen möglich als auch die Erweiterung auf Sprachen mit nichtlateinischem Alphabet.

Die Struktur wurde so gewählt, dass die erfassten Daten später einfach weiterverarbeitet und analysiert werden können. Dieses Verfahren ermöglicht unter anderem die Speicherung beliebig vieler Kollokationen, mit Angabe des Kollokationstyps und einem Verweis auf den Term. Somit bilden die Kollokationstabellen für Deutsch und Englisch eine eigene umfangreiche Kollokationsdatenbank.

collID	termID	Collocation	collTypeID
1	1	to convert hemoglobin	1
35	1	adult hemoglobin	2
36	1	fetal hemoglobin	2
57	1	hemoglobin transports	3
58	1	hemoglobin consists	3
60	1	hemoglobin binds	3
61	1	hemoglobin combines	3
84	1	hemoglobin's affinity	4
85	1	hemoglobin's ability	4
86	1	hemoglobin molecule	4

Abbildung 5: Kollokationen zum Term *hemoglobin* in der Datenbank

#### 4.2 Das Suchformular auf der Website

Zur Terminologiesuche steht den Nutzern der Website ein Online-Formular zur Verfügung. Es enthält eine kurze Erklärung zur verwendbaren Syntax und ist auf Deutsch und Englisch verfügbar. Als Suchkriterien stehen die Sprachrichtung, der Term, die Wortklasse, das Fachgebiet und der Autor eines Eintrags zur Verfügung. Bis auf das Feld *Term* handelt es sich um Felder, die nur die Auswahl aus einer definierten Liste erlauben. Im Feld *Term* kann mit Platzhaltern nach Teilwörtern gesucht werden, indem das Zeichen \* eingegeben wird. Um die Belastung des Webservers durch Suchanfragen zu minimieren, werden unabhängig von der Zahl der tatsächlichen Treffer nur die ersten 25 Ergebnisse angezeigt. Außerdem ist eine Mindestzahl von Zeichen zur Suche erforderlich.

### 4.3 PHP als verbindendes Element

Die Skriptsprache PHP wird genutzt, um bei der Suchanfrage eines Nutzers alle Treffer aus der Datenbank in tabellarischer Form auf einer Webseite anzuzeigen. Im entsprechenden Skript werden die SQL-Befehle abhängig vom eingegebenen Suchbegriff ausgeführt. Eine Sicherheitsfunktion blockiert schädlichen Code, den ein Angreifer als Suchbegriff getarnt an die Datenbank senden könnte.

In PHP wurden auch Funktionen implementiert, die der Benutzerfreundlichkeit und Übersichtlichkeit dienen. Dazu gehören die anklickbaren Hyperlinks und die gegliederte Anzeige der Suchergebnisse (eine Seite pro Eintrag). Auf jeder Ergebnisseite ist unten ein Link enthalten, der zum Suchformular zurückführt.

**Suchergebnisse**

Zu Ihrer Suche gab es **4** Ergebnisse:

[1] [2] [3] [4]

Term	Magnetfeldgradient
Synonym	Gradient

Abbildung 6: Verteilung der Suchergebnisse auf eine Seite pro Eintrag

Weiterhin dient PHP zum Einpflegen neuer Datensätze in die Datenbank. Dazu werden zuerst die von den Studierenden eingereichten Terminologiedaten im Microsoft-Excel-Format in die Datenbank importiert, wonach sie in Form einer (temporären) Hilfstabelle vorliegen. Durch das Aufrufen eines Update-Skriptes durch den Administrator des Systems werden die Datensätze in die eigentliche Datenbank überführt, sofern von diesem Autor noch kein Term existiert.

Die Hilfstabelle kann anstelle eines automatischen Imports auch durch eine manuelle Eingabe der Studierenden mit Inhalt gefüllt werden. Eine mögliche Erweiterung der Datenbank, bei der dieser Weg genutzt wird, ist im Folgenden beschrieben.

## 5 Möglichkeiten zur Erweiterung

Für eine zukünftige Erweiterung des Datenbestandes wäre ein webbasiertes Eingabesystem für die Terminologie eine sinnvolle Ergänzung oder Alternative zur MS-Excel-Vorlage. Studierende könnten so unmittelbar ihre Einträge online einpflegen und wären nicht mehr auf ein Programm angewiesen, das an einen Computer gebunden ist und dessen Lizenz gegebenenfalls erst erworben werden muss. Wie etwa bei der E-Learning-Plattform Moodle lassen sich nicht autorisierte Zugriffe über eine Anmeldung mit Passwort unterbinden.

Jeder Kursteilnehmer könnte innerhalb eines festgelegten Zeitraums neue Einträge anlegen und diese jederzeit einsehen und bearbeiten. Die Betreuer hätten nach Ablauf der Frist die Möglichkeit, alle erfassten Termini zu validieren und danach in die Datenbank aufzunehmen. Für diese Erweiterung bietet sich die Verwendung der existierenden PHP-Skripte in Verbindung mit der im vorigen Abschnitt erwähnten Hilfstabelle an.

### 5.1 Auswertung der Morphologiedaten

Die angesprochene morphologische Notation ist in der Datenbank bereits mit einem Eintragsfeld enthalten. Sie ermöglicht bei einer späteren Erweiterung der Funktionen, computergestützt den Typ jedes einzelnen Morphems zu erkennen und unterschiedliche Analysen der Termini zu erstellen. Beispielsweise wäre eine Abfrage aller Termini mit einem bestimmten Suffix denkbar. Die Ergebnisse können dann wiederum nach bestimmten Kriterien ausgegeben werden, etwa nach Präfixen oder nach Fachgebieten sortiert.

### 5.2 Auswertung der Kollokationsdaten

Die Kollokationen werden so gespeichert, dass sie im Kontext ihrer Termini möglichst einfach maschinell ausgewertet werden können. Da die Basis einer Kollokation dem Term entspricht, ist es sinnvoll, eine automatische Erkennung zu programmieren, in der Basis und Kollokator voneinander unterschieden werden. Mit diesen Informationen ließe sich beispielsweise eine Suche nach allen Basen realisieren, die einen bestimmten Kollokator gemeinsam haben. Da die Datenbank auf mögliche Erweiterungen durch zukünftige terminologische Arbeiten

ausgelegt wurde, kann der Kollokationsbestand bei Fortführung des Projekts TMT noch vergrößert werden.

### 5.3 Export von Terminologie

Vor allem im praktischen Übersetzungsprozess ist es bequem und zeitsparend, die medizinische Terminologie in vorhandene Übersetzungssoftware zu integrieren, um häufiges Nachschlagen auf der Website zu vermeiden. Die Nutzer könnten beispielsweise über einen Link bei den Suchergebnissen die gefundene Terminologie als Datei herunterladen.

Für dieses Verfahren bietet es sich an, die Suchergebnisse neben der Anzeige im Browser auch im Format TBX (*TermBase eXchange*) bereitzustellen. Als Standard zum Austausch von Terminologiedaten ermöglicht TBX die Nutzung des Datenbestandes in Translation-Memory-Systemen wie *SDL Trados* und dessen Terminologieverwaltung *MultiTerm*.

## 6 Zusammenfassung und Ausblick

Die Datenbank im Projekt TMT ist ein Versuch, das webbasierte Nachschlagen von Fachtermini mit einer Eintragsstruktur zu verbinden, die über ein klassisches Online-Wörterbuch hinaus geht und Übersetzer bei der Vorbereitung auf medizinische Fachtexte unterstützt. Ein weiteres Ziel des Projekts ist die Zusammenstellung eines möglichst umfangreichen Bestandes an fachspezifischen Kollokationen und an morphologischen Daten, die durch die muttersprachliche Introspektion der Studierenden – beziehungsweise durch eine sorgfältige Recherche in der fremden Sprache – validiert wurden.

Nach der organisatorischen und technischen Umsetzung können in einer Testphase die Rückmeldungen von Nutzern zur Verbesserung des Systems genutzt werden. Bei wachsenden Eintragsmengen im Zuge einer Erweiterung durch künftige Abschlussarbeiten kann später eine Konsolidierung des Datenbestands erforderlich werden, damit nicht zu viele Suchergebnisse zu einem bestimmten Term angezeigt werden.

Eine Online-Eingabe der Terminologie durch die Studierenden bietet den Vorteil, dass für Datenerzeugung und Datenabfrage dieselbe Plattform genutzt werden kann. Sie erfordert

aber auch weitergehende Sicherheitsmaßnahmen zum Schutz der Datenintegrität. Angesichts der Möglichkeit, qualitativ hochwertige und validierte Termini als Datenquelle zu nutzen, sollte diese Funktion bei einer Weiterentwicklung der Datenbank vorrangig umgesetzt werden.

## 7 Literaturverzeichnis

Benson, Morton, Evelyn Benson und Robert Ilson (1986). *The BBI Combinatory Dictionary of English. A Guide to Word Combinations*. Amsterdam/Philadelphia: John Benjamins.

Chabner, Davi-Ellen (2009). *Medical Terminology. A Short Course*. 5th ed. St. Louis: Saunders Elsevier.

Hausmann, Franz Josef (1989). „Le dictionnaire de collocations.“ Franz Josef Hausmann et al. (Hgg.). *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Handbücher zur Sprach- und Kommunikationswissenschaft 5.1. Berlin/New York: Walter de Gruyter. 1010-1019.

Hüging, Anna-Katharina (2011). *Übersetzerisches Handeln im Kontext der medizinischen Fachkommunikation*. Heidelberger Studien zur Übersetzungswissenschaft 15. Trier: Wissenschaftlicher Verlag Trier.

Keller, Nicole (2006). *Neue Wege in der Hilfsmittelkunde der Übersetzungswissenschaft: Zur Herleitung webbasierter Terminologiedatenbanken im Kontext von CAT-Systemen*. Heidelberger Studien zur Übersetzungswissenschaft 5. Trier: Wissenschaftlicher Verlag Trier.

Kimmes, Anne und Joachim Kornelius (2011). „Übersetzen V / Methodisch-thematische Hinführung zur Bachelor-Arbeit: Ressourcen im Bereich der übersetzungsbezogenen Terminologiearbeit“. Sommersemester 2011, Universität Heidelberg, Seminar für Übersetzen und Dolmetschen. <http://elearning.uni-heidelberg.de/course/view.php?id=4414>. [21.07.2011]

*Lighthouse Unlimited*. <http://www.lighthouse-unlimited.de/>.



Orbán, Wencke (2008). *Über die Entlehnung konstruktivistischer Lerntheorien in die Praxis der Übersetzungswissenschaft – Kooperatives Übersetzen als Kommunikations- und prozessorientierte Handlungsform des Übersetzens*. Heidelberger Studien zur Übersetzungswissenschaft 10. Trier: Wissenschaftlicher Verlag Trier.

TMT – *Teaching Medical Translation*. Ein Forschungsprojekt der Englischen Abteilung des Seminars für Übersetzen und Dolmetschen der Ruprecht-Karls-Universität Heidelberg. <http://www.tmt-heidelberg.com/>.

U.S. National Library of Medicine. *Outline of the NLM Classification*.  
<http://www.nlm.nih.gov/class//OutlineofNLMClassificationSchedule.html> [21.04.2011]

W<sup>3</sup>Techs (2011). *Usage of server-side programming languages for websites*.  
[http://w3techs.com/technologies/overview/programming\\_language/all](http://w3techs.com/technologies/overview/programming_language/all). [04.08.2011]

W<sup>3</sup>Techs (2011). *Usage of web servers for websites*.  
[http://w3techs.com/technologies/overview/web\\_server/all](http://w3techs.com/technologies/overview/web_server/all). [04.08.2011]

Wikimedia Commons, [http://commons.wikimedia.org/wiki/Main\\_Page](http://commons.wikimedia.org/wiki/Main_Page).

## 8 Abbildungsverzeichnis

Abbildung 1: Die Eintragsstruktur der Datenbank auf Grundlage von N. Keller (2006: 66).....	4
Abbildung 2: Auswahl des Fachgebiets in MS-Excel .....	6
Abbildung 3: Auswahl einer lokal gespeicherten Grafik .....	8
Abbildung 4: Eingabe der morphologischen Daten zum Term <i>anemia</i> .....	9
Abbildung 5: Kollokationen zum Term <i>hemoglobin</i> in der Datenbank.....	11
Abbildung 6: Verteilung der Suchergebnisse auf eine Seite pro Eintrag .....	12
Abbildung 7: Ein terminologischer Eintrag auf der Website.....	17

## 9 Tabellenverzeichnis

Tabelle 1: Die erweiterte Kollokationstypologie auf der Basis von Hausmann .....	6
Tabelle 2: Die Klammern zur Notation der Morphemtypen nach Chabner.....	9
Tabelle 3: Beispiel einer morphologischen Notation .....	9

## 10 Anhang

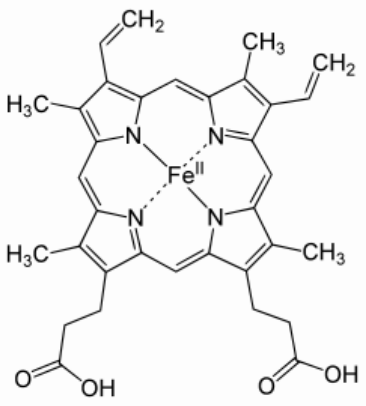
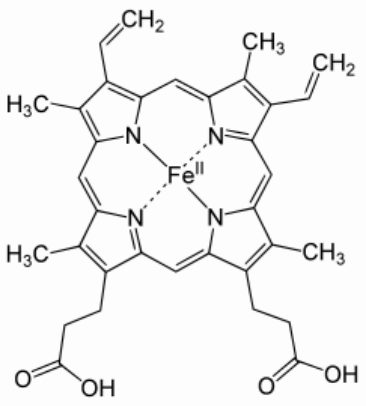
Term	Hämoglobin	hemoglobin
Synonym	(roter) Blutfarbstoff	Blood pigment; Haemoglobin (GB)
Plural	kein Plural	no plural
Fachgebiet	Blut- und lymphatisches System	Hemic and Lymphatic Systems
Definition	[...]	[...]
Quelle Definition	<a href="http://www.rheumatoide-arthritis.de/index.php?csrc=http://www.rheumatoide-arthritis.de/de/glossar/inews/1040301710-32207.html">http://www.rheumatoide-arthritis.de/index.php?csrc=http://www.rheumatoide-arthritis.de/de/glossar/inews/1040301710-32207.html</a>	<a href="http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&amp;va=hemoglobin">http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&amp;va=hemoglobin</a>
Wortklasse	Substantiv	noun
Genus	neutrum	
Kontext	[...]	[...]
Quelle Kontext	<a href="http://www.medsana.ch/artikel.php?id=9992711&amp;box=1">http://www.medsana.ch/artikel.php?id=9992711&amp;box=1</a>	<a href="http://www.everettclinic.com/kbase/topic/medtest/hw39098/descrip.htm">http://www.everettclinic.com/kbase/topic/medtest/hw39098/descrip.htm</a>
Morphologie	Hämo: gr. haima (Blut) glob: lat. globus (Kugel) in: chemische Verbindung	hemo: Greek 'haima' (blood) glob: Latin 'globus' (sphere) in: chemical compound
Grafik		
Quelle Grafik	By Yikrazuul (Own work) [Public domain], via <a href="#">Wikimedia Commons</a>	By Yikrazuul (Own work) [Public domain], via <a href="#">Wikimedia Commons</a>
Eintragsdatum	2012-11-20	2001-10-20
Autor	Johannes Akkach	
Kollokationen	~ bilden, enthalten V + S	(to) convert ~ V + S
	embryonale ~e, adulte ~e, glykiertes ~ Adj + S	adult ~, fetal ~ Adj + S
	binden (an) ~ S + V	~ + transport, + consist, + be restricted, + bind, + combine S + V
	Konzentration von ~ S (+ Präp) + S	~'s affinity, ~'s ability, ~ molecule S (+ Präp) + S

Abbildung 7: Ein terminologischer Eintrag auf der Website<sup>2</sup>

<sup>2</sup> Grafik links und rechts: [http://commons.wikimedia.org/wiki/File%3AHeme\\_b.svg](http://commons.wikimedia.org/wiki/File%3AHeme_b.svg), By Yikrazuul (Own work) [Public domain], via Wikimedia Commons; Zugriff jeweils am 14.09.2011

## **T21N - Translation in Transition**

T21N offers a cutting-edge electronic publishing venue, created by experts for both young talent and established researchers from the worlds of translation and interpreting.

T21N provides a stage for emerging ideas and new academic talent to present their ideas in a digital reading site, where speed and ease meet enjoyment.

T21N is exclusively published online at <http://www.t21n.com>.

Articles in compliance with our style sheet may be submitted at any time and will be published at short notice.

T21N editors research and teach at the Institute of Translation and Interpreting at the University of Heidelberg in Germany.

Editors:

Dipl.-Übers. Viktorija Bilić, Dr. Anja Holderbaum,  
Dr. Anne Kimmes, Prof. Dr. Joachim Kornelius,  
Dr. John Stewart, Dr. Christoph Stoll