

STEFAN FALKE/CHRISTIAN LANG

LANIAKEA: ENTWICKLUNG EINES GRAPHENTHEORETISCHEN ANALYSETOOLS ZUR ÜBERARBEITUNG DER TERMINOLOGIE-KOMPONENTE IN GRAMMIS

Abstract: Gegenstand dieses Beitrags ist die Entwicklung des graphentheoretischen Analysetools *Laniakea*, das zur Visualisierung von Phänomenen und Veränderungen in terminologischen Netzwerken entwickelt wurde. Wir führen theoretische Grundlagen, Designentscheidungen und technische Details der Implementierung des Tools aus. Darüber hinaus wird auch eine Beschreibung von Erfahrungen im Fokus des Beitrages stehen, die bei der Anwendung von *Laniakea* bei der Überarbeitung der terminologischen Ressourcen des Grammatischen Informationssystems *grammis*, gesammelt wurden.

Keywords: Terminologie, Visualisierung, Graphentheoretische Analyse, Deutsche Grammatik

1. Einleitung

Visualisierungen, also die graphische Repräsentation von Daten und Datenstrukturen, sind mit dem Aufkommen immer leistungsfähigerer Computer auch für große Datenmengen möglich geworden und können sogar interaktiv gestaltet werden. Nach Einschätzung von Ware (2012) können interaktive Visualisierungen als Schnittstelle zwischen umfassenden Informationsressourcen einerseits und dem visuellen System des Menschen andererseits dienen und so einen kognitiv adäquaten Zugang zu komplexen Daten bieten, immer vorausgesetzt, die Daten werden entsprechend gut aufbereitet (ebd., S. 2). Überdies, so Ware weiter, lassen Visualisierungen auch etwaige Probleme innerhalb der Daten offensichtlich werden (ebd., S. 3).

Vor diesem Hintergrund beschreiben wir in diesem Beitrag die Entwicklung des graphentheoretischen Analysetools *Laniakea*, das am Leibniz-Institut für Deutsche Sprache, Mannheim (IDS) entwickelt wurde. Das Tool visualisiert sowohl Begriffsstrukturen als auch Strukturen, die durch Verlinkungen zwischen Hypertexten entstehen. Es wurde entwickelt, um die Autor/innen bei der inhaltlichen Überarbeitung und Weiterentwicklung der *Wissenschaftlichen Terminologie*, der Terminologiekomponente von *grammis*,¹ zu unterstützen. *grammis* ist ein Online-Informationssystem zur deutschen Grammatik und bündelt auf über 3.000 Webseiten Informationen und Ressourcen zu unterschiedlichen Aspekten dieses Themas (Schneider/Schwinn 2014).

¹ <http://grammis.ids-mannheim.de>, siehe auch Schneider/Schwinn (2014).

Der Verlauf dieses Beitrages gliedert sich wie folgt: Abschnitt 2 gibt einen Überblick über den Einsatz von Visualisierungen im Bereich Terminologien und Wortnetze und stellt graphentheoretische Grundlagen dar. In Abschnitt 3 folgt eine kurze Darstellung der *Wissenschaftlichen Terminologie* mit einem Fokus auf den beiden unterschiedlichen visualisierbaren Netzstrukturen der Ressource. Abschnitt 4 beschreibt dann die Umsetzung von *Laniakea*. In Abschnitt 5 folgen Erfahrungen, die im Laufe der Anwendung von *Laniakea* bei der Überarbeitung der Terminologiekomponente von *grammis* gesammelt wurden, bevor in Abschnitt 6 das Fazit und ein Ausblick auf weitere Entwicklungen den Beitrag abschließen.

2. Theoretischer Hintergrund

2.1 Einsatz von Visualisierungen

Laniakea wird bei der inhaltlichen Überarbeitung der Terminologiekomponente von *grammis* eingesetzt. Eine *Terminologie* ist definiert als der „Gesamtbestand der **Begriffe** [...] und ihrer **Bezeichnungen** [...] in einem Fachgebiet“ (DIN 2342: 2011-08, S. 16). Diese Begriffe können in einem Begriffssystem zueinander in Beziehung gesetzt werden.² Der Einsatz von Visualisierungstechniken zur Darstellung von Begriffssystemen ist – zumindest in der Theorie – bedeutender Bestandteil von Terminologearbeit. Entsprechend finden Visualisierungstechniken Anwendung für die Darstellung des terminologischen Inventars unterschiedlicher Fachrichtungen, bspw. Rechtswissenschaften (Culy/Chiocchetti/Ralli 2013), Anatomie (Momota/Ohtsuka 2018) und Verhaltensanalyse (Critchfield 2017). Darüber hinaus finden sich auch einige Tools für Terminologiemanagement mit Visualisierungsfunktion³ wie bspw. Coreon⁴ oder auch Termweb.⁵ Auch für die Überarbeitung der terminologischen Komponente von *grammis* wurde bereits ein Visualisierungstool entwickelt und eingesetzt (Suchowolec/Lang/Schneider 2018), dieses ist allerdings nur auf Ebene der Begriffsstruktur einsetzbar (vgl. Abschnitt 3).

Visualisierungen stehen darüber hinaus in der Darstellung von Lexika und Wortnetzen wie dem WordNet (Miller 1995) im Fokus der Forschung. Bei der Visualisierung von Wortnetzen geht es einerseits um die Navigierbarkeit innerhalb eines Wortnetzes wie bei der Entwicklung des WordNet-Explorers

² DIN 2342:2011-08 definiert Begriffssystem als „Menge von **Begriffen** [...] eines **Begriffsfeldes** [...], die entsprechend den **Begriffsbeziehungen** [...] geordnet sind.“ (S. 7).

³ Siehe aber Drewer/Massion/Pulitano, die auf einen Mangel an „Tools für die [...] grafische Visualisierung“ (Drewer/Massion/Pulitano, 2017, S. 21) hinweisen.

⁴ <https://coreon.com>.

⁵ www.interverbumtech.de.

(Collins 2006). Andererseits sollen große Datenmengen wie das WordNet dadurch zugänglicher gemacht werden (Caldarola/Picariello/Rinaldi 2015).

Generell ist eine tabellarische Auflistung im Fall komplexer Datenstrukturen, in denen die Begriffe untereinander über mehrere Relationen verbunden sein können, schnell unübersichtlich. Dies gilt umso mehr für polyhierarchische Strukturen, also Strukturen, in denen Knoten über mehrere Mutterknoten verfügen können. Eine Visualisierung als Graph zeigt gegenüber einer tabellarischen Darstellung deutliche Vorteile, was die Lesbarkeit, Verständlichkeit und den Zugang zu diesen Daten angeht (Katifori et al. 2007; Keim et al. (Hg.) 2010). Voraussetzung dafür ist allerdings, dass die Visualisierung bestimmte Kriterien erfüllt.

Shneiderman hat entsprechende Forderungen formuliert, bei deren Einhaltung die Informationsvisualisierung auch großer Datenmengen für den Anwender von Nutzen ist. Diese Forderungen lauten „Overview first, zoom and filter, then details-on-demand“ (Shneiderman 2003, S. 337). *Overview* (‘Überblick’) meint die Möglichkeit, sich einen Gesamtüberblick über alle Datenpunkte zu verschaffen. Diese sollen zunächst in einer Ansicht komplett dargestellt werden. Mittels einer *Zoom*-Funktion kann sich der Nutzer einzelne Datenpunkte näher in den Fokus holen. Über *Filter* soll es möglich sein, die dargestellten Informationen einzuschränken und nur Daten anzuzeigen, die gewisse Eigenschaften erfüllen. *Details-on-demand* (‘Details bei Bedarf’) bedeutet, detailliertere Informationen zu bestimmten Datenpunkten einblenden zu können, wenn diese erwünscht sind, dies kann bspw. mittels Tooltips, also kleiner Fenster, die beim Überfahren mit der Maus erscheinen, realisiert werden.

Da Terminologien deutliche Ähnlichkeiten zu Computernetzwerken aufweisen (Garrido/Gutierrez 2016; Polguère 2014), können sie entsprechend formal dargestellt werden. Die einzelnen Begriffe werden als Knoten angesehen, die Beziehungen dieser Begriffe zueinander werden als Kanten dargestellt (Hypo-/Hyperonyme etc.). Sie können daher als Graph visualisiert werden.

Graphen sind definiert als eine Menge aus Knoten, die durch Kanten miteinander verbunden sind, und werden beschrieben mit $G = (V, E)$ (Newman 2003). Wobei $V = [v_1, v_2, v_3, \dots, v_n]$ für eine endliche Anzahl an Knoten (engl. vertices) und $E = [e_1, e_2, e_3, \dots, e_n]$ für eine endliche Anzahl an Kanten (engl. edges) steht. Die Kanten können dabei eine Richtung vorgeben, bspw. $e_n \rightarrow e_{n+1}$, d. h., Knoten e_n ist der Vorgänger von Knoten e_{n+1} . Analog wird Knoten e_{n+1} Nachfolger von Knoten e_n genannt. Da die Kanten in diesem Fall gerichtet sind, spricht man auch von einem gerichteten Graphen (siehe Abbildung 1a) (Struckmann/Wätjen 2016). Gerichtete Graphen werden verwendet, um die Richtung einer Beziehung, wie z. B. bei Kommunikationsrichtungen,

darzustellen. Ungerichtete Graphen zeichnen sich dadurch aus, dass die Kanten keine Richtung vorgeben, sondern eine beidseitige Beziehung zwischen zwei Knoten darstellen (siehe Abbildung 1b).

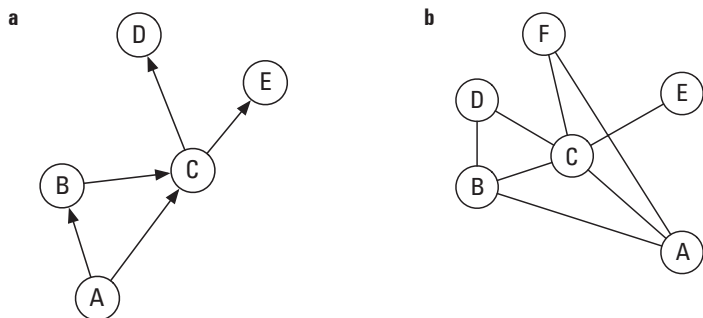


Abb. 1: Graphenstrukturen (eigene Darstellung)

2.2 Graphenanalyse

Sollen zwei Graphen miteinander verglichen werden, ist dies aufgrund der Visualisierung alleine nicht immer einfach. Einige quantitative Eigenschaften, die den Graphen charakterisieren, können in einer visuellen Darstellung nur schwer erfasst werden. Diese sind aber nötig, um grundlegende strukturelle Charakteristika eines Graphen zu erhalten.

So gibt die *Anzahl der Knoten* Auskunft über die Größe des Graphen. Die *Anzahl der Kanten* hingegen gibt einen Eindruck der Vernetzung des gesamten Graphen und auch einzelner Knoten. Die *Dichte* ist hierbei das Verhältnis der Anzahl aller Kanten, die in diesem Graphen möglich sind, zu deren tatsächlicher Anzahl. Je dichter ein Graph ist, desto geringer ist seine *mittlere Weglänge*, also die durchschnittliche Anzahl an Schritten, die nötig sind, um von einem zufällig gewählten Knoten zu einem anderen zufällig gewählten Knoten zu gelangen (Newman 2003, S. 171). In gerichteten Graphen sind außerdem zwei Arten von Knoten zu unterscheiden: *Hubs* und *Authorities*. Authorities sind Knoten, die zu einem bestimmten Thema Informationen bereitstellen. Hubs hingegen sind Knoten, die auf diese Authorities verweisen und dadurch Authorities zu einem bestimmten Thema bündeln (Kleinberg 1999, S. 606ff.). Um Knoten nach der Anzahl und Wichtigkeit ihrer Verlinkungen einzustufen, stellt der *PageRank*-Algorithmus eine Möglichkeit des Rankings dar (Page et al. 1999). Entscheidend ist hierbei, wie viele eingehende Links auf einen Knoten gerichtet sind und ob diese Links wiederum von Knoten kommen, die ihrerseits selbst über einen hohen PageRank-Wert verfügen. Eine weitere wichtige Kennzahl für Netzwerke ist die *Transitivität* (auch als *Clusterkoeffi-*

ziert bezeichnet). Wenn eine gerichtete Kante von einem Knoten v_n auf einen Nachfolger-Knoten v_{n+1} und von diesem wiederum eine gerichtete Kante auf den Nachfolger v_{n+2} zeigt, also $v_n \rightarrow v_{n+1} \rightarrow v_{n+2}$, dann gibt die Transitivität die Wahrscheinlichkeit an, mit der dann auch eine gerichtete Kante von Knoten v_n zu Knoten v_{n+2} existiert, also $v_n \rightarrow v_{n+2}$ (Watts/Strogatz 1998). Diese Eigenschaften sollten in einer Visualisierung dem Nutzer bei Bedarf zugänglich gemacht werden, um genauere Beurteilungen bestimmter Ausschnitte zu ermöglichen.

3. Die Terminologie-Komponente in *grammis*

3.1 Überblick

Eine der zentralen Funktionen der *Wissenschaftlichen Terminologie* besteht darin, den Benutzerinnen und Benutzern von *grammis* als Nachschlageressource zu dienen. Über die entsprechende Rubrik *Wissenschaftliche Terminologie* kann auf der *grammis*-Webseite ein Wörterbuch angesteuert werden, in dem über eine alphabetisch geordnete Lemmaliste kurze Erläuterungstexte zu Begriffen deutscher Grammatik abgerufen werden können. Darüber hinaus finden sich in den Texten der übrigen *grammis*-Komponenten Verbindungen in das Wörterbuch in der Form von Hyperlinks, die bei der Erstellung des Textes von den Autor/innen auf die entsprechenden Termini gesetzt wurden. Ein Klick auf diese Links öffnet den entsprechenden Erläuterungstext in einem Modalfenster. Dies sind Fenster, die sich über der Webseite öffnen, so dass beim Schließen dieses Fensters die ursprünglich gewählte Seite wieder angezeigt wird.

Die Wörterbuchkomponente ist dabei jedoch nur ein Teil eines komplexeren Begriffssystems. Dieses ist in seiner heutigen Form das Ergebnis einer umfassenden methodisch-infrastrukturellen Überarbeitung, im Zuge derer ehemals separate terminologische Ressourcen in einer gemeinsamen Ressource zusammengeführt wurden. Diese vereint sowohl Begriffsrelationen als auch die Beschreibungstexte der Wörterbuchkomponente in einem System und zwar dergestalt, dass die Erläuterungstexte eines der Attribute sind, die einem Begriff im Begriffssystem zugewiesen werden können (vgl. Suchowolec et al. (2017) für eine umfassendere Darstellung des Überarbeitungsprozesses). Über die Funktion als Nachschlageressource hinaus fungiert die *Wissenschaftliche Terminologie* als Wissensorganisationssystem im Sinne von Mazzocchi (2017) und wird beispielsweise zur Synonymdisambiguierung sowie zur Unterstützung des Suchalgorithmus verwendet (vgl. dazu auch Suchowolec/Lang/Schneider 2018). In Anlehnung an lexikographische Terminologie nennen wir die kodierten Begriffsrelationen *Makrostruktur*, wohingegen wir auf die Struktur der Erläuterungstexte mit dem Terminus *Mikrostruktur* referieren.

Die Erläuterungstexte sind untereinander durch Hyperlinks verknüpft. Diese Verlinkungen werden bei der Texterstellung oder -überarbeitung von den Autoren und Autorinnen händisch vorgenommen. Insgesamt ergeben sich also zwei Vernetzungsstrukturen. Zum einen das Begriffssystem an sich, in dem Relationen zwischen Begriffen abgebildet sind. Zum anderen die Vernetzung der Erläuterungstexte untereinander. Diese beiden Vernetzungsstrukturen sind prinzipiell nicht deckungsgleich, wie Abbildung 2 anhand nicht weiter konkretisierter Begriffe und deren assoziierter Erläuterungstexte illustriert.

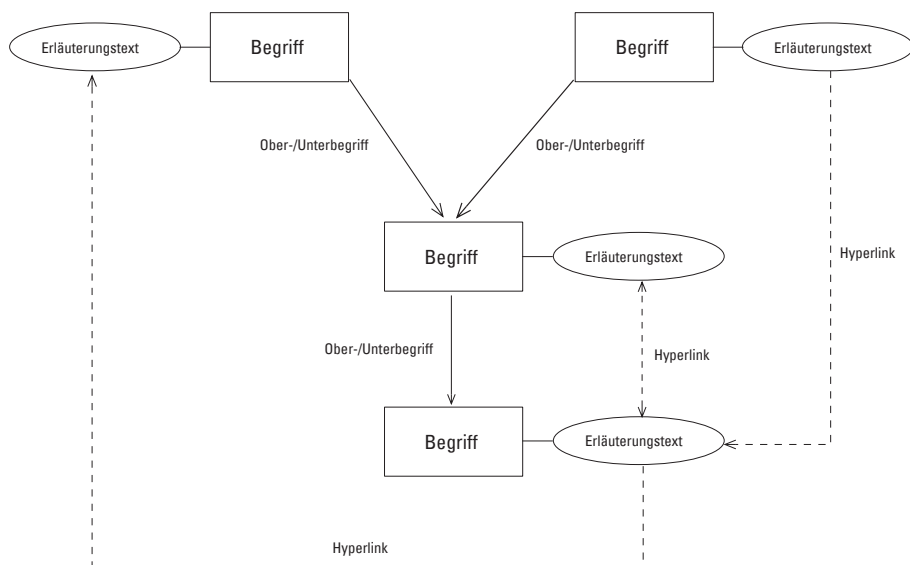


Abb. 2: Illustration von Makro- und Mikrostrukturvernetzungen, gestrichelte Linien symbolisieren Verlinkungen auf Mikrostrukturebene, durchgezogene Linien symbolisieren Begriffsbeziehungen auf Makrostrukturebene

3.2 Die Überarbeitung der Wissenschaftlichen Terminologie – Anforderungen

Die *Wissenschaftliche Terminologie* befindet sich naturgemäß in einem permanenten inhaltlichen Überarbeitungsprozess. Dieser Prozess betrifft sowohl die Erläuterungstexte und deren Vernetzung untereinander als auch das Begriffssystem, also die Makrostruktur.

Der inhaltlichen Revision kommt insbesondere deshalb besondere Wichtigkeit zu, da Erläuterungstexte und Begriffssystem ursprünglich disjunkte und teilweise disparate Ressourcen waren, ehe sie im Zuge der methodisch-infrastrukturellen Überarbeitung zusammengeführt wurden. Die Autor/innen müssen sich also im Revisionsprozess mit Vernetzungen auf sowohl der

mikrostrukturellen als auch der makrostrukturellen Ebene befassen und sollten idealerweise leichten Zugriff auf beide Strukturen haben.

Über eine separate Darstellung von Makro- und Mikrostrukturvernetzung hinaus ist die Möglichkeit einer Gegenüberstellung beider Strukturen wünschenswert, da sich aus der makrostrukturellen Vernetzung Hinweise auf die Verlinkung der Erläuterungstexte gewinnen lassen. Dies gilt umso mehr, da wir bei der Revision der *Wissenschaftlichen Terminologie* theoretischen Ansätzen der Terminologielehre folgen, die vom Primat der Begriffsstruktur (Makrostruktur) ausgehen und einen Begriff stets in seiner struktureller Umgebung betrachten (Suchowolec 2018). Das bedeutet, dass im Überarbeitungsprozess stets ein konzeptionelles Feld und nicht ein einzelner Begriff im Fokus steht. Folglich steht eine Überprüfung bzw. Überarbeitung der Makrostruktur am Anfang des Prozesses, die Erstellung, Bearbeitung und Verlinkung der Erläuterungstexte (Mikrostruktur) orientiert sich dann an der Makrostruktur (vgl. Suchowolec/Lang/Schneider 2018).⁶ Entsprechend den Ausführungen von Drewer/Massion/Pulitano (2017, S. 18) sind hierbei vornehmlich die hierarchischen Beziehungen von Wichtigkeit, weniger die nicht-hierarchischen Assoziativbeziehungen.

Hinsichtlich der makrostrukturellen Vernetzung ist ein visueller Zugang bereits mithilfe eines Visualisierungstools möglich, das einen umfassenden und stark parametrisierbaren visuellen Zugang zur Makrostruktur bereitstellt (für eine ausführlichere Darstellung dieses Tools siehe Lang/Schwinn/Suchowolec 2018; Suchowolec/Lang/Schneider 2018). Ein vergleichbarer Zugang zur mikrostrukturellen Vernetzung war jedoch bislang nicht gegeben und demzufolge auch keine Möglichkeit eines Vergleichs beider Vernetzungsstrukturen. Um einen einfachen Zugang zu Makro-, Mikrostrukturvernetzung und deren Vergleich zu gewährleisten, wurde zur Unterstützung der Autor/innen bei der Überarbeitung von Makro- und Mikrostruktur das graphentheoretische Analysetool *Laniakea* entwickelt. Die Informationsvisualisierung dient dabei mehreren Zielen, unter anderem der Qualitätssicherung und dem schnellen Zugriff auf große Datenmengen (Ware 2012). Zentrale Funktionen sind hierbei der Zugriff auf die mikrostrukturelle Vernetzung und die Möglichkeit, Mikro- und Makrostruktur gegenüber zu stellen sowie den Status quo mit dem Status quo ante zu vergleichen. Darüber hinaus werden einige graphentheoretische Analysen durchgeführt. Graphentheoretische Ansätze finden bereits in funktional ähnlichen Kontexten Anwendung, um Strukturen in komplexen Wissenssystemen aufzudecken (Mehler 2008, 2017; Watts/Strogatz 1998). Ein solches Tool wird unserem Wissen nach erstmals im Kon-

⁶ Dies soll nicht bedeuten, dass eine 1:1-Abbildung der Makro- in der Mikrostrukturvernetzung angestrebt wird, vgl. Abschnitt 5.

text der linguistischen Terminologie eingesetzt. Im Folgenden wird auf Makro- und Mikrostruktur hinsichtlich ihrer technischen Eigenschaften näher eingegangen.

3.3 Makrostruktur

Die Makrostruktur wird durch die Vernetzung der einzelnen Begriffe der terminologischen Ressource gebildet, wie diese in der Datenbank hinterlegt ist (siehe Abbildung 3). Dabei gibt es eine Tabelle, in der alle Begriffe hinterlegt sind (MAOCONCEPTS). Alle Begriffe sind einem sog. Termset (Schneider 2007) zugeordnet (MAOTERMSETS), dabei können mehrere Begriffe zu demselben Termset gehören. In Abbildung 3 ist dies dargestellt durch die 1:n-Beziehung, ein Begriff ist zwingend einem Termset zugewiesen, wohingegen ein Termset keinen bis n Begriffe beinhalten kann. Die ID des Termsets wird dabei in der Tabelle der Begriffe in der Spalte TERMSET_ID abgelegt.

Die Termsets untereinander sind über hierarchische Relationen oder nicht-hierarchische Assoziativbeziehungen miteinander verknüpft in der Form von Abstraktionsziehungen (Oberbegriffe/Unterbegriffe; BT, broader term), in der Form von partitiven Relationen (Holonyme/Meronyme; BTP, broader term partitive), also in einer Teil-Ganzes-Beziehung, und in der Form von nicht-hierarchischen Assoziativbeziehungen (RT, related term). Jede dieser Relationen wird in einer eigenen Tabelle abgebildet. So gibt es eine Tabelle für die Hypo-/Hyperonyme (TERMSETSBTs), eine für die Holo-/Meronyme (TERMSETSBTps) und eine für nicht-hierarchische Assoziativbeziehungen (TERMSETSRTs) (siehe Abbildung 3). So kann ein Termset in der Tabelle der jeweiligen Relation gar nicht oder n -mal vorkommen. Die ID des Termsets wird dabei in der entsprechenden Tabelle in der Spalte MAO_TERMSET_ID abgelegt. Die ID des Termsets, auf das gezeigt wird, wird in der entsprechenden Spalte (z. B. MAO_TERMSETBT_ID) abgelegt. Durch diese Verlinkung ist es möglich, einen gerichteten Graphen zu erstellen.

Die Tabellen mit den aktuellen Daten werden für die Auswertung in eine Analyseumgebung gespiegelt. Einerseits werden hier nur die relevanten Daten übernommen, andererseits werden so die Zugriffe auf die Live-Datenbank gering gehalten.

Um einen Vergleich mit dem Status vor der Überarbeitung zu gewährleisten, wurden zu einem festgelegten Zeitpunkt vor der Überarbeitung historische Daten aus der Datenbank extrahiert und in einer identischen Tabellenstruktur abgelegt. Diese Tabellen erlauben nur noch Lesezugriff, damit die Daten nicht versehentlich überschrieben werden. Die Tabellenstruktur der historischen Daten ist identisch mit der Struktur der Live-Daten, lediglich die Tabellennamen unterscheiden sich in ihrem Präfix, diese haben das Präfix MAO

(MACro Old) vor dem Tabellennamen. Während die Live-Daten das Präfix MAN (MACro New) verwenden.

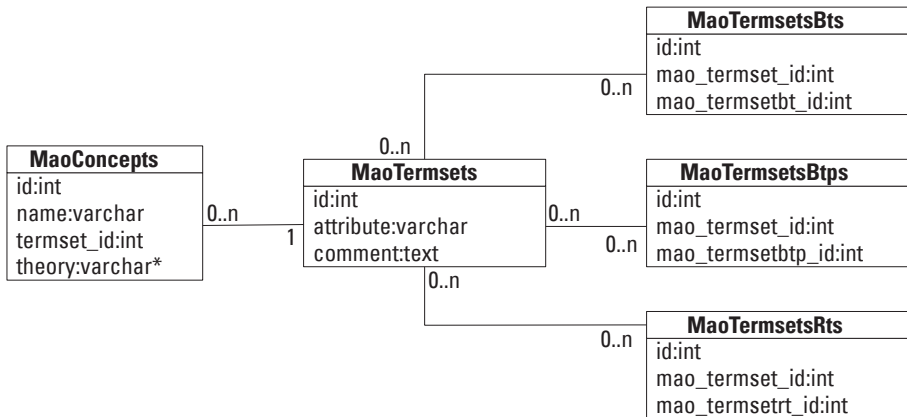


Abb. 3: Relationen der Makrostruktur

3.4 Mikrostruktur

Die mikrostrukturellen Erklärungstexte für jeden Begriff liegen in *grammis* als eigene Webseiten vor. Ihnen liegt ein Datenbankeintrag zugrunde, der diesen Text als XML-Datensatz vorhält. Beim Generieren der Webseite wird dieser mittels XSL-Transformation in HTML transformiert. In diesen Erklärungstexten finden sich Kurzdefinitionen, Erläuterungen, Korpusbelege etc. zu jedem Begriff. Andere Begriffe, auf die in diesen Texten Bezug genommen wird, sind mittels Hyperlinks verknüpft. Da es keine explizite Abbildung der makrostrukturellen Vernetzung in einem mikrostrukturellen Erklärungstext gibt (z. B. eine automatisch aus der Datenbank generierte Linkliste), müssen die Autor/innen beim Erstellen der Erläuterungen diese Vernetzung präsent haben und, falls gewünscht, im Text entsprechend hinterlegen.

Die Verlinkung geht dabei immer in eine Richtung, von dem Text in dem sich der Link befindet, zu dem im Link verwiesenen Text. Eine Verlinkung in die andere Richtung wird dabei nicht automatisch gesetzt, sondern muss bei Bedarf vom Autor vorgenommen werden. So bilden die mikrostrukturellen Texte einen gerichteten Graphen, bei dem jeder Text ein Knoten ist und die Links darin die gerichteten Kanten bilden.

Über einen Datenbank-Crawler können alle Links innerhalb der Erklärungstexte auf andere Datensätze extrahiert werden. Dadurch wird der aktuelle Stand (vor der Überarbeitung) extrahiert und in einer eigenen Tabelle ge-

sichert. Der Crawler speichert lediglich die ID des aktuellen Datensatzes und die in dem zugehörigen Text hinterlegten IDs der Datensätze, die verlinkt sind. Die Crawl-Tabelle (MINEDGES) besteht zunächst aus zwei Spalten SOURCE_CONCEPT_ID und TARGET_CONCEPT_ID. Wobei SOURCE_CONCEPT_ID die ID des Datensatzes enthält, der den Link in sich trägt, und TARGET_CONCEPT_ID das Linkziel darstellt, also der Datensatz ist, der mit dem ersten verknüpft ist. Diese Tabelle wird dann um die Bezeichnungen der verlinkten Begriffe erweitert, indem diese Daten aus der Begriffstabelle extrahiert werden. Die Bezeichnungen der Begriffe werden in den Spalten SOURCE_CONCEPT_NAME und TARGET_CONCEPT_NAME abgelegt.

Aus dieser Tabelle lassen sich sämtliche eindeutigen Begriffe extrahieren und in einer eigenen Tabelle ablegen. Diese ist über eine 1-n-Beziehung mit der Tabelle MINEDGES verknüpft (siehe Abbildung 4). Die in der Tabelle MINEDGES abgelegten Relationen sind gerichtet. Dies soll in der Visualisierung repräsentiert sein, so dass der Autor erkennen kann, ob und in welche Richtung zwei Texte miteinander verknüpft sind.

Für die Betrachtung der aktuellen Daten wird der Crawler jedesmal angestoßen, sobald ein Datensatz geändert wird. Es werden ausschließlich die Verlinkungen der geänderten Datensätze neu extrahiert. Alle anderen Verlinkungen bleiben bestehen.

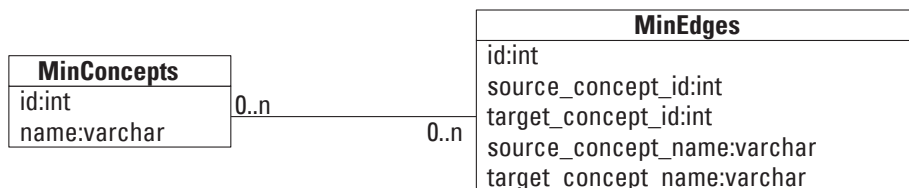


Abb. 4: Relationen der Mikrostruktur

Auch für die Mikrostruktur werden für den Vergleich mit dem aktuellen Überarbeitungsstand die historischen Daten vor der Überarbeitung herangezogen. Diese sind in einer identischen Tabellenstruktur vorgehalten. Die Tabellen sind durch ausschließlichen Lesezugriff vor Veränderungen geschützt und unterscheiden sich lediglich im Tabellenpräfix, das in diesem Fall Mro (Micro Old) lautet.

4. *Laniakea*

Das Autorentool *Laniakea* unterstützt die Autor/innen dabei, während der Bearbeitung eines Begriffes der Terminologie den Überblick über dessen Vernetzung zu behalten. Dazu visualisiert es die Makro- und Mikrostruktur dieses

Begriffes. *Laniakea* ist als eigene Webapplikation basierend auf dem PHP-MVC-Framework CakePHP (Cake Software Foundation, Inc. 2017) konzipiert. Die Visualisierung wird mittels der auf D3 basierenden JavaScript Library Vis.js (Almende B. V. 2018) realisiert.

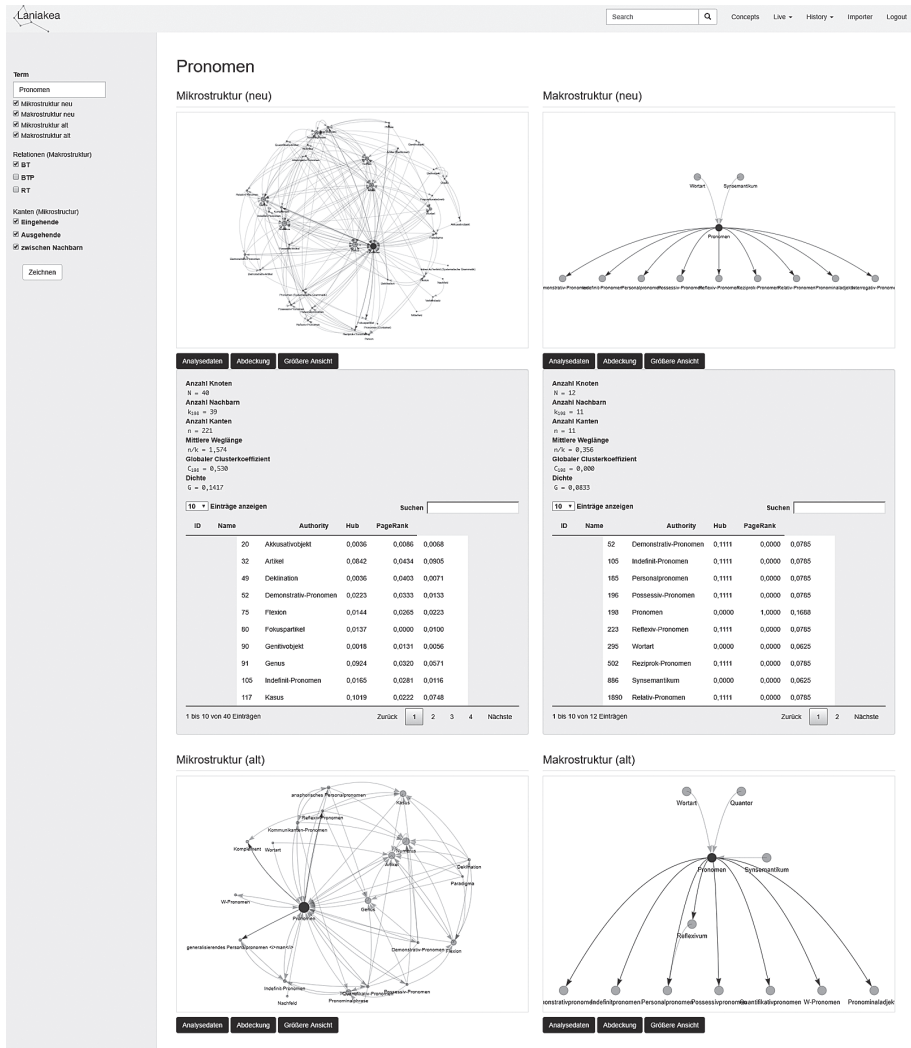


Abb. 5: Übersichtsseite *Laniakea*

In einer 4-Felder-Matrix werden die Makro- und Mikrostruktur zeilenweise nebeneinander gestellt (siehe Abbildung 5). In der oberen Zeile befindet sich der aktuelle Stand der Überarbeitung, in der unteren Zeile der Stand vor der Überarbeitung.

Über ein Suchfeld lassen sich Informationen über einen Begriff recherchieren. Beim Eintippen von mindestens drei Buchstaben öffnet sich eine Vorschlagsliste, aus der ein Vorschlag ausgewählt werden kann. Anschließend wird die Vernetzung dieses Begriffes aus der Datenbank extrahiert und angezeigt. Unterhalb des Suchfeldes befinden sich vier Checkboxes, über diese können die darzustellenden Strukturen gewählt werden.

Mittels dreier Auswahlfelder ist es möglich, die Darstellung der Relationen für die Visualisierung der Makrostruktur zu wählen, also Hypo-/Hyperonyme (BT), Holo-/Meronyme (BTP) oder nicht-hierarchische Assoziativbeziehungen (RT). Letztere sind aufgrund der Übersichtlichkeit standardmäßig ausgeblendet. Dies erfüllt das Kriterium *Filter* (siehe Abschnitt 2.1). Dabei ist jede beliebige Kombination möglich. Für die Visualisierung der Mikrostruktur steht eine Auswahl der zu zeichnenden Kanten zur Verfügung: vom gewählten Hauptknoten eingehende, ausgehende und wahlweise noch die Kanten zwischen den Nachbarn. Über mehrere Checkboxes können in der Mikrostruktur Knoten ein- bzw. ausgeblendet werden, die über ein- bzw. ausgehende Kanten mit dem Hauptknoten verbunden sind.

Die Makrostruktur wird als hierarchischer Baum dargestellt, während die Mikrostruktur als Netz dargestellt wird. In beiden Strukturen ist der zentrale Knoten in rot⁷ dargestellt ebenso wie alle Kanten, die von diesem Knoten ausgehen. Die Nachbarknoten sowie die eingehenden Kanten und diejenigen der Nachbarn untereinander werden in grün⁸ dargestellt. Bei einem Klick auf einen Knoten werden alle von ihm aus- und eingehende Kanten hervorgehoben.

Unter der Visualisierung befinden sich drei Schaltflächen: *Analysedaten*, *Abdeckung* und *Größere Ansicht*. Die Schaltfläche *Analysedaten* öffnet eine Box unterhalb der Visualisierung, die verschiedene Analysedaten bereithält, wie die Anzahl der Knoten und Nachbarn in dieser Darstellung, außerdem die Anzahl an Kanten zwischen den Nachbarn des gewählten Begriffes, die mittlere Weglänge, die Dichte sowie den globalen Clusterkoeffizienten. Darunter befindet sich eine Tabelle über alle Knoten mit den Werten für Authority, Hub und PageRank. Die Tabelle ist durchsuchbar und spaltenweise sortierbar. Diese Daten dienen dazu, einen schnellen Überblick über die wichtigsten Daten des visualisierten Graphen zu liefern.

Die Schaltfläche *Abdeckung* öffnet ebenfalls eine Box unterhalb der Darstellung und zeigt dort diejenigen Begriffe an, die bspw. in der Makrostruktur (neu) vernetzt sind, in der Mikrostruktur (neu) allerdings nicht. So wird auf

⁷ In Abbildung 5 als dunkelgrauer Knoten dargestellt.

⁸ In Abbildung 5 als hellgrauer Knoten dargestellt.

einen Blick ersichtlich, welche Begriffe in der jeweils zugehörigen Struktur vernetzt sind und welche nicht. Die Abdeckung wird nur angezeigt, wenn beide Strukturen eines Standes (alt oder neu) ausgewählt sind. Es ist außerdem möglich, die Visualisierung zu vergrößern (Kriterium Zoom, siehe Abschnitt 2.1). Dies ist auf zwei Arten möglich: Die Schaltfläche *Größere Ansicht* unterhalb der Visualisierung verdoppelt die Höhe und Breite des Visualisierungscontainers. Es besteht aber auch die Möglichkeit, innerhalb der Visualisierung durch Drehen des Mauseisens das Netz zu vergrößern oder zu verkleinern.

Lanikea folgt dem „overview first, zoom and filter, details on demand“-Modell von Shneiderman (Shneiderman 2003, S. 337), d. h. zunächst werden die gewählten Strukturen in ihrer Gesamtheit dargestellt. *Overview* ist in diesem Fall nur als Überblick über die Vernetzung eines einzelnen Begriffs zu sehen. Eine Darstellung aller Begriffe und ihrer Vernetzung würde zu unübersichtlich sein. Über Zoom- und Filter-Funktionen lässt sich die Darstellung auf die Ausschnitte beschränken, die für die jeweilige Fragestellung benötigt werden. Die Zoom-Funktion vergrößert derzeit nur die Darstellung des Netzes. Eine Zoom-Funktion in dem Sinne, dass in das gesamte Netz hinein- und herausgezoomt werden kann, also mehr oder weniger Knoten angezeigt werden, ist derzeit nicht realisiert. Die Filter-Funktionen sind derart umgesetzt, dass sich Nachbarknoten und Kanten ein- und ausblenden lassen.

5. Anwendungserfahrungen

Im Folgenden werden Anwendungsfälle dargestellt, die illustrieren, wie das graphentheoretische Tool *Laniaka* bei der inhaltlichen Überarbeitung der *Wissenschaftlichen Terminologie* als unterstützendes Hilfsmittel eingesetzt werden kann.

Der Umstand, dass die Makrostruktur als Orientierungspunkt bei der Revision bzw. Erstellung von Erläuterungstexten fungiert (vgl. Abschnitt 3.2), ist keinesfalls gleichbedeutend mit dem Bestreben, diese direkt in der Vernetzung der Texte abzubilden. Wenn ein Begriff in der Makrostruktur vor allem aus Modellierungsgründen zum Zweck einer fein granulierten Strukturierung der Daten eingeführt wird, muss es sich dabei nicht zwangsweise um einen linguistischen Terminus handeln, der in einem Erläuterungstext ausgeführt wird. Ein Beispiel dafür ist der in Abbildung 7 auftretende Begriff *Grundelemente des uniformen Modells*: Mithilfe des *uniformen Modells* lassen sich Sätze in sog. *topologische Felder* einteilen: *Vorfeld*, *Mittelfeld* und *Nachfeld*. Diese sind auch als Grundelemente des Modells bekannt. Darüber hinaus existieren (je nach theoretischer Prägung) zusätzliche Erweiterungen des Modells wie beispielsweise das *linke Außenfeld* etc. In der Makrostruktur wird der Unter-

scheidung zwischen Grundelementen und Erweiterungen des Modells Rechnung getragen. Jedoch schlägt sich die Differenzierung nicht in einem eigenen Erläuterungstext für Grundelemente und Erweiterungen des *Uniformen Modells* nieder. Vielmehr wird im Erläuterungstext bspw. des *Vorfeldes* darauf hingewiesen, dass es sich dabei um ein Grundelement des *uniformen Modells* handelt, während im Erläuterungstext zum *linken Außenfeld* darauf hingewiesen wird, dass es sich um eine Erweiterung des *uniformen Modells* handelt. In beiden Fällen erfolgt eine Verlinkung zum Erläuterungstext des *uniformen Modells*. Außerdem: Umgekehrt existieren mikrostrukturelle Relationen, die auf makrostruktureller Ebene nicht erwünscht sind. Dies gilt vornehmlich für Relationen zwischen Kohyponymen, die auf makrostruktureller Ebene redundant (und bspw. nach SKOS-Kriterien ein Qualitätsproblem darstellen),⁹ auf mikrostruktureller Ebene hingegen aus Gründen der Verdeutlichung bisweilen erwünscht sind. Ein Beispiel für einen solchen Fall ist das Verhältnis von *linker Satzklammer* und *rechter Satzklammer*. Während in letztgenannter ausschließlich verbale Elemente (inkl. *Verbpartikeln*) auftreten können, gilt dies für die erstgenannte nicht. Dieser Umstand ist so bedeutsam, dass in den entsprechenden Erläuterungstexten darauf hingewiesen wird und die Erläuterungstexte untereinander verlinkt sind. Das heißt: Mikrostrukturell sind die beiden Erläuterungstexte miteinander vernetzt. Auf makrostruktureller Ebene hingegen sind beide Begriffe als Unterbegriffe desselben Mutterknotens, nämlich *Satzklammer*, modelliert. Eine direkte Relation zwischen *linker Satzklammer* und *rechter Satzklammer* als Kohyponyme ist aus oben genannten Gründen also nicht erwünscht.

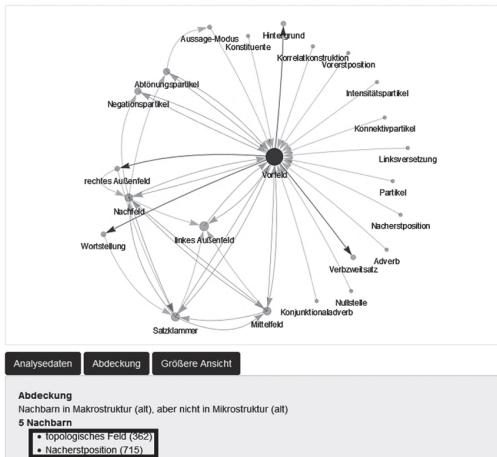
Auch wenn es nicht unser Anspruch ist, dass jede (hierarchische) Relation der Makrostruktur ihren Niederschlag in der Vernetzung der Texte findet, erfüllt die Makrostruktur doch eine wertvolle Orientierungsfunktion für die Überarbeitung und Verlinkung der Erläuterungstexte untereinander wie das folgende Beispiel zeigt.

5.1 Vollständigkeit mikrostruktureller (und makrostruktureller) Vernetzung

Im Rahmen des Handbuchs der deutschen Konnektoren (Pasch et al. 2003) werden die Begriffe *Vorerstposition* und *Nacherstposition* als Positionen im Inneren des *Vorfeldes* beschrieben, also als Meronyme des Holonyms *Vorfeld*. Eine Betrachtung der Makrostruktur vor der inhaltlichen Überarbeitung (siehe das rechte Panel in Abbildung 6) zeigt, dass von diesen beiden nur eines, nämlich die *Nacherstposition*, makrostrukturell repräsentiert war.

⁹ Siehe bspw. https://github.com/cmader/qSKOS/wiki/Quality-Issues#Valueless_Associative_Relations (Stand: 13.6.2019).

Mikrostruktur (alt)



Makrostruktur (alt)

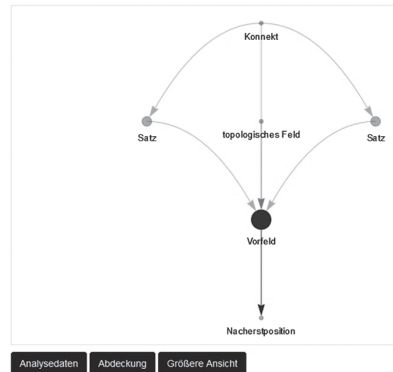


Abb. 6: Gegenüberstellung Makrostruktur und Mikrostrukturvernetzung vor der Revision für den Begriff *Vorfeld*. Basis für die Makrostruktur und damit den Abgleich sind nur die hierarchischen Relationen. Die Mikrostruktur enthält ausgehende und eingehende Links sowie Verlinkungen der Nachbarnoten untereinander¹⁰

Das linke Panel in Abbildung 6 zeigt unter dem Rubrum *Abdeckung*, welche der unmittelbar mit *Vorfeld* verbundenen Knoten auf Makrostrukturebene keine Entsprechung in der mikrostrukturellen Vernetzung haben (siehe die Hervorhebung durch das Rechteck). Es zeigt sich, dass vor der Überarbeitung keine Hyperlinks im Eintrag zu *Vorfeld* zu einem entsprechenden Eintrag zum Begriff *Nacherposition* leiteten, obwohl ein solcher Wörterbucheintrag im System vorhanden war.¹¹ Das bedeutet, dass im Hinblick auf das Verhältnis von *Vorfeld* und *Nacherposition* die Vernetzung in der Makrostruktur keine Entsprechung in der Mikrostrukturvernetzung hatte. Abgesehen davon, dass es Fälle gibt, in denen bewusst eine Entscheidung gegen die Erstellung eines Erläuterungstextes zu einem Begriff (und damit automatisch gegen eine mi-

¹⁰ Ein weiteres Problem, das im Laufe der Überarbeitung behoben werden konnte (vgl. Abbildung 7), betrifft ausschließlich die makrostrukturelle Vernetzung. *Konnekt* ist ebenso wie *Satz* als direkter Oberbegriff von *Vorfeld* modelliert. Gleichzeitig ist *Vorfeld* als Unterbegriff von *Satz* modelliert. Widersprüche dieser Art werden durch eine grafische Aufbereitung auf den ersten Blick erkennbar. Für eine detailliertere Diskussion derartiger Fälle siehe Suchowolec et al. (2017).

¹¹ Auf die fehlende Verlinkung zum Begriff *topologisches Feld* wird im weiteren Verlauf nicht eingegangen, da zu diesem vor der Revision kein Erläuterungstext im System vorhanden war.

strukturelle Vernetzung mit diesem Begriff) getroffen wird, können die Gründe dafür, dass Erläuterungstexte, die Begriffen zugewiesen sind, die in der Begriffsstruktur als unmittelbare Nachbarn in einer hierarchischen Beziehung zueinander stehen, nicht miteinander verlinkt sind, vielfältig sein. Da zur Entstehungszeit der meisten Erläuterungstexte Mikro- und Makrostruktur disjunkte Ressourcen waren und es kein dezidiertes Autorenteam für die Terminologie gab, d. h. verschiedene Texte von verschiedenen Autor/innen verfasst wurden, ist es möglich, dass beim Verfassen des Textes die benachbarten Begriffe schlicht nicht berücksichtigt wurden. Dies ist vor allem dann möglich, wenn es sich um Begriffe einer bestimmten theoretischen Richtung handelt. Überdies kann es vorkommen, dass der benachbarte Begriff im Text zwar erwähnt wird, letztlich aber vergessen wurde, einen entsprechenden Link zu setzen. Für Nutzende unseres Systems wird dadurch im besten Fall die Navigation zwischen verschiedenen Erläuterungstexten erschwert, im schlechtesten Fall entgeht ihnen jedoch Information.

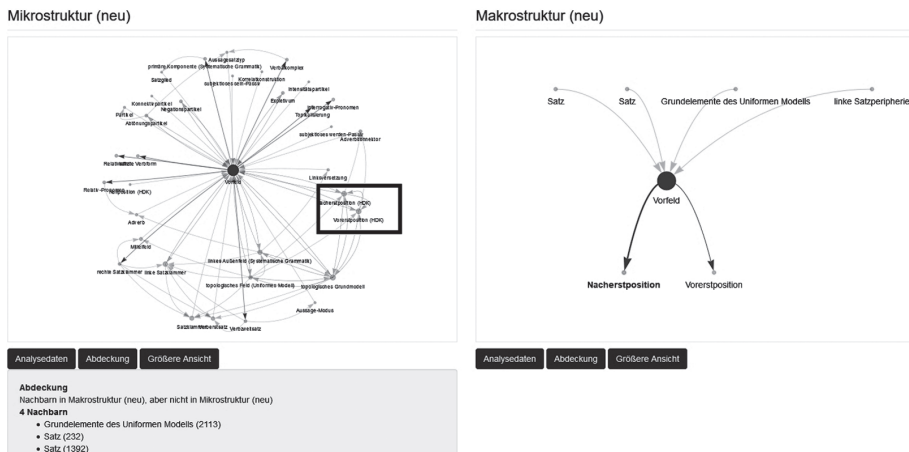


Abb. 7: Gegenüberstellung Makrostruktur und Mikrostrukturvernetzung nach der Revision für den Begriff *Vorfeld*. Basis für die Makrostruktur und damit den Abgleich sind nur die hierarchischen Relationen. Die Mikrostruktur enthält ausgehende und eingehende Links sowie Verlinkungen der Nachbarknoten untereinander

Durch die direkte Gegenüberstellung von Makro- und Mikrostrukturvernetzung in Graphenform, ergänzt durch eine tabellarische Aufstellung derjenigen Begriffe, die in der jeweils anderen Struktur nicht abgedeckt sind, gestaltet sich das Auffinden fehlender Links mithilfe der Software schneller und unkomplizierter, als dies der Fall wäre, müssten die Links jedes Textes gesammelt und dann mit der Makrostruktur abgeglichen werden. Dies er-

leichtert die Arbeit der Autor/innen, auch wenn die Entscheidung darüber, welche Begriffe und Relationen auf Mikrostrukturebene aufgenommen werden, letztlich von den Autor/innen getroffen werden muss. *Laniakea* liefert in dieser Hinsicht eine übersichtliche und leicht zugängliche Entscheidungsgrundlage, auf deren Basis im konkreten Fall im Laufe des Überarbeitungsprozesses z. B. das Problem der fehlenden Verlinkung zwischen *Vorfeld* und *Nacherstposition* erkannt und behoben wurden (siehe die Markierung im linken Panel von Abbildung 7). Darüber hinaus wurde die fehlende Relation zwischen *Vorfeld* und *Vorerstposition* makrostrukturell ergänzt und mikrostrukturell abgebildet.

5.2 Navigierbarkeit

Neben der Vollständigkeit der mikrostrukturellen Vernetzung in dem Sinne, dass zentrale Relationen der Makrostruktur mikrostrukturell in Form von Hyperlinks abgebildet werden, ist die Navigierbarkeit zwischen den Erläuterungstexten der Mikrostruktur ein weiteres Anliegen der Autor/innen im Revisionsprozess.

Jeder Erläuterungstext enthält als verbindliches XML-Fragment eine Kurzdefinition des jeweils beschriebenen Begriffs. Diese Definitionen sind klassische intensionale Definitionen, die den jeweiligen Oberbegriff des Begriffs und konstitutive Charakteristika nennen, die den Begriff von jeweils anderen Kohyponymen abgrenzen. Darüber hinaus sollen Erläuterungstexte, die einem Begriff mit Unterbegriffen zugewiesen sind, Hyperlinks zu den Texten dieser Unterbegriffe enthalten (so diese denn im System vorhanden sind). Infolgedessen ist im Sinne der Navigierbarkeit für die Benutzer/innen des Systems erwünscht, dass in der mikrostrukturellen Vernetzung zwischen Oberbegriff und Unterbegriff gegenseitige Bezugnahme herrscht, diese also gegenseitig aufeinander verlinken, so dass mit nur einem Klick vom Erläuterungstext des Ober- zu den jeweiligen Unterbegriffen und umgekehrt navigiert werden kann.

Die Darstellung als Graph in *Laniakea* ermöglicht es zu erkennen, wenn zwischen Ober- und Unterbegriff eine unilaterale anstelle der gewünschten bilateralen Verlinkung besteht. Bilaterale Verlinkungen werden über den Umstand hinaus, dass zwei Texte (repräsentiert durch zwei Knoten) durch zwei verschiedene Pfeile im Graphen verbunden sind, zusätzlich farblich abgesetzt. Auf diese Weise werden jeweils nicht vorhandene bilaterale Verlinkungen von Oberbegriff und Unterbegriff für die Gesamtheit der Unterbegriffe eines Begriffes auf den ersten Blick erkennbar und müssen nicht durch die Überprüfung der Einzeltexte erschlossen werden.

Abbildung 8 zeigt, dass vor der inhaltlichen Überarbeitung ein Hyperlink aus dem Erläuterungstext zu *Pronomen* zum Text zu *generalisierendes Personalpronomen* führte, jedoch fand sich kein entsprechender Link in die entgegengesetzte Richtung. Im Sinne der Benutzbarkeit der *Wissenschaftlichen Terminologie* hatte das zur Folge, dass Nutzende nicht direkt aus dem Text zum Begriff *generalisierendes Personalpronomen* heraus zum Text zu *Pronomen* navigieren konnten, sondern stattdessen einen Umweg in Kauf nehmen mussten (über die Suchfunktion oder die Lemmaliste).

Mikrostruktur (alt)

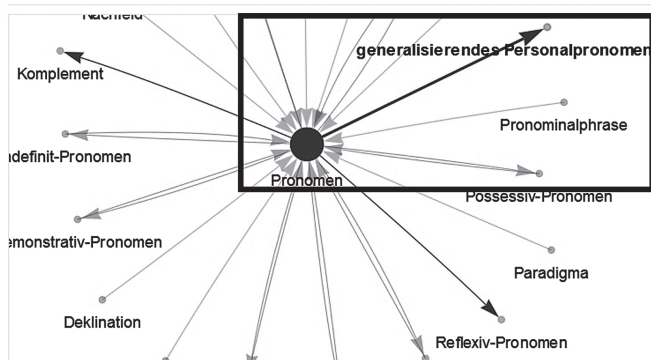


Abb. 8: Unilaterale Verlinkung von Ober- zu Unterbegriff vor der Revision

Über die Tatsache hinaus, dass in diesem Fall zwei Texte nicht bilateral miteinander verlinkt waren, ist dieser Umstand indikativ dafür, dass der Text zu *generalisierendes Personalpronomen* keine intensionale Definition mit Nennung des jeweiligen Oberbegriffs enthielt. Auf diese Weise kann eine unilaterale Verlinkung zwischen Ober- und Unterbegriff ein Hinweis für eine inhaltliche Überarbeitung eines Textes über die Verlinkungen hinaus sein.¹² Abbildung 9 zeigt, dass im Zuge der Revision eine Verlinkung in die entgegengesetzte Richtung hinzugefügt wurde, so dass zwischen Oberbegriff und Unterbegriff nun die gewünschte bilaterale Verlinkung besteht.

¹² Dies ist natürlich nicht zwingend der Fall. Der entsprechende Text kann eine intensionale Definition enthalten, in der einfach vergessen wurde, den entsprechenden Hyperlink zu setzen.

zeigt, dass nach der Überarbeitung der Erläuterungstext zu *semantische Rolle* die gewünschten Funktionen übernimmt: Zum einen fungiert der Text als Verteiler, von dem aus mit nur einem Klick zu den Erläuterungstexten der Unterbegriffe gelangt werden kann. Zum anderen können Nutzende in die entgegengesetzte Richtung direkt aus einem der Texte der Unterbegriffe zum Oberbegriff *semantisches Feld* navigieren.

Zusätzlich zur Darstellung der Vernetzungsstrukturen als Graphen und dem damit verbundenen visuellen Eindruck bietet *Laniakea* unter dem Rubrum *Analysedaten* unterhalb des Graphen einen Überblick über eine Reihe graphentheoretischer Kennzahlen (siehe Abschnitt 2 für eine detailliertere Darstellung derselben). Hinsichtlich der Funktion der einzelnen Knoten im Graphen geben deren Authority-, Hub- und PageRank-Werte weiteren Aufschluss. Überdies werden mit Knoten-, Kanten und Nachbaranzahl sowie mittlerer Weglänge, Globalem Clusterkoeffizient und Dichte Kennzahlen aufgeführt, die den Graphen als Ganzes betreffen.

6. Fazit und Ausblick

Wie die dargestellten Anwendungsfälle illustrieren, kann *Laniakea* eine wertvolle Unterstützung der Autoren bei der Terminologiarbeit leisten, indem auf makrostruktureller Ebene, auf mikrostruktureller Ebene und im Vergleich beider Ebenen mögliche qualitative Probleme sichtbar gemacht werden. Über die bereits implementierten Funktionen hinaus sehen wir noch weiteres Potenzial für eine Fortentwicklung des Tools.

Um das Kriterium *Overview* (siehe Abschnitt 2.1) zu erfüllen, ist wie oben ausgeführt für die nächste Version geplant, eine alternative Methode der Graphenerstellung zu implementieren. Bisher erstellt das Tool einen Graphen basierend auf einem ausgewählten Begriff und bildet dessen direkte Nachbarn sowie auf Mikrostrukturebene deren Verlinkungen untereinander ab. Künftig soll es möglich sein, eine Menge an Begriffen auszuwählen, auf deren Basis *Laniakea* dann einen Graphen erstellt. Auf diese Weise entsteht ein semantisch motivierter Teilgraph, der beispielsweise alle Erläuterungstexte und deren Verlinkungen zum Themenkomplex *Pronomen* enthält.

Eine Darstellung des gesamten Netzes ist wegen der dann entstehenden Unübersichtlichkeit bei großen Datenmengen nicht geplant. Da die Darstellung aller Termini und deren Verlinkung zu unübersichtlich ist, wurde darauf verzichtet. Allerdings lässt sich mit Hilfe zweier Schieberegler (Höhe, Breite) ein

übernimmt der Text zu *semantische Rolle* nicht die gewünschte Verteiler-Funktion in der mikrostrukturellen Vernetzung.

stellen, wie weit nach oben und nach unten Termini dargestellt werden sollen, Oberbegriffe von Oberbegriffen etc. Über die Breite lässt sich einstellen, wie viele Nachbarn von Nachbarn auf gleicher Ebene dargestellt werden sollen. In der Standardeinstellung wird nur jeweils eine Ebene nach oben und nach unten angezeigt und auch nur die direkten Nachbarn.

Das Kriterium *Details-on-Demand* soll ebenfalls erfüllt werden, d. h. beim Überfahren eines Knotens mit der Maus sollen per Tooltip weitere Informationen zu diesem Knoten eingeblendet werden. Ebenso soll das Kriterium *Verhältnis (Relate)* umgesetzt werden, d. h. bei einem Doppelklick auf einen Knoten, der nicht im Fokus steht, wird dieser in den Fokus gestellt und die Visualisierung neu errechnet.

Literatur

- Almende B. V. (2018): vis.js – a dynamic, browser based visualization library. Rotterdam: Almende B. V. (Online: <http://visjs.org>, Stand: 2.10.2018).
- Cake Software Foundation, Inc. (2017): CakePHP: Build fast, grow solid. Las Vegas: Cake Software Foundation, Inc. (Online: <https://cakephp.org>, Stand: 2.5.2017).
- Caldarola, Enrico. G./Picariello, Antonio/Rinaldi, Antonio M. (2015): Big graph-based data visualization experiences: The WordNet case study. In: Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K). Bd. 1. Piscataway: Institute of Electrical and Electronics Engineers (IEEE), S. 104-115.
- Collins, Christopher (2006): WordNet Explorer: Applying visualization principles to lexical semantics. Toronto: Computational Linguistics Group, Department of Computer Science, University of Toronto.
- Critchfield, Thomas S. (2017): Visuwords®: A handy online tool for estimating what nonexperts may think when hearing behavior analysis jargon. In: Behavior Analysis in Practice 10, 3, S. 318-322.
- Culy, Chris/Chiocchetti, Elena/Ralli, Nastacia (2013): Visualizing conceptual relations in legal terminology. In: Proceedings of the 17th International Conference on Information Visualisation (IV), July 16-18, London, UK. Piscataway: Institute of Electrical and Electronics Engineers (IEEE), S. 333-338.
- DIN 2342:2011-08 = Deutsches Institut für Industrienormierung e. V. (Hg.) (2011): Begriffe der Terminologielehre. Berlin: Beuth.
- Drewer, Petra/Massion, François/Pulitano, Donatella (2017): Was haben Wissensmodellierung, Wissensstrukturierung, künstliche Intelligenz und Terminologie miteinander zu tun? Köln: Deutsches Institut für Terminologie.
- Garrido, Camilo/Gutierrez, Claudio (2016): Dictionaries as networks: Identifying the graph structure of Ogden's Basic English. In: Matsumoto, Yuji/Prasad, Rashmi (Hg.): Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka: The COLING 2016 Organizing Committee, S. 3565-3576.

- Katifori, Akrivi/Halatsis, Constantin/Lepouras, George/Vassilakis, Costas/Giannopoulou, Eugenia (2007): Ontology visualization methods – a survey. In: ACM Computing Surveys (CSUR) 39, 4, S. 10.
- Keim, Daniel/Kohlhammer, Jörn/Ellis, Geoffrey/Mansmann, Florian (Hg.) (2010): Mastering the information age: Solving problems with visual analytics. Goslar: Eurographics Association. (Online: <https://vismaster.eu/wp-content/uploads/2010/11/VisMaster-book-lowres.pdf>, Stand 13.5.2019).
- Kleinberg, Jon M. (1999): Authoritative sources in a hyperlinked environment. In: Journal of the ACM (JACM) 46, 5, S. 604-632.
- Lang, Christian/Schwinn, Horst/Suchowolec, Karolina (2018): Grammatische Terminologie am IDS – ein terminologisches Online-Wörterbuch als ein vernetztes Begriffssystem. In: Sprachreport 34, 1, S. 16-26.
- Mazzocchi, Fulvio (2017): Knowledge Organization System (KOS). (Online: www.isko.org/cyclo/kos, Stand: 19.9.2017).
- Mehler, Alexander (2008): Structural similarities of complex networks: A computational model by example of Wiki graphs. In: Applied Artificial Intelligence 22, 7-8, S. 619-683.
- Mehler, Alexander (2017): Zur Struktur und Dynamik von kollaborativ erzeugten lexikalischen Netzwerken. Vortrag gehalten auf der 53. Jahrestagung des Instituts für Deutsche Sprache (IDS), Mannheim, 14.3.-16.3.2017. (Online: www1.ids-mannheim.de/fileadmin/org/tagungen/jt2017/IDS_JT2017_Abstract_24_Mehler.pdf, Stand: 13.6.2017).
- Miller, George A. (1995): WordNet: a lexical database for English. In: Communications of the ACM 38, 11, S. 39-41.
- Momota, Ryusuke/Ohtsuka, Aiji (2018): Network of Anatomical Texts (NAnaTex), an open-source project for visualizing the interaction between anatomical terms. In: Anatomical Science International 93, 1, S. 149-153.
- Newman, Mark E. J. (2003): The structure and function of complex networks. In: SIAM Review 45, 2, S. 167-256.
- Page, Lawrence/Brin, Sergey/Motwani, Rajeev/Winograd, Terry (1999): The PageRank Citation Ranking: Bringing order to the web. Stanford: Stanford InfoLab.
- Pasch, Renate/Brauße, Ursula/Breindl, Eva/Waßner, Ulrich H. (2003): Handbuch der deutschen Konnektoren. (= Schriften des Instituts für Deutsche Sprache 9). Berlin/New York: De Gruyter.
- Polguère, Alain (2014): From writing dictionaries to weaving lexical networks. In: International Journal of Lexicography 27, 4, S. 396-418.
- Schneider, Roman (2007): A database-driven ontology for German grammar. In: Rehm, Georg/Witt, Andreas/Lemnitzer, Lothar (Hg.): Datenstrukturen für linguistische Ressourcen und ihre Anwendungen/Data structures for linguistic resources and applications. Proceedings of the Biennial GLDV Conference 2007. Tübingen: Narr, S. 305-314.
- Schneider, Roman/Schwinn, Horst (2014): Hypertext, Wissensnetz und Datenbank: die Web-Informationssysteme *grammis* und *ProGr@mm*. In: Institut für Deutsche Spra-

- che (Hg.): *Ansichten und Einsichten. 50 Jahre Institut für Deutsche Sprache*. Mannheim: Institut für Deutsche Sprache, S. 337-346.
- Shneiderman, Ben (2003): *The eyes have it: A task by data type taxonomy for information visualizations*. In: Bederson, Benjamin B./Shneiderman, Ben (Hg.): *The craft of information visualization. Readings and reflections*. Amsterdam u. a.: Morgan Kaufmann, S. 364-371.
- Struckmann, Werner/Wätjen, Dietmar (2016): *Mathematik für Informatiker: Grundlagen und Anwendungen*. 2., überarb. u. erw. Aufl. Berlin/Heidelberg: Springer.
- Suchowolec, Karolina (2018): *Sprachlenkung – Aspekte einer übergreifenden Theorie*. (= Sprachwissenschaft 38). Berlin: Frank & Timme.
- Suchowolec, Karolina/Lang, Christian/Schneider, Roman (2018): *An empirically validated, onomasiologically structured, and linguistically motivated online terminology. Re-designing scientific resources on German grammar*. In: *International Journal on Digital Libraries*, S. 1-16.
- Suchowolec, Karolina/Lang, Christian/Schneider, Roman/Schwinn, Horst (2017): *Shifting complexity from text to data model. Adding machine-oriented features to a human-oriented terminology resource*. In: Garcia, Jorge/Bond, Francis/McCrae, John P./Buitelaar, Paul/Chiarcos, Christian/Hellmann, Sebastian (Hg.): *Language, data, and knowledge. First International Conference, LDK 2017. Galway, Ireland, June 19-20, 2017. Proceedings*. (= Lecture Notes in Artificial Intelligence 10318). Cham: Springer, S. 203-212.
- Ware, Colin (2012): *Information visualization: Perception for design*. 3. Aufl. Amsterdam u. a.: Morgan Kaufmann.
- Watts, Duncan J./Strogatz, Steven H. (1998): *Collective dynamics of 'Small-World' networks*. In: *Nature* 393, 6684, S. 440-442.