

Harald H. Zimmermann *)

Der ‚Nietzsche-Thesaurus‘ des Nietzsche-Online-Portals

In: Hammwöhner/Rittberger/Semar (Hrsg., 2004): Wissen in Aktion.
Festschrift für Rainer Kuhlen. Konstanz: UVK, 281-290
Stand: 1.4.2004; Datei FNRK1

Zielsetzung

Die im vorliegenden Beitrag beschriebenen Arbeiten sind Teil des Vorhabens der Entwicklung eines sog. Online-Portals zu Friedrich Nietzsche. In einem früheren Beitrag (Zimmermann 2000) wurde dazu die Frage der Präsentation der Nietzsche-Texte diskutiert <Anm. 1>, ein weiterer Teilaspekt ist die Entwicklung eines Kommunikationssystems für Experten und interessierte Laien (Zimmermann 2003).

Im vorliegenden Beitrag wird ein ‚Thesaurussystem‘ zum Werk von Friedrich Nietzsche vorgestellt. Wesentliche Grundlage sind die Vorarbeiten, die Johann Prossliner in und mit seinem Zitatelexikon („Das Lexikon der Nietzsche-Zitate“, Ausgabe Dezember 2001, im Folgenden kurz ‚Zitatelexikon‘) geleistet hat (Prossliner 2001). Eine zweite Grundlage ist der sog. Wehrle-Eggers, ein nach dem Muster des Roget-Thesaurus gestaltetes deutschsprachiges Wörterbuch nach Sachgruppen (Wehrle-Eggers 1961, letzte Ausgabe).

Im Ergebnis soll ein Verfahren entstehen, das es erlaubt, mit Hilfe des Thesaurus-Systems online via Internet oder Intranet zumindest die Zitatstellen des Zitatelexikons in dessen elektronischer Variante zu recherchieren. <Anm. 2> <Anm.3>

Verfahrensweise

Es sind verschiedene Recherche-Zugänge vorgesehen <Anm. 4>:

- Zugang über die ‚Aufbaustruktur‘ des Zitatelexikons, d.h. über die Gliederung (Kapitel / Unterkapitel ...) des Zitatelexikons
- Zugang über das Registersystem des Zitatelexikons
- Zugang über ein am ‚Deutschen Wortschatz‘ von Wehrle/Eggers orientiertes semantisches Kategoriensystem unter Bezug zu den Stich- und Schlagwörtern im Register des Zitatelexikons.

Hinzu kommen ‚ergänzende‘ Zugänge, etwa über

- zitierte Werke
- erwähnte Personen
- erwähnte Orte
- Zitate zur Nietzsche-Biografie
- Zarathustra-Zitate (aus den Zarathustra-Ausgaben)
- Zitate mit Bezug zu Richard Wagner

Im Folgenden wird das Verfahren konzeptionell beschrieben. Eine Realisierung ist noch für 2004 vorgesehen.

Ausgangsmaterial

Ausgangspunkt sind zunächst die Materialien zum „Lexikon der Nietzsche-Zitate“, die mir vom Kastell-Verlag auf CD-ROM zur Bearbeitung zur Verfügung gestellt wurden:

- Word-Datei des Textteils (knapp 2.500 Zitationen = Text-Extrakte)
- Word-Datei des Registers (ca. 50.000 Schlag- und Stichwörter)

Der Wehrle-Eggers wurde bereits vor einiger Zeit zu Forschungszwecken elektronisch adaptiert und umgesetzt. (Anm. 5)

Zwischenziel ist es, aus diesen Materialien verschiedene sog ‚Relationenwörterbücher‘ (andere Bezeichnung für ‚Thesaurus-Beziehungen‘, im Folgenden auch kurz ‚Wörterbücher‘) zu erstellen:

- Ein Wörterbuch, das die Beziehungen zwischen den „Kapiteln“ des Zitatelexikons darstellt: Kapitel -> Unterkapitel -> ... -> Belegstelle, im Folgenden auch ‚Struktur-Wörterbuch‘ genannt.
- Ein Wörterbuch, das ‚möglichst nahe‘ die Struktur des vorhandenen Index (Register) zum Zitatelexikon (Druck-Ausgabe Dezember 2001) abbildet.
- Ein Wörterbuch, das die hierarchische Struktur des Wehrle-Eggers bis zu den Kategorien (Sachgruppen) darstellt.
- Ein Wörterbuch, das aus dem Abgleich der ‚Stichwörter‘ des Nietzsche-Registers mit den Kategorien (Sachgruppen) des Wehrle-Eggers erzeugt wird.

Kodierarbeiten

Die Aufbereitung des Index (Registers) muss intellektuell erfolgen. In der Druckausgabe (und dementsprechend in der zu bearbeitenden elektronischen Druckvorlage) stehen u.a. Tilden für Wortteile und werden Wortteile durch Abkürzungspunkte ersetzt. Zudem wird das ‚Stichwort häufig als Element des verwiesenen Teils ‚mitgedacht‘ (aber nicht dargestellt).

Dazu einige allgemeine Beispiele (Anm. 6):

- | | | |
|---|------------------------|----------------------------------|
| - | Begabung -> B. / Genie | wird zu: ... -> Begabung / Genie |
| - | begehren -> b.swerter | wird zu: ... -> begehrenswerter |
| - | BEI- -> -bringen | wird zu: bei- -> beibringen |

Aus dem Material des Zitate-Korpus wird eine Beziehung zwischen den Zitatnummern und der Belegstelle (Zarathustra ...) hergestellt. Die Weiterverarbeitung erfolgt analog der Vorgehensweise beim Index.

Beispiel für den Umsetzungs- und Aufbereitungsprozess

(1) „Original“-Registereintrag

Schauspiel

s.a. Theater

der Eitelkeit (Akteure) 950
 ewiges da capo 119
 ewiges, unerbittlich 333
 für Götter (Erdenmensch) 321
 furchtbarstes, hoffnungsreichstes 908
 Moral (Ende in 100 Akten) 0908
 Revolution 91, 784
 S. / Hörspiel (W.) 2232
 Stück / Akteure 950,1937
 Theater / Welt 2187
 Wagner (Niedersinken vor Kreuz) 38

(2) *Halbautomatische Zwischen-Umsetzung*

Schauspiel	\$s.a. Theater
der Eitelkeit	\$<Akteure>0950
ewiges da capo	\$0119
ewiges, unerbittlich	\$0333
für Götter	\$<Erdenmensch>0321
furchtbarstes, hoffnungsreichstes	\$0908
Moral	\$<Ende in 100 Akten>0908
Revolution	\$0091,0784
S. / Hörspiel	\$<W.>2232
Stück / Akteure	\$0950,1937
Theater / Welt	\$2187
Wagner	\$<Niedersinken vor Kreuz>0038

(3) *(Zwischen-)Ergebnis nach intellektueller Aufbereitung*

Schauspiel	\$s.a. Theater
Schauspiel+Schauspiel der Eitelkeit	\$<Akteure>0950
Schauspiel+ewiges da capo	\$0119
Schauspiel+ewiges Schauspiel, unerbittlich	\$0333
Schauspiel+Schauspiel für Götter	\$<Erdenmensch>0321
Schauspiel+furchtbarstes, hoffnungsreichstes Schauspiel	\$0908
Schauspiel+Moral	\$<Ende in 100 Akten>0908
Schauspiel+Revolution	\$0091,0784
Schauspiel+Schauspiel / Hörspiel	\$<W.>2232
Schauspiel+Stück / Akteure	\$0950,1937
Schauspiel+Theater / Welt	\$2187
Schauspiel+Wagner	\$<Niedersinken vor Kreuz>0038

(4) *Hierarchische Relationierung im Theasaurus (ohne die Umkehr-Relationen)*

Schauspiel s.a. Theater

Schauspiel -> Schauspiel der Eitelkeit (Akteure)

Schauspiel -> ewiges da capo
 Schauspiel -> ewiges Schauspiel, unerbittlich
 Schauspiel -> Schauspiel für Götter (Erdenmensch)
 Schauspiel -> furchtbarstes, hoffnungsreichstes Schauspiel
 Schauspiel -> Moral (Ende in 100 Akten)
 Schauspiel -> Revolution
 Schauspiel -> Schauspiel / Hörspiel (Wagner)
 Schauspiel -> Stück / Akteure
 Schauspiel -> Theater / Welt
 Schauspiel -> Wagner (Niedersinken vor Kreuz)

Schauspiel der Eitelkeit (Akteure) -> 0950
 ewiges da capo -> 0119
 ewiges Schauspiel, unerbittlich -> 0333
 Schauspiel für Götter (Erdenmensch) -> 0321
 furchtbarstes, hoffnungsreichstes Schauspiel -> 0908
 Moral (Ende in 100 Akten) -> 0908
 Revolution -> 0091
 Revolution -> 0784
 Schauspiel / Hörspiel (Wagner) -> 2232
 Stück / Akteure -> 0950
 Stück / Akteure -> 1937
 Theater / Welt -> 2187
 Wagner (Niedersinken vor Kreuz) -> 0038

Eitelkeit s.a. Schauspiel der Eitelkeit
 da capo s.a. ewiges da capo
 ewiges Schauspiel s. ewiges Schauspiel, unerbittlich
 furchtbarstes Schauspiel s. furchtbarstes, hoffnungsreichstes Schauspiel
 hoffnungsreichstes Schauspiel s. furchtbarstes, hoffnungsreichstes Schauspiel
 Akteure / Stück s. Stück / Akteure
 Akteure s.a. Stück / Akteure
 Stück s.a. Stück / Akteure

usf.

Konzept der Verbindung der Wehrle-Eggers-Daten mit dem Nietzsche-Material

An zwei Beispielen soll kurz dargestellt werden, wie der Zusammenhang zwischen den Daten des Wehrle-Eggers und des Nietzsche-Zitate-Registers hergestellt wird:

Das Hauptstichwort „Schauspiel“ ist ein Element der Sachgruppe 0599 („Bühnenkunst“) zu der übrigens auch eine Bedeutungsvariante von „Theater“ gehört. Mit Hilfe der Strukturdaten des Wehrle-Eggers lässt sich folgende Kette „zurückverfolgen“:

Schauspiel -> Bühnenkunst (599) -> Geschriebene Sprache -> Verständigung durch Sprache -> Verständigungsmittel -> Kommunikation -> Geistesleben

Der Registereintrag „Eitelkeit“ ergibt in analoger Weise (u.a.) die Kette „Eitelkeit“ -> Geziertheit (853) -> Allgemeingeschmack / Geziertheit -> Ästhetische Gefühle -> Individuelle Gefühle -> Gefühlsleben.

Natürlich wird man dem Nutzer / der Nutzerin nicht unbedingt die Kette (in umgekehrter Reihenfolge) als Zugangsweg anbieten. Es ist eher vorstellbar, dass mit Hilfe dieser Verknüpfung eine Art semantisches Netz über die ‚Nietzsche-Welt‘ gespannt und damit eine Art Browsing ermöglicht wird.

Status und Ausblick

Die an der Gliederung des Zitatelexikons orientierte „hierarchische Struktur“ der Themen des Zitatelexikons ist erstellt. (Anm. 7)

Die Kodierung des Zitate-Registers ist bis einschließlich „T“ erfolgt (März 2004), Teile des Registers sind bereits in die oben beschriebene hierarchische Relationierung überführt. Diese Arbeiten sollen bis Mitte 2004 fertig gestellt sein.

Eine Zuordnung der Zitatstellen zu den Textstellen der Nietzsche-Werke liegt in elektronischer Form vor, so dass im Prinzip auch von den Zitat-Extrakten zu den Volltexten verzweigt werden kann.

Nach Fertigstellung des am Registermaterial orientierten ‚Thesaurus‘ wird die Zuordnung zu den im Wehrle-Eggers verzeichneten Termini erfolgen. Dabei werden ggf. auftretende Lücken im Wortmaterial des Wehrle-Eggers in inhaltlich analoger Form zu schließen versucht. Ggf. werden intellektuell Bedeutungsvarianten, die in der ‚Nietzsche-Welt‘ nicht belegt sind, markiert bzw. eliminiert.

Nach den jetzigen Vorstellungen werden die Daten (Zugangssystem und Zitate) in eine relationale Datenbank übernommen und über eine Internet-Schnittstelle mit den beschriebenen Zugangsformen zugänglich gemacht.

Anmerkungen

- (1) Der etwas provokante Titel „So sprach Zarathustra“ (statt „also“ ...) sollte bereits auf das (dort diskutierte) Problem und die Möglichkeiten der Umsetzung der originalen Nietzsche-Schreibweise in eine moderne Rechtschreibung verweisen.
- (2) Im vorliegenden Zusammenhang soll nicht die grundsätzliche Frage behandelt werden, was unter einem ‚Thesaurus‘ oder – etwa im Sinne der KI-Forschung – unter einer Ontologie verstanden wird. Auf keinen Fall sind die entwickelten Beziehungen als Thesaurus im Sinne von DIN 1463 als hinreichend zu verstehen. Insbesondere ist auf eines hinzuweisen: Bereits die 2.453 ‚Zitate‘ (Extrakte) sind das Resultat einer subjektiven Auswahl (hier des Autors Prossliner), entsprechendes gilt für die Stichwörter des Registers. Gerade bei Nietzsche dürfte es schwer fallen, einen (wissenschaftlichen) Konsens über eine Textselektion und ein begriffliches Zugangssystem zu erreichen. Man muss sich also im vorliegenden Falle stets bewusst sein, dass beide Teilbereiche – auch wenn man dem Autor Prossliner das Bemühen um Objektivität unterstellen kann – Resultat individueller Überlegungen sind. Anders gesagt: Ganz im Sinne von Kunz / Rittel <Lit. 5> sind das Korpus und das Zugangssystem (sicherlich deutlicher, als dies etwa bei einem Thesaurus in naturwissenschaftlichen Bereichen der Fall wäre) Element von ‚jemandes Informationssystem‘.

- (3) Nicht wesentlich anders verhält es sich mit der Einbindung des „Deutschen Wörterbuchs“ von Wehrle-Eggers. Dabei wird allerdings von der Vorstellung ausgegangen, dass sich das vom vorliegenden Anwendungsfall (Nietzsche) losgelöste System gerade bei einem an philosophisch-psychologischen Fragen orientierten Autor gut eignet. Dennoch muss dieser Teil zunächst als ein Experiment angesehen werden. Erst die Praxis wird zeigen, ob diese Überlegungen fruchten.
- (4) Damit kein Missverständnis entsteht: Ein Zugang über die Einzelwörter des Registers soll *daneben* möglich sein, das gleiche gilt für einen Zugang über die Textstichwörter. Damit wird der Thesaurus-Zugang zu einem *Angebot* unter mehreren.
- (5) Grundlage ist die Ausgabe von 1961 (m.W. die letzte Ausgabe dieses Werkes). Es handelt sich um rd. 1000 Sachgruppen zu Grund-Bezeichnungen des gesamten auf sprachliche Benennungen abgebildeten ‚Weltwissens‘, die hierarchisch in 6 Hauptgruppen (Begriffliche Beziehungen, Gefühlsleben, Geistesleben, Raumbereich, Stoff, Wollen) geordnet sind, die wiederum weitere Untergruppen (etwa zu „Gefühlsleben“: Gemeinschaftsgefühle, Gemütsanlage, Individuelle Gefühle, religiöse Gefühle, sittliche Gefühle), Unter-Untergruppen usw. aufweisen, bis man schließlich zu einer der 1000 Sachgruppen gelangt. ...
- (6) Das sehr komplexe Verfahren der Prädizierung und Umsetzung kann hier nicht dargestellt werden. Um eine numerische Sortierung sicherzustellen, mussten beispielsweise auch in der Index-Umsetzung die Verweise auf die Zitatnummern im Textteil mit führenden Nullen versehen werden. Aus Zeitgründen und mangels tieferer Sachkenntnis kann es bei den Ergänzungen zu Interpretationsproblemen kommen. Eine Überprüfung (durch ‚Experten‘) ist daher nötig, wird aber zurückgestellt.
- (7) Beispielkette: Die seelisch-moralische Welt -> Tugend und Laster -> „Laster“ -> Egozentrik -> Eitelkeit -> 0952; 0953, 0954; 0955; 0956; 0957;
Beispiele für Zitatextrakte:
0952: „jeder Mann <hat> gerade so viel Eitelkeit <>, als ihm an Verstand fehlt.“ <HIS 6 (1,293)>
0953: „Wir begegnen oft einem Wohlwollen, welches uns unerklärlich ist; verstehen wir es aber, so beleidigt es uns, weil es zeigt, dass man uns nicht ernst, nicht wichtig nimmt.“ <MEN 1,337>

Literatur

- Eggers, Hans; Wehrle, Hugo (1961): Deutscher Wortschatz. 12. Aufl., Stuttgart
- Kunz, Werner; Rittel, Horst (1972): Die Informationswissenschaften: Ihre Ansätze, Probleme, Methoden und ihr Ausbau in der Bundesrepublik Deutschland. München: Wien. Auch online: URN: urn:nbn:de:bsz:291-scidok-314; URL: <http://scidok.sulb.uni-saarland.de/volltexte/2000/31/>
- Prossliner, Johann (2001): Licht wird alles, was ich fasse. Das Lexikon der Nietzsche-Zitate. München
- Zimmermann, Harald H. (2000): 'So sprach Zarathustra' oder: Über die Frage der Normierung von Nietzsche-Texten. In: Thomas A. Schröder (Hrsg., 2000): Auf dem Weg zur Informationskultur. Wa(h)re Information? Schriften der ULB Düsseldorf.
- Zimmermann, Harald H. (2003): Zur Gestaltung eines Internet-Portals als offenes Autor-zentriertes Kommunikationssystem. In: Hennings/Grudowski/Ratzek

(Hrsg., 2003). Leben in der Informationsgesellschaft. Festschrift für Gernot Wersig. Frankfurt: DGI, 127-138

- *) Harald H. Zimmermann, geb. 1941; Studium der Germanistik, Geschichte, Pädagogik und Philosophie (Staatsexamen 1967, Saarbrücken); Promotion in Sprachwissenschaft und Informatik (1972; Saarbrücken), Professur (H3) für Nichtnumerische Datenverarbeitung / Informationslinguistik an der Universität Regensburg (1974-1980); Professur (C4) für Informationswissenschaft an der Universität des Saarlandes, Saarbrücken (1980 bis heute). Forschungsschwerpunkte: Maschinelle Sprachdatenverarbeitung und Information Retrieval; Betriebliche Information und Kommunikation; Fachinformation.