

**Dokumentation des Projektteiles „Extrakte“ des
Informationssystems „Nietzsche Online“ der
Informationswissenschaft Saarbrücken**

Inhalt

1. Einleitung	Seite 2
2. Der Projektteil „Extrakte“	Seite 2
3. Dokumentation des Projektteils „Werkextrakte“	Seite 3
3.1 Struktur des Lexikons	Seite 3
3.2 Zielsetzung bei der Umsetzung in elektronische Form	Seite 4
3.3 Die Umsetzung der Werkextrakte in elektronische Form – Aufbau der xml-Datei	Seite 5
3.3.1 Der „Head“	Seite 6
3.3.2 Der „Body“	Seite 7
4. Dokumentation des Projektteils „Briefextrakte“	Seite 11
4.1 Voraussetzungen – Die Struktur der Briefextrakte	Seite 11
4.2 Zielsetzung bei der Aufnahme in das Informationssystem „Nietzsche Online“	Seite 12
4.3 Umsetzung der Briefextrakte in eine xml-Struktur	Seite 12
4.3.1 Der „Head“	Seite 13
4.3.2 Der „Body“	Seite 15
5. Kurze Zusammenfassung – Überblick über gegenwärtigen Stand und Probleme des Projektteils „Extrakte“	Seite 15
6. Anhang	Seite 18
6.1 Darstellung der Klassifikation der Werkextrakte am Beispiel von Werkextrakt ex_li_0595	Seite 18
6.2 Ausführlich kommentierte xml-Datei zu Werkextrakt ex_li_0595	Seite 18
6.3 Aufbau der Klassifikation der Briefextrakte	Seite 20
6.4 Ausführlich kommentierte xml-Datei zu Briefextrakt ex_br_0237	Seite 20

1. Einleitung

Das Informationssystem „Nietzsche Online“ (<http://nietzsche.ps.uni-sb.de/>¹) ist ein Projekt, das im Rahmen des Forschungsseminars „Elektronisches Publizieren am Beispiel von Friedrich Nietzsche“ im Wintersemester 2000/01 am Fachbereich 5.6 Informationswissenschaft der Universität des Saarlandes (in Zusammenarbeit mit Johann Prossliner, Herausgeber diverser Schriften von und über Friedrich Nietzsche) entstanden ist, und gegenwärtig vervollständigt wird.

„Nietzsche Online“ hat im Rahmen des Forschungsseminars modellhaften Charakter und soll die Grundlage für ein globales, permanentes und professionell gepflegtes Informationssystem bilden. Dementsprechend sind die einzelnen Projektteile (FAQ, Briefe, Porträts, Extrakte, Werke, Leben, Bibliographie, Orte, Linksammlung) lediglich beispielhaft realisiert und verzichten auf eine umfassende Bereitstellung der gesamten Inhalte.

Das gesamte Projekt wird mit Hilfe der Beschreibungssprache xml (Extensible Markup Language, ein Sgml Subset) in elektronische Form umgesetzt, da diese größtmögliche Flexibilität hinsichtlich der verwendeten Auszeichnungsbefehle (die im Prinzip von den Gestaltern des Informationssystems selbst erstellt werden) garantiert, und zudem eine Trennung von Inhalt und Darstellung vorgibt, so daß die inhaltliche Erschließung weitgehend unabhängig von der späteren Präsentation ist (näheres zu xml ist auf den Seiten des „Nietzsche Online“ Projekts zu finden).

Im Folgenden wird der Projektteil „Extrakte“, der gegenwärtig in die Bereiche „Werkextrakte“ und „Briefextrakte“ unterteilt ist, in Hinblick auf die Grundlagen, die Ziele und die deren Umsetzung beschrieben.

2. Der Projektteil Extrakte

Unter „Extrakten“ sind innerhalb dieses Projektes im weitesten Sinn Auszüge aus Schriften aller Art von oder an Friedrich Nietzsche zu verstehen. Gegenwärtig ist der Projektteil „Extrakte“ in zwei klar voneinander abgesetzte Teile getrennt.

Zum ersten sind dies „Werkextrakte“; dabei handelt es sich um Auszüge aus Werken Friedrich Nietzsches, die satzenhaften bzw. sprichwortartigen Charakter aufweisen. Die Auswahl dieser Extrakte folgt dem von Johann Prossliner herausgegebenen Nietzsche-

¹ Letzter Besuch: 30.03.2001

Zitatenlexikon „Licht wird alles, was ich fasse“². Im Rahmen des Forschungsseminars beschränkt man sich auf die elektronische Umsetzung von Extrakten, die aus „Ecce homo“ und „Also sprach Zarathustra“ stammen, da zu diesen Auszügen innerhalb des Projektes die Volltexte (bei „Nietzsche Online“ unter dem Menüpunkt „Werke“ greifbar) bearbeitet wurden³.

Bei dem zweiten Teil der Extrakte handelt es sich um eine Umsetzung von Auszügen aus Briefen, die Friedrich Nietzsche verfaßt hat. Die Auswahl dieser Briefextrakte geht ebenfalls auf Herrn Prossliner zurück, die Quelle dieser Extrakte lag bis zum Abschluß dieser Dokumentation lediglich in elektronischer und unvollständiger Form vor; sie wird zur Zeit noch bearbeitet und wird bis zu ihrer offiziellen Veröffentlichung noch einige Erweiterungen erfahren.

Die Auswahl der gegenwärtig in das Informationssystem übernommenen Briefextrakte orientiert sich an den in „Nietzsche Online“ als Volltexte verfügbaren Briefen (unter dem Menüpunkt „Briefe“ zu finden), beschränkt sich jedoch auf einige prägnante Beispiele.

Brief- und Werkextrakte unterscheiden sich hinsichtlich ihrer Struktur, ihres Umfangs und ihrer Intention in solchem Maße voneinander, daß eine gemeinsame, einheitliche Umsetzung in elektronische Form nicht sinnvoll ist. Aus diesem Grund werden diese beiden Projektteile getrennt und weitgehend unabhängig voneinander bearbeitet und hier entsprechend einzeln dokumentiert.

3. Dokumentation des Projektteils „Werkextrakte“

Wie zuvor schon erwähnt beschäftigt sich dieser Teil des Projekts mit der Umsetzung des Nietzsche-Zitatenlexikons „Licht wird alles, was ich fasse“, vorläufig jedoch nur mit den aus „Ecce homo“ und „Also sprach Zarathustra“ entnommenen Auszügen.

3.1 Struktur des Lexikons

Das Nietzsche-Zitatenlexikon in der hier relevanten Ausgabe von 1999⁴ beinhaltet 2309 Auszüge aus Werken Nietzsches, die vom Textumfang etwa die Hälfte des Buches ausmachen. Die zweite Hälfte des Lexikons wird von einem umfangreichen Register (bestehend aus über 35000 Stichwörtern) eingenommen.

² Prossliner, Johann (Hrsg.): „Licht wird alles, was ich fasse“. Das Lexikon der Nietzsche-Zitate“. München 1999.

³ Siehe dazu die Dokumentation des Projektteils „Werke“ von Dorit Günther

⁴ Eine überarbeitete Ausgabe ist in Vorbereitung

Die Werkextrakte sind durch Überschriften und Unterüberschriften thematisch in eine Klassifikation eingeordnet, wobei ein Extrakt prinzipiell in jeder Klasse ab der 2. Hierarchieebene der Klassifikation direkt zugeordnet sein kann.⁵

Jedes Extrakt ist mit einer Numerierung und einem Verweis zu dem Volltext, aus dem es stammt, versehen. Die Zitate sind von verschiedener Länge, im allgemeinen aber recht kurz (ein bis wenige Sätze), wobei der Herausgeber aus dem Original oftmals Textpassagen herausgestrichen hat (im der Vorlage durch „<>“ gekennzeichnet). Die Orthographie wurde weitgehend an die Regeln der Rechtschreibreform von 1996/98 angeglichen.

Die Bedeutung des Registers ist allein schon aus dessen Umfang ersichtlich. Die Schlagworte sind hier alphabetisch geordnet, wobei auch hier Unterbegriffe in eine – in diesem Falle zweistufige – Klassifikation eingegliedert sind. Jedem dieser Unterbegriffe ist mindestens eine Zitatnumerierung zugeordnet. Das Register wurde in eine Datenbank übertragen und konnte bei der Umsetzung der Extrakte in elektronische Form genutzt werden.

3.2 Zielsetzung bei der Umsetzung in elektronische Form

Zweierlei Gesichtspunkte stehen bei der Zielsetzung im Mittelpunkt:

1. Möglichst alle Informationen, die das Buch vermittelt, sollen erhalten werden; also neben den Extrakten als eigentlichem Text auch noch die im Lexikon enthaltenen Metainformationen, nämlich die Numerierung der Extrakte, die Quellenangaben der einzelnen Auszüge, sowie die durch das Register geschaffenen Bezüge und Suchmöglichkeiten.
2. Es sollten informationelle Mehrwerte bei der Einbindung in das Informationssystem „Nietzsche Online“ entstehen, vor allem durch die Verlinkung der Extrakte zu den entsprechenden Volltexten (also zu den Stellen der Werke, aus denen dieses Extrakt stammt). Darüber hinaus sollten Links zu Personen, Orten oder Schlüsselbegriffen umfassende Auskünfte über Hintergründe des Extraktes liefern, und so Verständnisschwierigkeiten beseitigen bzw. das Wissen über das Leben und das Umfeld Friedrich Nietzsches vergrößern.

Darüber hinaus fiel allgemein die Entscheidung, die Rechtschreibung in sämtlichen Texten gänzlich an die Regeln der neuen Rechtschreibreform anzugleichen.

⁵ Siehe auch Anhang 6.1

3.3 Die Umsetzung der Werkextrakte in elektronische Form – Aufbau der xml-Datei

Jedes Werkextrakt bildet eine eigenständige Datei. Diese Dateien sind zweiteilig aufgebaut; sie bestehen aus einem „Head“, in dem Informationen über das Extrakt festgehalten sind, die im allgemeinen später nicht in den dem Nutzer präsentierten Text einfließen, und einem „Body“, in dem der eigentliche Text des Extrakts – versehen mit den für die Präsentation nötigen Auszeichnungsbefehlen („Tags“) – steht.

Darüber hinaus stehen zu Beginn des jeder dieser xml-Dateien zwei Befehlszeilen, wovon die erste rein technische Informationen über die verwendete xml-Version und ISO-Norm angibt, und in keinem inhaltlichen Zusammenhang mit dem Extrakt steht, und die zweite dazu dient, das Extrakt als solches kenntlich zu machen, es als Werkextrakt zu identifizieren (typ=“werk“) und ihm eine ID zur Ordnung und Unterscheidung von anderen Werkextrakten zuzuweisen. Dabei steht „ex_“ für Extrakt, „li_“ für das Nietzsche-Zitatenlexikon „Licht wird alles, was ich fasse“; die vierstellige Zahl entspricht der Numerierung, die Prossliner in seinem Lexikon vorgibt.

Grundgerüst der xml-Datei für Werkextrakte

```
<?xml version="1.0" encoding="iso-8859-1" ?>

<extrakt typ="werk" id="ex_li_XXXX">
  <head>
    <autor id="per_prossliner_johann_00" typ="herausgabe">
      <datum>
        <jahr>1999</jahr>
      </datum>
    </autor>
    <autor id="per_nietzsche_friedrich_00" typ="verfassung">
      <datum>
        <jahr>numerische Jahresangabe</jahr>
      </datum>
    </autor>
    <design id="per_bieg_oliver_00" typ="erstellung">
      <datum>
        <jahr>numerische Jahresangabe</jahr>
        <monat>numerische Monatsangabe</monat>
        <Tag>numerische Angabe des Tages</Tag>
      </datum>
    </design>
```

```

<werk_ref id="we_zar_Buchnr._Kapitelnr." typ="volltext"/>
<quelle_ref id="quelle_prossliner_johann_000001"/>
<prossliner nummer="Nummer des Extraktes" verweis="natürlichsprachiger Verweis aus dem
Lexikon" seite="numerische Seitenangabe"/>
<klasse name="Jeweils niedrigste Klasse von Überschriften" />
<sw>Schlagwort</sw>
</head>
<body sprache="DE" form="prosa">
<p>
Eigentlicher Extrakttext, versehen mit Auszeichnungsbefehlen
</p>
</body>
</extrakt>

```

3.3.1 Der „Head“

Die erste Angabe im head (`<autor>`) identifiziert Johann Prossliner – in Form einer ID, die auf ein Porträt innerhalb des Informationssystems verweist – als ursprünglichen Herausgeber des Extraktes, mit der Angabe, in welchem Jahr er in dieser Funktion tätig war.

Der darauf folgende *Autoren-Tag* kennzeichnet Friedrich Nietzsche als Verfasser, also als eigentlichen Autor des Extraktes, ebenfalls in Form einer ID, verbunden mit der Jahresangabe der Ersterscheinung der Schrift, aus der das Extrakt „gewonnen“ wurde.

Als nächstes werden Designer (in diesem Fall vom Typ „erstellung“, „verschlagwortung“ und „modifikation“ sind ebenfalls möglich) und letztes Bearbeitungsdatum – mit numerisch genauer Angabe von Jahr, Monat und Tag – beschrieben. Sollten mehrere Designer an dem Extrakt gearbeitet haben, so muß für jeden Designer solch ein Eintrag in den Head vorgenommen werden.

Daran schließt sich die Referenz des dazugehörigen Volltextes innerhalb des Informationssystems an. Im Beispiel wird auf den Volltext von „Also sprach Zarathustra“ („zar“ für „Zarathustra“, „ecc“ für „Ecce Homo“) mit der Angabe, in welchem Buch, und in welchem Kapitel die entsprechende Textstelle gefunden werden kann, verwiesen. Daraus wird dann in der endgültigen Präsentation ein Link zu dem entsprechenden Kapitel im „Zarathustra“ generiert. Die Verweise beschränken sich zur Zeit auf ganze Kapitel, das heißt, der erstellte Link kann nur auf den jeweiligen Kapitelanfang verweisen. Es besteht allerdings die Überlegung, in Zukunft durch im Volltext eingebundene „Anker“ (HTML-Standards entsprechend) oder eine detailliertere Gliederung der Volltexte einen Link zu genau der

Textpassage zu erstellen, wo das Extrakt – unter Umständen graphisch hervorgehoben – vorzufinden ist.

Der nächste *Tag* referenziert, in Form einer ein-eindeutigen ID, die „physische“ Quelle - hier also das Nietzsche-Zitatenlexikon – die im Projektteil „Bibliographie“ genau beschrieben wird.

Der sich anschließende „Prossliner“-*Tag* verzeichnet folgende Metainformationen aus dem Zitatenlexikon: Die Nummer, die das Extrakt dort trägt, den Verweis auf die physische Quelle des Zitates, und die Seite, auf der das Extrakt in „Licht wird alles, was ich fasse“ vorzufinden ist. Gegenwärtig dienen diese Angaben lediglich zur leichteren Nachprüfbarkeit der Extraktquellen.

Von größter Wichtigkeit dagegen ist die Angabe der „klasse“ in der darauffolgenden Befehlszeile. Hier ist zu jedem Extrakt die jeweils nächste, rangniedrigste Überschrift verzeichnet. Dadurch kann jedes Extrakt in die Klassifikation der Überschriften eingeordnet, und die thematische Gliederung des Zitatenlexikons in das Informationssystem übernommen werden. Somit können die Extrakte über ein noch zu erstellendes, der Buchvorlage ähnliches Inhaltsverzeichnis zugänglich gemacht werden, so daß auch bei „Nietzsche Online“ eine den ursprünglichen Vorstellungen des Herausgebers folgende Herangehensweise an die Zitatensammlung möglich bleibt.

Ein anderes Ordnungskriterium stellen die Schlagworte (<sw>) dar, die in beliebiger Häufigkeit – mindestens jedoch einmal – im Head stehen müssen. Diese orientieren sich weitgehend an den im Buchregister verzeichneten Schlagworten. Allerdings mußten diese meist etwas umgestaltet werden, da sie im Original häufig nicht als einzelne Begriffe, sondern als phrasenhafte Ausdrücke verzeichnet sind. Da die Schlagworte jedoch mit dem Ziel in die xml-Struktur aufgenommen werden, die Extrakte über Suchmaschinen zugänglich zu machen, erwies sich diese „Phrasenhaftigkeit“ als ungünstig, und die ursprünglichen Schlagworte mußten großenteils in einzelne Begriffe aufgegliedert werden.

3.3.2 Der „Body“

Am Anfang jedes „Bodies“ muß die Sprache (in aller Regel „DE“ für deutsch, daneben können auch noch „DEO“, für deutsch in der Originalschreibweise, „FR“ für französisch, „EN“ für englisch, „IT“ für italienisch, „LT“ für lateinisch und „GR“ für griechisch als Standardsprachen festgelegt werden), in der das Extrakt vorwiegend geschrieben ist, sowie die literarische Gattung des Textes (im allgemeinen „prosa“, aber auch „lyrik“) festgelegt werden.

Der eigentliche Text der Extrakte folgt, so weit wie möglich, der Quelle („Licht wird alles, was ich fasse“). Dabei können folgende *Tags* verwandt werden:

`<p>text</p>`

Mit diesem *Tag* werden einzelne Absätze (**paragraphen**) eines Textes kenntlich macht. Jedem Extrakttext steht der Eröffnungst*ag* `<p>` vor, der Abschluß jedes Textes wird durch den Schluß*tag* `</p>` markiert.

`<auslassung/>`

zur Kennzeichnung von Stellen, die Prossliner in seinem Buch mit „◊“ gekennzeichnet hat. Dabei handelt es sich um Auslassungen bzw. Streichungen von Textpassagen des Quelltextes.

`listenpunkt n...`

Hiermit werden ungeordnete Aufzählungen bzw. nicht numerierte Listen (HTML entsprechend) markiert.

`listenpunkt n...`

Mit diesen Befehlen wird eine geordnete Aufzählung bzw. numerierte Liste erstellt.

`<luecke/>`

zum Markieren von nicht lesbaren oder fehlenden Stellen im Originaltext.

`<sprache="XX">fremdsprachentext</sprache>`

Dieser *Tag* umschließt Textstellen, die in einer anderen Sprache als der anfangs als Standard festgelegten verfaßt sind. Es können die gleichen Sprachen gewählt werden, wie schon beim Eröffnungst*ag* erwähnt. Theoretisch ist es möglich, diese Textstellen graphisch hervorzuheben (z.B. zu kursivieren) oder mit der jeweiligen Übersetzung zu verlinken.

`<rede id="per_XX_00" typ="direkt">text der rede</rede>`

Hiermit wird die (in diesem Falle „direkt“, „indirekt“ wäre auch noch möglich) Rede eines Protagonisten gekennzeichnet, die von ihrem Umfang weniger als einen Absatz einnimmt. Der Redner muß mittels einer eindeutigen ID identifiziert werden. Somit ist eine Verlinkung zu dem jeweiligen Porträt des Redners möglich, und jede Rede kann – ihrem Typen entsprechend – graphisch verändert werden. Über eine spezielle Suchanfrage könnte man sich beispielsweise auch alle Extrakte ausgeben lassen, bei denen Zarathustra als Redner auftritt.

`<rede_p id="per_XX_00" typ="direkt"><p>text der rede</p></rede_p>`

Im Prinzip hat dieser *Tag* eine ähnliche Funktion wie der vorangegangene; allerdings wird dieser ausschließlich für Reden gebraucht, die sich zumindest über einen ganzen Absatz erstrecken. Dies ist insbesondere bei den aus „Also sprach Zarathustra“ entnommenen Extrakten häufig der Fall. Die überaus kompliziert verschachtelte Struktur der Reden aus diesem Werk verursacht – bedingt durch die Kürze der Extrakte – nur in wenigen Fällen Probleme. In diesen Zweifelsfällen dient der Volltext der Werke innerhalb des Informationssystems, der sich intensiv mit dem Aufbau der Redestruktur im „Zarathustra“ auseinandersetzt, als Referenz. Da es in diesem Projekt allerdings um die Umsetzung des Nietzsche-Zitatenlexikons in das Informationssystem geht, und der Gestalter die Extrakte inhaltlich möglichst unverändert in elektronische Form umsetzen sollte, sind Abweichungen vom Volltext innerhalb des Extrakttextes in Kommentaren (*<!--kommentar -->*) vermerkt. Im einzelnen ist noch abzuklären, wie mit diesen Abweichungen letztendlich umzugehen ist.

```
<eingriff original="ursprungstext oder kein eintrag, falls nur neuer text  
hinzukommt">neuer / eingesetzter text</eingriff>
```

Dieser *Tag* wird verwendet, um Eingriffe oder anderweitige Abweichungen des Extraktes vom Ursprungstext kenntlich zu machen, verbunden mit der Möglichkeit, dem Nutzer auf Wunsch das Original oder den veränderten Text zu präsentieren.

In aller Regel werden die Texteingriffe vorgenommen, um syntaktische Fehler zu beseitigen oder semantische Lücken zu schließen, die dadurch entstehen, daß viele Extrakte aus einem größeren Textzusammenhang herausgerissen sind.

```
<zitat_ref id="XX">zitierter text</zitat_ref>
```

Versieht Textpassagen, die nicht von Nietzsche stammen, über eine eindeutige ID mit einem Link zu dem Werk aus dem das Zitat stammt, oder verweist zumindest auf dessen bibliographische Angaben.

```
<person_ref id="XX">name</person_ref>
```

Dieser *Tag* dient dazu, Personen, die in den Extrakten erwähnt werden, mittels ihrer ID mit deren Porträts zu verlinken.

<ort_ref id="XX">*ort*</ort_ref>

Verweist mittels ID auf eine im Informationssystem vorliegende Datei, die über diesen Ort informiert.

<faq_ref id="XX">*begriff, zu dem eine faq vorliegt*</faq_ref>

Mittels dieses *Tags* können Zentralbegriffe, zu denen im Projektteil „faq“ Erläuterungen vorliegen, mit diesen Erklärungen verlinkt werden.

<komm_ref id="XX">*begriff, zu dem ein kommentar vorliegt*</komm_ref>

Dieser *Tag* bietet die Möglichkeit, Kommentare zu Texten über Verlinkungen mit den entsprechenden Stellen zu verknüpfen.

<e_werk_ref id="XX">*werk*</e_werk_ref>

Wird in einem Extrakt ein Werk erwähnt, dessen elektronische Fassung im Informationssystem vorliegt, so kann mit Hilfe dieses *Tags* ein Link zu dem entsprechenden Text erstellt werden.

<brief_ref id="XX">*brief*</brief_ref>

Hierüber wird ein Link zu einem Brief generiert, der im Informationssystem vorliegt.

<link_ref id="XX">*thema, zu dem ein link verfügbar ist*</link_ref>

Liegt zu einem Thema, einer Person, einem Ort etc. eine bekannte Informationsquelle außerhalb des Informationssystems „Nietzsche Online“ vor, so kann diese über die Linkliste, auf die dieser *Tag* verweist die entsprechende Site erreicht werden.

<bib_ref id="XX">*werk*</bib_ref>

Dieser *Tag* schafft einen Verweis zu einem bibliographischen Eintrag im Projektteil „Bibliographie“.

<!--kommentar →

Für Anmerkungen innerhalb des Textes, die nur für interne Zwecke bestimmt sind, ist ein Kommentar-*Tag* vorgesehen. Die darin festgehaltenen Bemerkungen tauchen nicht im für den Nutzer einsehbaren Text auf.

<!--k -->

Kommentare dieser Form zeigen kursivierte Stellen in der Buchvorlage von „Licht wird alles, was ich fasse“ an. Steht ein solcher Kommentar allein vor einem Wort, so ist nur dieses eine Wort im Original kursiv gedruckt. Folgt nach einigen Worten der abschließende <!-- /k --> *Tag*, so ist der ganze

Text, der von den beiden Kommentar-*Tags* umschlossen wird, ursprünglich kursiv gedruckt⁶.

Da es sich bei Kursivierungen um ein Gestaltungsmoment handelt, das jeweils verschiedene Bedeutungen haben kann (beispielsweise die Hervorhebung von Fremdworten, die Betonung bedeutungsschwerer Begriffe usw.), und im Zuge der elektronischen Umsetzung unter Umständen Worte in der Präsentation kursiviert werden, die es ursprünglich nicht gewesen sind (etwa, wenn alles, was Zarathustra in einem Kapitel sagt, hervorgehoben werden soll), konnte noch keine befriedigende Lösung gefunden werden.

4. Dokumentation des Projektteils Briefextrakte

Im Gegensatz zu den Extrakten im zuvor dokumentierten Projektteil handelt es sich hier nicht um ein abgeschlossenes, bereits erschienenes Buch, sondern um einen unfertigen, bisher noch nicht veröffentlichten Text, der – zumindest teilweise – in xml umgesetzt werden soll. Dies schafft eine neue Situation, für die die Erfahrungen mit der Umsetzung der Werkextrakte in vielen Bereichen unbrauchbar sind.

4.1 Voraussetzungen – Die Struktur der Briefextrakte

In der Datei, die als Textgrundlage für diese Extrakte zur Verfügung steht, sind die Briefauszüge chronologisch geordnet, beginnend mit dem Jahr 1859, endend mit dem Jahr 1889. Dieser Zeitraum ist in 4 große Epochen unterteilt: „Der Gymnasiast“, „Der Student“, „Der Professor“, „Der Frührentner“. Diese Epochen zerfallen teilweise wiederum in (Unter-)Epochen. Somit findet man hier eine zweistufige Klassifikation vor, nach der die Briefextrakte gegliedert sind⁷. Ein weiteres Ordnungselement, etwa in Form eines Registers, liegt nicht vor.

Noch mehr als der äußere Aufbau unterscheiden sich die Struktur und die Intention der Briefextrakte von denen der Werkextrakte. Die Briefe sind mit eigenem Datum (zumindest Monat und Jahr sind angegeben), Adressaten und gelegentlich dem Absendeort versehen. Der Extrakttext ist meist viel umfangreicher als der der Werkextrakte. Zudem weicht die Struktur der Briefe häufig von „reiner“ Prosa ab, es finden sich Elemente wie Anreden, Grüße, Widmungen, aber auch Passagen, die in Lyrik verfaßt sind, oder Aufzählungen und Listen. Auch inhaltlich setzen sich die Briefextrakte klar von den Werkextrakte ab. Sie zielen in

⁶ vgl. Anhang 6.4

erster Linie darauf ab, Einblicke in die Entwicklung und den Lebenslauf Friedrich Nietzsches zu geben, seinen Werdegang als Philosoph und Schriftsteller zu dokumentieren. Sie haben also stark biographische Züge, Sentenzen oder Aphorismen, wie bei den Werkextrakten, finden sich dort kaum.

Darüber hinaus trifft man bei den Briefextrakten auf die Probleme einer noch nicht veröffentlichungsreifen Schrift: sie ist „unkorrigiert, unvollständig“⁸ und die Numerierungen, mit denen jedes Briefextrakt versehen ist, sind „unrichtig“.

Die Orthographie ist schon an die Regeln der Rechtschreibreform 1996/98 angeglichen.

4.2 Zielsetzung bei der Aufnahme in das Informationssystem „Nietzsche Online“

1. Die Klassifikation der Extrakte nach Epochen soll im Informationssystem erhalten bleiben. Zudem sollte eine Stichwortsuche ermöglicht, sowie eine Verlinkung mit den Volltexten der Briefe gewährleistet werden.
2. Generell sollten durch den Einsatz von Hypertext Verlinkungen – ähnlich wie bei den Werkextrakten – zur Schaffung informationeller Mehrwerte beitragen. Wegen des starken biographischen Bezugs der Briefextrakte zu Friedrich Nietzsche ist hier die Vermittlung von Hintergrundinformationen von besonderer Wichtigkeit.

4.3 Umsetzung der Briefextrakte in eine xml-Struktur

Analog zu den Werkextrakten wird für jedes Briefextrakt eine eigene Datei – aufgeteilt in „Head“ und „Body“ – erstellt. Auch hier stehen zu Beginn jeder dieser Dateien zwei Befehlszeilen, wobei die erste genau mit der bei den Werkextrakten verwendeten Anfangszeile übereinstimmt, also lediglich Auskunft über die xml-Version und ISO-Norm gibt, die in diesen Dateien benutzt wird. Die darauffolgende Zeile gibt an, daß es sich hier um eine Datei von Typ Briefextrakt handelt, und weist ihr eine ein-eindeutige ID zu. Diese ist folgendermaßen aufgebaut: „ex_“ identifiziert die Datei als Extrakt, „br_“ weist ihr den Typen „Brief“ zu, die vierstellige Numerierung stimmt gegenwärtig mit der in der Vorlage verwendeten überein. Diese ist möglicherweise nur provisorisch, abhängig davon, ob Herr Prossliner die gegenwärtige Numerierung beibehält, oder ob er gezwungen ist, sie abzuändern.

⁷ vgl. Anhang 6.3

Struktur der Briefextrakte-xml

```
<?xml version="1.0" encoding="iso-8859-1" ?>

<extrakt typ="brief" id="ex_br_0026 ">
  <head>
    <autor id="per_prossliner_johann_00" typ="herausgabe"/>
    <design id="per_bieg_oliver_00" typ="ersteller">
      <datum>
        <jahr>2001</jahr>
        <monat>01</monat>
        <Tag>31</Tag>
      </datum>
    </design>
    <klasse name="An der Universität Bonn. Oktober 1864 - August 1865"/>
    <brief_ref id="br_k0467"/>
    <sw></sw>
  </head>
  <body sprache="DE" form="prosa">
    <p>
    </p>
  </body>
</extrakt>
```

4.3.1 Der „Head“

Der Head beginnt auch hier wiederum mit dem „*autor*“-*Tag*. Darin wird Johann Prossliner – per ID – als Herausgeber beschrieben. Anders als bei den Werkextrakten wird hier kein Datumseintrag vorgenommen, da die Textgrundlage der Extrakte offiziell noch nicht erschienen ist.

Der darauffolgende „*design*“-*Tag* ist analog zu dem entsprechenden *Tag* der Werkextrakte aufgebaut: der Designer wird mittels seiner ID schriftlich festgehalten und ihm wird ein „*typ*“ zugeordnet; daran schließt sich die genaue Datumsangabe der letzten Bearbeitung an.

Es sind sowohl für Autoren (hier als Sammelbegriff für Verfasser, Herausgeber oder Bearbeiter verwendet, also Personen, die aktiv den eigentlichen Text schaffen oder verändern), als auch für Designer (die für die technische Umsetzung der Texte zuständig sind) mehrere Einträge möglich.

Gegenwärtig ist es nicht nötig, Nietzsche als Verfasser zu kennzeichnen, da bisher ausschließlich Briefe, die dieser selbst geschrieben hat, als Extraktgrundlage genutzt werden.

⁸ Zu Beginn des elektronischen Briefextrakteskripts von Prossliner vorzufinden

Generell sollten redundante Angaben vermieden werden, um Nicht zu viel „Datenballast“ im Informationssystem entstehen zu lassen. Wo Sachverhalte eindeutig sind (wie hier etwa, daß Nietzsche Verfasser der Briefe ist), oder aus anderen Quellen erschlossen werden können, sollten sie nicht mehr explizit genannt, sondern durch Verlinkungen verfügbar gemacht werden.

Aus diesem Grund wird beim „klasse“-*Tag* auch nur die Überschrift der jeweils niedrigsten Ordnung, also die kleinstmögliche Epochenbenennung, angegeben. Im Beispiel ist dies "An der Universität Bonn. Oktober 1864 - August 1865", die in der Klassifikation der Epoche „Der Student 1864-1869“ untergeordnet wird. Somit geht aus der untersten Klasse die darüberliegende hervor. Über diese Klassifikation wird es möglich sein, die Extrakte wie über das Inhaltsverzeichnis eines Buches zugänglich zu machen⁹.

Die nächste Befehlszeile „brief_ref“ ist von äußerster Wichtigkeit. Durch diese Briefreferenz wird nicht nur eine Verlinkung mit dem Volltext des betreffenden Extraktes ermöglicht, sondern auch sämtliche Informationen, die in der referenzierten Datei enthalten sind, erschlossen und via Stylesheet für die Präsentation des Briefextraktes verfügbar gemacht werden.

Daher müssen im „Head“ des Extraktes weder der Empfänger, das Datum, noch der Ort, an dem der Brief verfaßt wurde, angegeben werden, denn diese Angaben sind bereits im „Head“ des Briefes, auf den verwiesen wird, enthalten. Sollte ein Nutzer eine Suchanfrage stellen, in der nur solche Briefextrakte, die an einem bestimmten Ort geschrieben worden, oder die an eine bestimmte Person gerichtet sind, ausgegeben werden sollen, so ist dies ebenfalls durch diesen Verweis gewährleistet.

Um eine Durchsuchbarkeit der Briefextrakte nach Schlüsselbegriffen zu ermöglichen – eine Volltextrecherche kann in näherer Zukunft voraussichtlich nicht realisiert werden – muß zu jedem Briefextrakt eine Schlagwortliste erstellt werden. Da auf kein Register zurückgegriffen werden kann, bleibt die Auswahl dieser Schlagworte momentan dem Extraktersteller vorbehalten.

⁹ vgl. Anhang 6.3

4.3.2 Der „Body“¹⁰

Wie bei den Werkextrakten beginnt auch hier jeder „Body“ mit der Angabe der Standardsprache des Extraktes und der Typenangabe, die den Text einer bestimmten Gattung zuweist. Darauf folgt grundsätzlich der Eröffnungs-*Tag* eines Absatzes <p> .

Der eigentliche Extrakttext setzt sich aus zwei klar voneinander unterscheidbaren Teilen zusammen, nämlich dem Briefkopf, in dem Adressat (evtl. mit Aufenthaltsort) und Briefdatum festgehalten sind, und dem Briefkörper, also dem Extrakt an sich.

Die im Briefkopf enthaltenen Daten entsprechen also den Angaben, die – wie oben erwähnt – über den Verweis zu den Briefvolltexten für die Präsentation der Extrakte zur Verfügung stehen. Der Briefkopf kann also maschinell generiert werden, und die entsprechenden Daten müssen – als redundante Angabe – nicht in der Briefextraktdatei eingetragen werden.

Der Briefkörper kann im allgemeinen analog zu den Werkextrakten verarbeitet werden. Allerdings ist der Briefextrakttext meist wesentlich umfangreicher als der der Werkextrakte und enthält viele Anspielungen auf Werke anderer Künstler und Schriftsteller, viele Personen und Orte werden genannt. Dementsprechend dicht ist das Netz von Verweisen zu Orts- oder Personenporträts.

Weiterhin weicht der Körper des Briefextraktes auch in der Hinsicht von den Werkextrakten ab, als daß häufig Listen oder Aufzählungen darin vorkommen, und daß man Elemente wie Unterschriften, Widmungen etc., die für einen Brief typisch sind, vorfindet. Allerdings kann sämtlichen dieser Besonderheiten mit den schon bei den Werkextrakten beschriebenen Auszeichnungsbefehlen Rechnung getragen werden.

Da somit das gleiche Kontingent an *Tags* für die Briefextrakte wie für die Werkextrakte zur Verfügung steht, brauchen diese hier nicht noch einmal aufgeführt und erklärt zu werden.

5. Kurze Zusammenfassung – Überblick über gegenwärtigen Stand und Probleme des Projektteils „Extrakte“

Die Zielsetzung für die Aufnahme der Brief- und Werkextrakte konnte weitgehend erfüllt werden.

Die Orthographie der Werkextrakte ist vollkommen an die neue deutsche Rechtschreibreform angeglichen.

¹⁰ vgl Anhang 6.4

Der Text sowohl der Werk- als auch der Briefextrakte konnte, versehen mit den entsprechenden Auszeichnungsbefehlen und zahlreichen Referenzierungen, was insbesondere bei der Bearbeitung eine intensive Auseinandersetzung mit dem Text erfordert, inhaltlich weitgehend unverändert in das Informationssystem „Nietzsche Online“ übernommen werden. Auf der Textebene (also dem Extrakt-„Body“) bestehen zur Zeit vor allem wegen der mangelnden Genauigkeit der Angabe von Sprechern bei den „Zarathustra“-Extrakten Schwierigkeiten.

Momentan wird auf diese und andere kleine Unstimmigkeiten lediglich in Kommentaren hingewiesen. Eine befriedigende Lösung hat sich noch nicht ergeben.

Ebenso ist noch keine Lösung für den Umgang mit in der Vorlage kursiv gedruckten Textpassagen gefunden worden. Auch auf diese Stellen weisen bislang lediglich Kommentare hin.

Die in den Vorlagen bestehenden Klassifikationen der Werk- und der Briefextrakte können praktisch unverändert in das Informationssystem übernommen werden, und werden so später einen Zugang zu den Extrakten ermöglichen, der dem Zugang über ein Inhaltsverzeichnis eines Buches sehr nahe kommt.

Die Verlinkung der einzelnen Extrakte zu den Volltexten konnte in großen Teilen realisiert werden. Bei den Werkextrakten ist diese Verlinkung bislang jedoch noch nicht differenziert genug umgesetzt. Durch die Referenz der Briefvolltexte wird bei den Briefextrakten nicht nur eine Verlinkung generiert, sondern sämtliche in der Volltextdatei enthaltenen Daten für die Briefextrakte verfügbar gemacht. Somit wird Datenredundanz vermieden und der Arbeitsaufwand reduziert.

Das Register des Zitateslexikons wird für die Erstellung einer Schlagwortliste genutzt, kann jedoch nicht in seiner ursprünglichen Form verwandt werden. Durch die Aufsplittung der Schlagwortphrasen sind die Werkextrakte später für elektronische Suchanfragen verfügbar.

Der Mangel eines Registers bei den Briefextrakten wird durch eine Schlagwortliste im „Head“ eines jeden Extraktes ausgeglichen. Da die Auswahl dieser Schlagworte jedoch dem Designer überlassen bleibt, mangelt es an Systematik und Wissenschaftlichkeit bei der Erstellung dieser Listen.

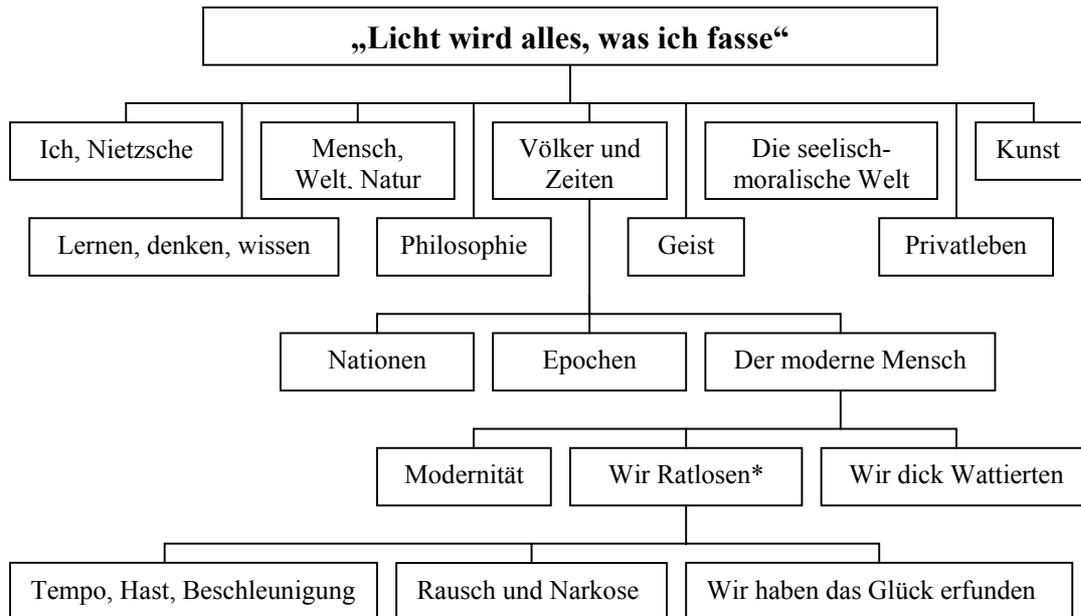
Es ist möglich, daß in näherer Zukunft die Numerierung sowohl der Brief- als auch der Werkextrakte völlig verändert werden muß, da bei der geplanten Neuauflage des Zitateslexikons und beim Erscheinen der Briefextraktesammlung in Buchform die momentan gültige Numerierung wahrscheinlich nicht übernommen wird.

Da bis zur Fertigstellung dieser Dokumentation noch kein Stylesheet und keine xsl für die Extrakte erarbeitet worden ist, konnte bisher weder ein Brief- noch ein Werkextrakt in seiner endgültigen Form präsentiert werden. Die hier beschriebenen Umsetzung der Extrakte in xml mit all ihren geplanten Funktionen, wie etwa Verlinkungen und Durchsuchbarkeit konnten also in noch keiner praktischen Anwendung erprobt und überprüft werden. Dementsprechend sind Umarbeitungen der xml-Struktur, die Einführung neuer Auszeichnungsbefehle oder anderweitige Änderungen noch immer möglich.

6. Anhang

6.1 Darstellung der Klassifikation der Werkextrakte am Beispiel von Werkextrakt ex_li_0595

Im Folgenden wird die Struktur der Überschriften des Zitateslexikons „Licht wird alles, was ich fasse“, die eine Klassifikation darstellt, in Bezug auf das Werkextrakt ex_li_0595 graphisch veranschaulicht. Ab der 2. Hierarchieebene können einer Klasse einzelne Extrakte zugeordnet sein. Hier ist nur der Zweig, der zu diesem Extrakt führt, komplett ausgearbeitet.



* Den Extrakten Nr. 0579, 0580 und 0581 steht „Wir Ratlosen“ direkt als Überschrift vor.

Extrakt Nr. 0595 wird im Head die Klasse „Wir haben das Glück erfunden“ zugeordnet. Aus der Graphik ist nun leicht ersichtlich, daß das Extrakt durch diese Zuweisung unter „Völker und Zeiten/Der moderne Mensch/Wir Ratlosen/Wir haben das Glück erfunden“ in die Klassifikation eingeordnet werden kann.

6.2 Ausführlich kommentierte xml-Datei zu Werkextrakt Nr. 0595

Kommentare zur Erläuterung der Befehle sind von `<!-- -->` eingeschlossen, und in der kursivierten Schriftart „Times New Roman“ verfaßt.

```

<?xml version="1.0" encoding="iso-8859-1" ?>
<!-- rein technische Datenangabe -->

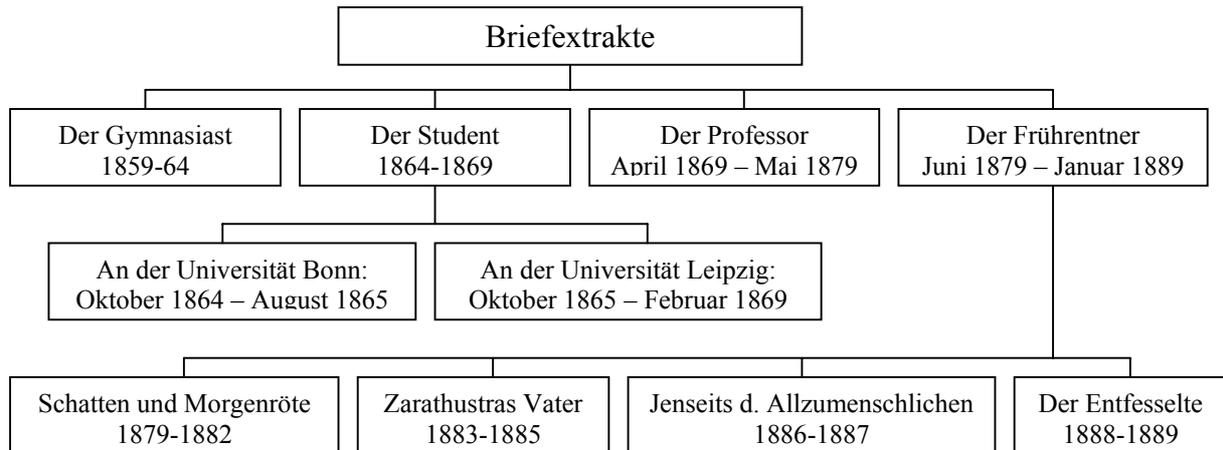
<extrakt typ="werk" id="ex_li_0595">
<!-- Es handelt sich hier um ein Extrakt vom Typ „werk“, also ein Werkextrakt mit der ID „ex_li_0595“ („ex“ für „Extrakt“, „li“ für „Licht wird alles, was ich fasse“ und die Zahl für die Numerierung, die es im Lexikon trägt; somit handelt sich um Extrakt Nr.0595 aus dem Nietzsche-Zitateslexikon)-->
  <head>
<!-- Beginn des „Heads“-->
    <autor id="per_prossliner_johann_00" typ="herausgabe">
<!-- Johann Prossliner wird mittels ID als Herausgeber identifiziert-->
      <datum>
        <jahr>1999</jahr>
      </datum>
<!-- Angabe des Jahres, in dem Prossliner als Herausgeber tätig war-->
  
```

```

</autor>
<autor id="per_nietzsche_friedrich_00" typ="verfassung">
  <datum>
    <jahr>1883</jahr>
  </datum>
</autor>
<!-- Friedrich Nietzsche wird per ID als Verfasser kenntlich gemacht, ebenfalls verbunden mit der Angabe, in
welchem Jahr er diese Tätigkeit ausgeübt hat-->
  <design id="per_bieg_oliver_00" typ="erstellung">
    <datum>
      <jahr>2001</jahr>
      <monat>03</monat>
      <Tag>20</Tag>
    </datum>
  </design>
<!-- Angabe des Designers Oliver Bieg, der am 20. März 2001 zum letzten Mal etwas an diesem Werkextrakt
verändert hat-->
  <werk_ref id="we_zar_01_00" typ="volltext"/>
<!-- Verweis zum Volltext, aus dem das Extrakt entnommen wurde. Hier ist dies die Vorrede (_00) des ersten
Buches (_01) des Werkes „Also sprach Zarathustra“ (_zar) -->
  <quelle_ref id="quelle_prossliner_johann_000001"/>
<!-- Verweis auf die Quelle „Licht wird alles, was ich fasse“ mittels ID-->
  <prossliner nummer="0595" verweis="ZAR I, Vorrede 3" seite="75"/>
<!--Angabe von Metainformationen aus dem Zitatenlexikon: Nummer des Extraktes, Verweis auf den dritten Teil
der Vorrede im ersten Buch von „Also sprach Zarathustra“; das Extrakt ist auf Seite 75 zu finden-->
  <klasse name="Wir haben das Glück erfunden"/>
<!-- Die rangniedrigste Überschrift des Extraktes. Mittels dieser Angabe ist eine Einordnung des Zitates in die
Klassifikation möglich (vgl. Anhang 6.1) -->
  <sw>Geiz in der Sünde</sw>
  <sw>Genügsamkeit schreit gen Himmel</sw>
  <sw>Zufriedenheit</sw>
<!-- Schlagwortliste, für Suchmaschinen bestimmt, und aus dem Register des Lexikons gewonnen -->
</head>
<!-- Ende des „Heads“ -->
  <body sprache="DE" form="prosa">
<!-- Beginn des „Bodies“ -->
  <p>
<!-- Der Anfang eines Absatzes wird markiert (paragraph) -->
Nicht eure Sünde - eure Genügsamkeit schreit gen Himmel, euer Geiz selbst
in eurer Sünde schreit gen Himmel! <!-- Zarathustra ist Sprecher -->
<!-- Der eigentliche Text des Extrakts – versehen mit einem Kommentar, daß Zarathustra eigentlicher Sprecher
ist, dies allerdings von Prossliner nicht kenntlich gemacht wurde-->
  </p>
<!-- Markiert das Ende eines Absatzes-->
</body>
<!-- Hier endet der „Body“ -->
</extrakt>

```

6.3 Aufbau der Klassifikation der Briefextrakte



6.4 Ausführlich kommentierte xml-Datei zu Briefextrakt Nr. 0237

Wiederum sind erläuternde Kommentare kursiviert, in „Times New Roman“ verfaßt, und von <!-- --> eingeschlossen.

```

<?xml version="1.0" encoding="iso-8859-1" ?>
<!-- rein technische Datenangabe-->

<extrakt typ="brief" id="ex_br_0237">
<!-- Die Datei ein Extrakt vom Typ „brief“ mit der ID „ex_br_0237“ (Briefextrakt Nr. 0237)-->
  <head>
<!-- Beginn des „Heads“-->
    <autor id="per_prossliner_johann_00" typ="herausgabe"/>
<!-- Johann Prossliner wird per ID als Herausgeber identifiziert, da die Textgrundlage nicht als Buch
veröffentlicht ist, ohne Datum-->
    <design id="per_bieg_oliver_00" typ="erstellung">
      <datum>
        <jahr>2001</jahr>
        <monat>02</monat>
        <Tag>27</Tag>
      </datum>
    </design>
<!-- Angabe des Designers Oliver Bieg, der am 27. Februar 2001 zum letzten Mal etwas an diesem Werkextrakt
verändert hat-->
    <klasse name="Zarathustras Vater 1883 - 1885"/>
<!-- Mittels dieses Tags kann das Extrakt in die Klassifikation eingegliedert werden -->
    <brief_ref id="br_f0587"/>
<!-- Diese Referenz zum Volltext des Briefes, aus dem das Extrakt stammt, ermöglicht die Generierung eines
Links zu dem Brief, sowie die automatische Erstellung des Briefextraktkopfes aus den dort im „Head“
enthaltenen Metadaten -->
    <sw>schlechte Musik</sw>
    <sw>Philosophie</sw>
    <sw>Pessimismus</sw>
    <sw>Lehrling</sw>
    <sw>Schüler</sw>
    <sw>misstratener Mensch</sw>
<!-- Vom Designer erstellte Schlagwortliste, die von Suchmaschinen genutzt werden kann-->
  </head>
<!-- Ende des „Heads“ -->
  
```

```

<body sprache="DE" typ="prosa">
<!-- Beginn des „Bodies“ mit der Standardsprache deutsch, in Prosa verfaßt -->
  <p>
<!-- Der Beginn eines neuen Paragraphen wird ausgezeichnet-->
Was Musik angeht: so habe ich letzten Herbst gewissenhaft und neugierig die
Probe gemacht, wie ich <!-- k -->jetzt <!-- Das Wort „jetzt“ war in der Quelle kursiviert -->
zu <per_ref id="per_wagner_richard_00">R. Wagners</per_ref> <!-- „R. Wagner“ wird
mit einem Link zum Porträt Richard Wagners versehen --> Musik stehe. Was mir diese wolkige,
schwüle, vor allem schauspielerische und prätentiose Musik zuwider ist! So
sehr zuwider als - als - als - tausend Dinge, zum Beispiel <per_ref
id="per_schopenhauer_arthur_00">Schopenhauers</per_ref> <!-- „Schopenhauer“ wird
mit einem Link zum Porträt Arthur Schopenhauers versehen --> Philosophie. Das ist Musik eines
<!-- k -->missratenen <!-- Das Wort „missratenen“ war in der Quelle kursiviert --> Musikers und
Menschen, aber<!-- k -->eines großen Schauspielers<!-- /k --> <!-- die drei Worte
“eines großen Schauspielers“ sind im Original kursiviert--> - darauf will ich schwören. Da
lobe ich mir die tapfere und unschuldige Musik meines Schülers und Freundes
<per_ref id="per_gast_peter_00">Peter Gast</per_ref> <!-- „Peter Gast“ wird mit
einem Link zum Porträt Peter Gasts versehen -->, eines <!-- k --> echten <!-- Das Wort „echten“
war in der Quelle kursiviert --> Musikers: der mag einmal für seinen Teil dafür sorgen,
dass die Herrn Schauspieler und Schein-Genies nicht mehr zu lange den
Geschmack verderben. - Der arme <!-- k -->Stein<!-- Das Wort „Stein“ war in der Quelle
kursiviert --> ! Er hält R. Wagner sogar für einen Philosophen!
  </p>
<!-- Das Ende des Absatzes wird markiert -->
</body>
<!-- Ende des „Bodies“ -->
</extrakt>

```