

Elektronische Kanttexte

IKP-Arbeitsbericht NF 18

Hans-Christian Schmitz, Marcel Diel, Gert Keul, Simon Lüert, Winfried Lenders

schmitz@lingua.uni-frankfurt.de, Lenders@uni-bonn.de

Die Universität Bonn verfügt über eine elektronische Version von Immanuel Kants Schriften gemäß den Abteilungen 1-3 der Akademie-Ausgabe. Auf dieses Kant-Korpus kann über die Webseite des ehemaligen¹ Instituts für Kommunikationsforschung und Phonetik zugegriffen werden: <http://www.ikp.uni-bonn.de/kant>. Im vorliegenden Arbeitsbericht werden der Umfang, der Zustand und die Verwendung des Korpus beschrieben. Der Arbeitsbericht ergänzt und aktualisiert den bereits im Jahr 2003 erschienenen Bericht über die *Bereitstellung und Pflege der Werke Immanuel Kants in elektronischer Form* ([3]).

1 Das Bonner Kant-Korpus

Seit den 1960er Jahren wurden an der Universität Bonn im Rahmen der Arbeit am Allgemeinen Kant-Index die Werke, der Briefwechsel und der handschriftliche Nachlass Kants systematisch digitalisiert. Die Datenerfassung wurde mehrfach unterbrochen und zog sich über Jahrzehnte hin.² Es kamen verschiedene Techniken zu Einsatz – manuelles Stanzen von Lochkarten, Erfassung mittels OCR, ... –, außerdem wurden die Daten auf verschiedene Weisen mit Metainformation annotiert. Das Ergebnis war ein Korpus, das alle Kanttexte aus den Abteilungen 1-3 der Akademie-Ausgabe umfasste, das allerdings unheitlich kodiert und z.T. mit Fehlern belastet war. Im Zuge eines von der DFG geförderten Projekts zur *Bereitstellung und Pflege der Werke Immanuel Kants in elektronischer Form* wurde das elektronische Kant-Korpus nun ergänzt, weitgehend korrigiert und in ein standardisiertes XML-Format überführt. (Zum Umfang und Zustand der Daten zu Projektbeginn vgl. [3].)

¹ Das Institut für Kommunikationsforschung und Phonetik (IKP) ist im September 2005 im neu gegründeten Institut für Kommunikationswissenschaften (IfK) aufgegangen.

² Zur Geschichte des Bonner Kant-Korpus vgl. [2] und Lenders: *Zur elektronischen Edition der Werke Immanuel Kants* in [3].

1.1 Umfang des Korpus

Die Akademieausgabe der Schriften Kants besteht aus vier Abteilungen, nämlich erstens den schon zu Lebzeiten veröffentlichten Werken (AA 1-9), zweitens dem Briefwechsel (AA 10-13), drittens dem handschriftlichen Nachlass (AA 14-23) und viertens den Vorlesungsmitschriften (AA 24, 25, 27-19). Der handschriftliche Nachlass zerfällt in die von Erich Adickes und seinem Assistenten Friedrich Berger herausgegebenen Bände AA 14-19³ und die von Gerhard Lehmann herausgegebenen Bände AA 20-23. (Vgl. [4].) Das Bonner Kant-Korpus umfasst die Abteilungen 1-3 der Akademie-Ausgabe, d.h. die Bände AA 1-23. Die Vorlesungsmitschriften (Abt. 4, AA 24-29) sind nicht Teil des Korpus.

- *Abteilung 1 (Werke)*: Zum Korpus gehören alle von Kant verfassten Texte und die Inhaltsverzeichnisse der Bände AA 1-9. Die Herausgeberanmerkungen der gedruckten Bände sind nicht vorhanden.
- *Abteilung 2 (Briefwechsel)*: Das Korpus enthält den kompletten Briefwechsel aus der zweiten Auflage der Bände AA 10-13⁴, darunter sowohl die von Kant verfassten als auch die an Kant gerichteten Briefe. Das Korpus enthält außerdem die Inhaltsverzeichnisse der Bände AA 10-13. Der Band AA 13 besteht bis auf einen Brief fast ausschließlich aus Herausgeberanmerkungen zu den Briefen der Bände AA 10-12. Diese Anmerkungen sind nur auszugsweise vorhanden.⁵

³ Der alleinverantwortliche Herausgeber der Bände AA 14-18 ist Erich Adickes. Adickes hat auch weitgehend AA 19 zu verantworten, die Arbeit an dem Band wurde allerdings erst nach Adickes' Tod von seinem Assistenten Friedrich Berger beendet.

⁴ Die zweite Auflage der Abteilung 2 der Akademie-Ausgabe hat einen größeren Umfang – d.h.: sie umfasst mehr Briefe – als die erste Auflage.

⁵ Die Herausgeberanmerkungen zu den meisten Briefen sind nicht vorhanden. Die Anmerkungen, die überhaupt ins elektronische Korpus aufgenommen wurden, wurden nicht vollständig aufgenommen. Ein Beispiel: Mit Brief 2 (AA 10, S. 1f) verschickte Kant ein Exemplar seiner Erstlingsschrift *Gedanken von der wahren Schätzung der Lebendigen Kräfte* zusammen mit der Bitte, eine Rezension zu veröffentlichen. Der Empfänger des Briefes war vermutlich Albrecht von Haller, der Herausgeber der *Göttingischen Zeitungen von Gelehrten Sachen*. In der Anmerkung zu Brief 2 (AA 13, S. 1f) wird u.a. bemerkt, dass 1750 eine rein referierende Rezension in den *Göttingischen Zeitungen* veröffentlicht wurde und dass Lessing vermutlich auf Grund dieser Rezension das folgende Gedicht verfasst habe: „Auf des Herrn K* Gedanken\\ K* unternimmt ein schwer Geschäfte,\\ Der Welt zum Unterricht.\\ Er schätzt die lebendgen Kräfte,\\ Nur seine schätzt er nicht.“ Im elektronischen Kant-Korpus ist von der Herausgeberanmerkung zu Brief 2

- *Abteilung 3, AA 14-19 (von Erich Adickes respektive Friedrich Berger herausgegebener Teil des handschriftlichen Nachlasses):* Der von Erich Adickes und seinem Assistenten Friedrich Berger herausgegebene Teil des handschriftlichen Nachlasses ist in 8141 Einzelnotizen aufgeteilt.⁶ Jede Notiz beginnt mit einem Notizkopf, in dem die jeweilige Notiznummer, die Datierung, die Fundstelle und ggf. weitere Anmerkungen stehen (s. Anhang A). Zum Kant-Korpus gehören alle Notizen zusammen mit den Notizköpfen und den Herausgeberanmerkungen, welche innerhalb der Notizen stehen.⁷ Das Korpus umfasst nicht die Herausgeberanmerkungen, die als Fußnoten in den gedruckten Bänden erscheinen. Ausnahmen: Einige wenige Fußnoten aus den Bänden AA 17 und 18 sind vorhanden. Schließlich gehören die Inhaltsverzeichnisse der Bände AA 14-19 zum Korpus.
- *Abteilung 3, AA 20-23 (von Gerhard Lehmann herausgegebener Teil des handschriftlichen Nachlasses):* Zum Korpus gehören alle Texte Kants aus den Bänden AA 20-23, inklusive der Herausgeberanmerkungen, die in diese Texte integriert sind (d.h. inklusive der Anmerkungen, die nicht als Fußnoten abgedruckt sind). Ferner sind die Fußnoten des Herausgebers aus den Bänden AA 20-22 vollständig und die Fußnoten aus AA 23 zu einem kleinen Teil vorhanden. Auch die Inhaltsverzeichnisse der Bände AA 20-23 gehören zum Kant-Korpus.
- *Bezugstexte:* Die Texte aus Kants handschriftlichen Nachlass beziehen sich z.T. auf eigene, bereits veröffentlichte Texte oder auf Texte fremder Autoren. Ein Teil dieser Texte ist in der Akademie-Ausgabe abgedruckt, nämlich erstens alle von Kant selbst verfassten Bezugstexte – Kant hat handschriftliche Notizen in seinen Handexemplaren von *Das Gefühl des Schönen und Erhabenen* (1764), *Die Kritik der reinen*

ausschließlich das Lessing-Gedicht vorhanden. Die Erklärung, was es mit dem Gedicht auf sich hat, fehlt.

⁶ Die Notizen sind von 1 bis 8112 durchnummeriert. Der Unterschied zwischen der Anzahl der Notizen – 8141 – und der letzten Notiznummer – 8112 – kommt dadurch zustande, dass einzelne Notiznummern mehrfach vergeben und durch einen Buchstaben ergänzt wurden. Beispiele sind die Notizen 158 und 158a in AA 15.

⁷ Es handelt sich dabei sehr häufig um Angaben, wie die einzelnen Textstellen der Notizen in der Original-Handschrift angeordnet sind. Beispielsweise werden in den Reflexionen zur Mathematik aus Band AA 14 (S. 1-61) die relativen Positionen von Formeln und Skizzen durch den Herausgeber spezifiziert.

Vernunft (1781) und *Die Kritik der praktischen Vernunft* (1788) hinterlassen – und zweitens die folgenden Bezugstexte fremder Autoren:⁸

- *Gottfried Achenwall: Juris naturalis pars posterior complectens jus familiae, jus publicum et jus gentium. Editio quinta emendatior, Göttingen 1763.*
Der Text in in AA 19 nur teilweise abgedruckt. Das elektronische Kant-Korpus enthält nur die in AA 19 abgedruckten Teile.
- *Alexander Gottlieb Baumgarten, Metaphysica, Ed. III, 1757.*
Der Text ist verteilt auf die Bände AA 14, 15 und 17 vollständig abgedruckt. Er ist auch vollständig im elektronischen Korpus enthalten.
- *Alexander Gottlieb Baumgarten: Initia Philosophiae Practicae, Halle 1760.*
Der Text ist vollständig abgedruckt in AA 19 und vollständig im elektronischen Korpus enthalten.
- *Friedrich Bouterwek: Rezension von Kants Metaphysische Anfangsgründe der Rechtslehre, Göttingische Anzeigen von gelehrten Sachen unter der Aufsicht der Königl. Gesellschaft der Wissenschaften 18.02.1797.*
Der Text ist vollständig abgedruckt in AA 20 und vollständig im elektronischen Korpus enthalten.
- *Johan August Eberhard: Vorbereitung zur natürlichen Theologie zum Gebrauch akademischer Vorlesungen, Halle 1781.*
Der Text ist vollständig abgedruckt in AA 18 und vollständig im elektronischen Korpus enthalten.
- *Johann Gottlieb Kreutzfeld: Dissertatio Philologico-Poetica de Principiis Fictionum Generalioribus, zweite Dissertation vor Antritt zur Professur der Dichtkunst an der Universität Königsberg, verteidigt am 28.02.1777.*
Der Text ist vollständig abgedruckt in AA 15 und vollständig im elektronischen Korpus enthalten.
- *George Friedrich Meier: Auszug aus der Vernunftlehre, Halle 1752.*
Der Text ist vollständig abgedruckt in AA 16 und vollständig im elektronischen Korpus enthalten.

⁸ Vgl. das Vorwort von Erick Adickes in AA 14.

- *Johannes Schulz: Rezension von Johann August Eberhard, Philosophisches Magazin Jenaer Litteraturzeitung 1790, 281-284. (abgedruckt in Cassirers Kant-Ausgabe, Bd. VI (Berlin 1914, 73-117))*

Der Text ist vollständig abgedruckt in AA 20 und vollständig im elektronischen Korpus enthalten.

- *Schüleraufsatz, unbetitelt*

Der Text ist vollständig (?) abgedruckt in AA 17 und ebenso vollständig im elektronischen Korpus enthalten.

- *Personenindex 1.Stufe*: Der von Katharina Holger, Eduard Gerresheim, Antje Lange und Jürgen Goetze bearbeitete und im Jahr 1969 bei de Gruyter erschienene Personenindex 1. Stufe zu Kants gesammelten Schriften (Abt. 1-3 der AA) gehört vollständig zum elektronischen Kant-Korpus.

Im Rahmen des von 2001-2005 DFG-geförderten Projekts zur *Bereitstellung und Pflege von Kants gesammelten Werken in elektronischer Form* wurden die schon vorhandenen Daten ergänzt. Der Schwerpunkt lag dabei auf den Daten des handschriftlichen Nachlasses (Abt. 3 der Akademie-Ausgabe).⁹ Insbesondere wurden die Notizköpfe der Notizen aus den Bänden AA 14-19, sämtliche Bezugstexte fremder Autoren und der Personenindex 1. Stufe neu aufgenommen. Außerdem wurden – soweit sie nicht schon vorhanden waren – die Herausgeberanmerkungen, die in der gedruckten Ausgabe im Text und nicht als Fußnoten erscheinen, ergänzt.

Die Bezugstexte wurden digitalisiert, indem sie eingescannt, mittels des OCR-Programms Abbyy Fine Reader 7.0 zeichenweise erkannt und schließlich in ein XML-Format transformiert wurden. Die Akademie-Ausgabe diente als Vorlage, weil in ihr die Bezugstexte mit lateinischen Lettern (Antiqua) abgedruckt sind. In den Originalausgaben sind die Bezugstexte in Frakturschrift gedruckt, Frakturschrift lässt sich mit dem Fine Reader 7.0 nicht verlässlich erkennen, ergo konnten die Originalausgaben nicht als Vorlagen verwendet werden. Der Achenwall-Text ist in der Akademie-Ausgabe nur auszugsweise abgedruckt; er konnte deshalb auch nur auszugsweise ins Korpus aufgenommen werden. Mittlerweile gibt es mit dem Abbyy Fine Reader XIX ein OCR-Programm, das Frakturschrift verlässlich erkennt.¹⁰

⁹ Die Daten der Abteilungen 1 und 2 (AA 1-13) brauchten nicht ergänzt zu werden. Auch ihre Bearbeitung hielt sich in Grenzen: Die Daten lagen bereits SGML-kodiert vor, sie konnten direkt nach XML transferiert werden. Allein Überschriften und tabellarische Strukturen mussten ‚von Hand‘ überarbeitet werden.

Mithilfe dieses Programms können die fehlenden Teile des Achenwall-Textes nachträglich digitalisiert werden.

Die Daten des handschriftlichen Nachlasses waren zu Projektbeginn einheitlich kodiert (vgl. [5] und [3]): In den Daten der Bände AA 14-16 und 23 wurde nicht zwischen Groß- und Kleinschreibung unterschieden, Umlaute, ‚ß‘ und andere Sonderzeichen waren nicht kodiert, und die Daten waren kaum mit Metainformation annotiert. In den Daten der Bände AA 17-22 wurde zwischen Groß- und Kleinschreibung unterschieden, Umlaute und ‚ß‘ waren kodiert, und die Daten waren in größerem Umfang mit Metainformation annotiert. Die Nachlassdaten wurden nun so überarbeitet, dass durchweg zwischen Groß- und Kleinschreibung unterschieden wird und Umlaute, ‚ß‘ und alle anderen Sonderzeichen gemäß UTF-8 kodiert sind. Außerdem wurden fehlende Metainformationen in Form von XML-Tags hinzugefügt, so dass die Daten jetzt weitestgehend ähnlich annotiert sind. (S.u., Abschnitt 1.2.)

Die Überarbeitung der Bände AA 14-23 wurde mithilfe eines entsprechend konfigurierten VIM-Editors durchgeführt,¹¹ regelmäßige Modifikationen und Anreicherungen wurden durch Perl-Scripte erledigt.

Das Korpus war zu Projektbeginn nicht fehlerfrei. Darüber hinaus ist es bei der elektronischen Ergänzung und Bearbeitung des Korpus vereinzelt zu neuen Fehlern gekommen. Die Fehler wurden weitgehend, wahrscheinlich aber nicht vollständig korrigiert. Es muss davon ausgegangen werden, dass das Korpus noch Fehler enthält; die auf dem Korpus basierende elektronische Edition der Schriften Kants kann die gedruckte Ausgabe deshalb nicht uneingeschränkt verlässlich ersetzen.

1.2 Kodierung

Das elektronische Kant-Korpus besteht aus folgenden Dateien:

1. (a) XML-Dateien für die Bände der Akademie-Ausgabe (vgl. Tabelle 1). In diesen Dateien sind alle Kanttext und – soweit vorhanden – die Herausgeberanmerkungen gespeichert.
- (b) JPG-Dateien für alle Abbildungen aus den Bänden AA 1-23.

¹⁰ Vgl. <http://www.frakturschrift.de>.

¹¹ Für häufig vorkommende Modifikationen und Ergänzungen wurden Tastenkombinationen definiert, so dass diese Bearbeitungen schnell und einfach durchgeführt werden konnten.

<i>Akademie-Ausgabe</i>	<i>Dateien</i>	<i>DTD</i>
Abt. 1 (AA 1-9)	1.xml - 9.xml	kant.abt1.dtd
Abt. 2 (AA 10-13)	10.xml - 13.xml	kant.abt2.dtd
Abt. 3 (AA 14-19, Adickes-Bände)	14.xml - 19.xml	kant.abt3_1.dtd
Abt. 3 (AA 20-23, Lehmann-Bände)	20.xml - 23.xml	kant.abt3_2.dtd

Tabelle 1: Bände der Akademie-Ausgabe

2. XML-Dateien für die Inhaltsverzeichnisse der Bände AA 1-23 (vgl. Tabelle 2).
3. XML-, respektive HTML-Dateien für die Bezugstexte fremder Autoren (vgl. Tabelle 3).
4. HTML-Dateien für den Personenindex 1. Stufe.

Die Zeichen der XML- und HTML-Dateien sind gemäß UTF-8 kodiert; auf alle Nicht-ASCII-Zeichen wird mittels Gegenstandreferenzen (*entity references*) verwiesen.

1.2.1 Kodierung der Einzelbände (1.xml-23.xml)

In den Daten der Einzelbände sind die folgenden Textelemente kodiert:

- Überschriften, Paragraphen
- Seitenanfänge, Zeilenanfänge: Ein Paragraph kann auf einer Seite anfangen und auf einer anderen enden; ebenso kann z.B. eine fremdsprachliche Textstelle (s.u.) in einer Zeile beginnen und erst in einer späteren Zeile enden. Es ist mit XML nicht möglich, überlappende Textelemente zu kodieren,¹² wir haben deshalb Seiten und Zeilen als leere Elemente kodiert. Es wird jeweils der Seiten- respektive Zeilenanfang zusammen mit der jeweiligen Seiten- bzw. Zeilennummer markiert. Ein Paragraph kann einen neuen Seitenanfang enthalten, und eine fremdsprachliche Textstelle einen neuen Zeilenanfang enthalten; die verschiedenen Textelemente überlappen nicht.
- Randtexte Kants (nur in AA 1), Fußnoten Kants, Fußnotenreferenzen (das sind Verweise auf Fußnoten im Haupttext, also z.B. hochgestellte Sterne oder Ziffern)

¹² Mit SGML können solche Elemente als *concurrent structures* kodiert werden. Vgl. [1].

- Verweise auf Bilder im Text: Bilder sind separat als JPG-Dateien gespeichert (s.o.).
 - Eigennamen
 - fremdsprachliche Textstellen: Nicht-deutschsprachige Textstellen sind jeweils mit Angabe der Sprache (Latein, Griechisch, Französisch, Niederländisch, ...) und ggf. des verwendeten Alphabets – griechische Textstellen können entweder in griechischen Buchstaben oder in lateinischer Umschrift stehen – markiert.
 - mathematische Formeln
 - Worttrennungen: Wörter, die in der gedruckten Ausgabe getrennt sind, werden im elektronischen Korpus stets vollständig der Zeile zugeordnet, in der sie beginnen. Die Stelle, an der das jeweilige Wort in der gedruckten Ausgabe getrennt ist, ist markiert. Bsp. AA 1, Seite 5, Zeile 4: `Hochedelge<trenn/>bornen`
 - Hervorhebungen und Besonderheiten im Drucksatz: Antiqua, Fettdruck, Sperrdruck, Hochstellung, Petitdruck etc.
 - – Briefe, öffentliche Erklärungen, handschriftliche Erklärungen, Gedichte, Denkverse, Stammbucheinträge, amtliche Erklärungen
 - Beilagen zu Briefen, Datumsangaben(Nur in Bänden, in denen Briefe etc abgedruckt sind, d.h. in den Bänden AA 10-13 und 23.)
 - – handschriftliche Notizen, inklusive der Notizköpfe
 - innerhalb der Notizköpfe: Notiznummern, Datierungen, Fundorte und sonstige Anmerkungen
- (Nur in den von Adickes herausgegebenen Bänden des handschriftlichen Nachlasses AA 14-19.)
- Einschübe Kants: In den von Adickes herausgegebenen Bänden des handschriftlichen Nachlasses (AA 14-19) sind Zusätze, die Kant zu seinen Aufzeichnungen gemacht hat, durch Klammern und ein hochgestelltes ‚g‘ (für ‚gleichzeitig‘) oder ‚s‘ (für ‚später‘) markiert: (*g/s* ...). Diese Markierungen wurden übernommen und nach XML übersetzt: `<zusatz typ="gleichzeitig"/"später"> ... </zusatz>`.
- implizite Wiederholungen von Textstellen, z.B. mithilfe von Auslassungszeichen: Wir geben ein Beispiel in Abschnitt [2.2.1](#).

- tabellarische Strukturen: HTML-ähnlich kodierte tabellarische Strukturen werden verwendet, um besondere Anordnung von Textstücken zu markieren. Folgende Beispiele dienen der Veranschaulichung:

- In der *Kritik der reinen Vernunft* steht die *Tafel der Kategorien* (vgl. AA 3, Seite 93 und AA 4, Seite 66), welche abgekürzt folgendermaßen aussieht:

1. Der Quantität	
2. Der Qualität	3. Der Relation
4. Der Modalität	

Im elektronischen Kant-Korpus ist die Tafel als Tabelle mit neun Feldern kodiert. Die Tabelle hat drei Spalten und drei Tabellenzeilen; eine Tabellenzeile kann dabei mehr als eine Textzeile enthalten. In der Beispieltabelle sind die ersten und dritten Felder der ersten und dritten Zeile und das zweite Feld der zweiten Zeile leer:

1. Der Quantität		
2. Der Qualität		3. Der Relation
4. Der Modalität		

- In Abschlüssen von Briefen stehen zumeist links Ort und Datum und rechts eine Abschlussformel. Wir kodieren solche Abschlüsse als Tabellen mit mindestens zwei Spalten. Vgl. diesbezüglich den Abschluss von Brief Nr. 5 (AA 10, Seite 3) – der Ort und das Datum stehen in der linken Tabellenspalte, die Abschlussformel und der Name stehen in der rechten Tabellenspalte:

	Ew: Königl: Majestät
Koenigsberg	allerunterthänigster Knecht
d. 8ten April. 1756.	Immanuel Kant.

- In den von Adickes herausgegebenen Bänden des handschriftlichen Nachlasses (AA 14-19) wird an zahlreichen Stellen versucht, den Aufbau handschriftlicher Notizen im Drucksatz nachzuvollziehen. So kommen im Drucksatz Muster wie das folgende vor, demzufolge in der Originalhandschrift der XX-Text wohl um den yy-Text herum geschrieben ist:¹³

¹³ Vgl. z.B. AA 15, Seite 494 (Notiz 1111).

```

XX XX XX XX XX XX XX
                XX XX XX
yy yy yy      XX XX XX
yy yy yy      XX XX XX
yy yy yy      XX XX XX

```

Derartige Muster können als tabellarische Strukturen kodiert werden:

	XX XX XX XX XX XX XX
	XX XX XX
yy yy yy	XX XX XX
yy yy yy	XX XX XX
yy yy yy	XX XX XX

Dabei geht allerdings die Information, dass der XX-Text über dem yy-Text beginnt, verloren. Grundsätzlich lässt sich diese Information leicht in die Kodierung aufnehmen, indem man nämlich die Tabelle in zwei Tabellenzeilen aufteilt und die erste (lange) Zeile über beide Spalten reichen lässt:

XX XX XX XX XX XX XX	
	XX XX XX
yy yy yy	XX XX XX
yy yy yy	XX XX XX
yy yy yy	XX XX XX

Das Problem dieser Kodierung ist, dass der XX-Text in zwei Teile geteilt und dadurch seine Einheit aufgehoben wird. Wir haben uns – um diesem Problem zu entgehen – für die davor vorgeschlagene Kodierung entschieden, derzufolge der XX-Text und der yy-Text jeweils zusammenhängend in einer Tabellenspalte stehen. Die Information, dass die erste Zeile des XX-Texts im Druckbild über dem yy-Text beginnt, haben wir als Kommentar hinzugefügt.

In den Bänden des handschriftlichen Nachlasses (AA 14-23) gibt es recht häufig Stellen, deren Druckbild nur mit großem Aufwand und nicht ohne Nachteile in XML rekonstruiert werden kann. Wir haben uns entschieden, diese Stellen stets vereinfacht in XML zu kodieren und durch Anmerkungen zu ergänzen.¹⁴

- Anmerkungen und Textzusätze der Herausgeber: Es handelt sich dabei um Anmerkungen, die im Text oder als Fußnoten erscheinen. (Soweit überhaupt vorhanden nur in den Bänden des handschriftlichen Nachlasses AA 14-23.) Außerdem sind Ergänzungen unvollständiger Worte durch die Herausgeber markiert. (Auch in den Briefbänden 10-13. Bsp. AA 10, Seite 1, Zeile 17: Kräftenschätz<hg_zusatz>ung</hg_zusatz>¹⁵) Schließlich sind in den Bänden AA 20-22 fehlende Satzzeichen – zu meist Punkte, die eigentlich am Ende eines Satzes stehen sollten – markiert.
- In den Bänden des Opus postumum (AA 21, 22) sind die Convolute markiert. Außerdem sind Verweise auf Faksimileseiten gesetzt. (S.u., Abschnitt 2.3.)

Die Daten des Kant-Korpus haben wesentliche Unterschiede: So kommen beispielsweise Briefe nur in den Bänden AA 10-13 und 23 vor, und die Einteilung in kurze, handschriftliche Notizen gilt nur für die Bände AA 14-19. Die Daten sind durchgehend ähnlich kodiert; wegen ihrer Unterschiede war es aber sinnvoll, verschiedene *Document Type Definitions* (DTDs) zu spezifizieren. Wir haben eigene Dokumenttypen für die Bände der 1. Abteilung (AA 1-9, *kant.abt1.dtd*), für die Bände der 2. Abteilung (AA 10-13, *kant.abt2.dtd*), für die Adickes-Bände der 3. Abteilung (AA 14-19, *kant.abt3_1.dtd*) und für die Lehmann-Bände der 3. Abteilung (AA 20-23, *kant.abt3_2.dtd*) definiert. Die kommentierten DTDs des Kant-Korpus werden nicht eigens veröffentlicht, können aber auf Anfrage hin zur Verfügung gestellt werden.

1.2.2 Kodierung der Inhaltsverzeichnisse

Die Inhaltsverzeichnisse der Bände AA 1-23 wurden XML-kodiert. Die Inhaltsverzeichnisse der Briefbände AA 10-12 weichen in ihrer Struktur von den Inhaltsverzeichnissen der anderen Bände AA 1-9 und 13-23 ab. Wir ha-

¹⁴ Anmerkungen wurden als als XML-Kommentare oder als Werte eines fast universell zu vergebenden ‚anm‘-Attributs eingefügt. Wenn an einer Problemstelle kein Tag vorhanden war, dem sinnvoll eine Anmerkung hinzugefügt werden konnte, dann wurde ein eigens für diesen Zweck definiertes, leeres Problem-Element eingefügt. Vgl. z.B. AA 19, Seite 343: <problem anm=“In den Notizen 7357-7360 ist der Notiztext nicht durch einen eigenen Paragraphen vom Notizkopf getrennt. Zur Wahrung der Struktur ist eine Trennung aber noetig.\”>.

¹⁵ In den Originaldaten sind die Umlaute (,ä‘) durch Gegenstandreferenzen ersetzt. Diese Referenzen sind im Beispiel zugunsten der Lesbarkeit aufgelöst.

<i>Akademie-Ausgabe</i>	<i>Dateien</i>	<i>DTD</i>
AA 1-9	Inhalt1.xml - Inhalt9.xml	Inhalt1.dtd
AA10-12	Inhalt10.xml - Inhalt12.xml	Inhalt2.dtd
AA13-23	Inhalt13.xml - Inhalt23.xml	Inhalt1.dtd

Tabelle 2: Inhaltsverzeichnisse

ben deshalb zwei DTDs erstellt, hinsichtlich derer die XML-Dateien gültig sind (s. Tabelle 2).

Folgende Elemente sind in den XML-Dateien annotiert:

- Titel und Überschriften der verzeichneten Texte und Kapitel zusammen mit den entsprechenden Seitenzahlen. Die Hierarchie der Einträge (Kapitel – Unterkapitel – ...) wird dabei berücksichtigt.
- Inhaltsverzeichnis-interne Überschriften
- Fettdruck
- In Bänden des handschriftlichen Nachlasses: Verweise auf Schriftphasen und Fundstellen
- In den Briefbänden: Briefnummern, Datierungen, Empfänger respektive Absender

Wir stellen die kommentierten DTDs der Inhaltsverzeichnisse auf Anfrage gerne zur Verfügung.

1.2.3 Kodierung der Bezugstexte

Für die Bezugstexte wurde eine einheitlicher Dokumenttyp definiert (s. Tabelle 3). Folgende Merkmale der Texte sind entsprechend der Bezugstext-DTD annotiert:

- Überschriften, Paragraphen, Zeilenwechsel
- Seiten der Originalausgaben, Bände und Seiten, auf denen der Text in der Akademie-Ausgabe gedruckt ist (Seiten sind als *Seitenwechsel* kodiert, s.o.)
- Fußnoten, Inhaltsverzeichnisse, Indizes
- Hervorhebungen durch: Kapitälchen, Fettdruck, Kursivdruck, Sperrdruck, Zentrierung, ...

<i>Bezugstext</i>	<i>Datei</i>	<i>DTD</i>
Achenwall: Juris naturalis	achenwall.xml	bezug.dtd
Baumgarten: Metaphysica	baumgarten-metaphysica.xml	bezug.dtd
Baumgarten: Initia	baumgarten-initia.xml	bezug.dtd
Bouterwek: Rezension	bouterwek.xml	bezug.dtd
Eberhard: Vorbereitung	eberhard.xml	bezug.dtd
Kreutzfeld: Dissertatio	kreutzfeld.xml	bezug.dtd
Meier: Vernunftlehre	meyer.html	—
Schulz: Rezension	schulz.xml	bezug.dtd
Schüleraufsatz	schueleraufsatz.xml	bezug.dtd

Tabelle 3: Bezugstexte

- griechische und hebräische Textstellen
- Worttrennungen

Meyers *Auszug aus der Vernunftlehre* liegt nur HTML-kodiert vor. Für die entsprechende Datei gibt es also keine relevante DTD. Die kommentierte DTD für die anderen Bezugstexte wird auf Anfrage gerne zur Verfügung gestellt.

1.2.4 Kodierung des Personenindex 1.Stufe

Für den Personenindex 1. Stufe wurde keine eigene XML-Kodierung vereinbart. Der Index wurde, nachdem er gescannt und mithilfe des Fine Reader 7.0 automatisch erkannt wurde, direkt in HTML kodiert. Für jeden Anfangsbuchstaben der Indexeinträge wurde eine eigene HTML-Datei erzeugt: PI-A.html, PI-B.html, ..., PI-Z.html. Die Anleitung zur Benutzung des Index wurde unter dem Namen „Anleitung.html“ gespeichert.

2 Verwendung des Korpus

Das Bonner Kant-Korpus ist die zentrale Datenbasis für eine elektronische Edition der Werke Immanuel Kants. Ausgehend von dem Korpus wurden erstens verschiedene Darstellungen der Kanttexte erzeugt und über das WWW zugänglich gemacht (vgl. Abschnitt 2.1). Zweitens wurden die Texte durch Installation einer Suchmaschine und die Bereitstellung von Indizes durchsuchbar gemacht (vgl. Abschnitt 2.2). Drittens wurden aufeinander Bezug nehmende Stellen der Kanttexte durch Hyperlinks verknüpft;

er wurde dadurch begonnen, die Texte zu einem umfassenden Hypertext zu verbinden. (vgl. Abschnitt 2.3). Viertens und letztens wurden Korpus-Daten in verschiedenen Formaten für verschiedene Zwecke bereitgestellt; mithilfe des Bonner Korpus wurden dadurch u.a. philologische Arbeiten am Werk Kants unterstützt (vgl. Abschnitt 2.4).

2.1 Verschiedene Darstellungen

Der erste Zweck des Kant-Korpus besteht darin, die Daten über das WWW zugänglich und lesbar zu machen. Es wurden deshalb verschiedene Darstellungen der Korpusdaten erzeugt.

2.1.1 Druckähnliche Version der Bände AA 1-23

Die XML-kodierten Bände AA 1-23 wurden seiten- und zeilengetreu gemäß der Druckausgabe nach HTML transferiert. Für jede Druckseite, die Text enthält, der auch im elektronischen Korpus enthalten ist, wurde eine eigene HTML-Seite erzeugt. Für leere Seiten und für Seiten, die ausschließlich Herausgeberanmerkungen enthalten und deren Inhalt nicht Teil des elektronischen Korpus ist, wurden keine HTML-Seiten erzeugt.

Die Bezugstexte aus den Nachlassbänden sind separat gespeichert; sie befinden sich – anders als in der gedruckten Ausgabe – nicht unter den Kanttexten. Druckseiten, die ausschließlich einen Bezugstexte enthalten, wurden wie leere Seiten behandelt, d.h. für sie wurden keine druckähnlichen HTML-Seiten erstellt.

Fußnoten stehen an den Seitenenden, an denen sie auch in der gedruckten Ausgabe stehen. Es wurden nur die Fußnoten der Bände AA 20-22 in die HTML-Edition aufgenommen. Die Fußnoten der anderen Bände wurden nicht in die HTML-Edition aufgenommen, weil sie entweder nur teilweise (AA 17, 18, 23) oder gar nicht (alle anderen Bände) Teil des elektronischen Korpus sind.¹⁶

Die druckähnlichen HTML-Seiten sind als Tabellen aufgebaut. In der von links aus gesehen ersten Tabellenspalte stehen die Zeilennummern des nachfolgenden Textes. In der zweiten Tabellenspalte steht der Text der Akademie-Ausgabe. Es handelt sich dabei üblicherweise um einen Text Kants.¹⁷ Die dritte Tabellenspalte ist vorgesehen für Verknüpfungen zu

¹⁶ S.o., Abschnitt 1.1.

¹⁷ In der Spalte stehen natürlich auch an Kant gerichtete Briefe fremder Verfasser, Überschriften der Herausgeber etc.

verwandten Textstellen oder Faksimiles der Originaltexte etc. So sind beispielsweise die Kapitel des Werks *Antropologie in pragmatischer Hinsicht* (AA 7) mit den entsprechenden Notizen aus AA 14 verknüpft, und den Seiten des Opus Postumum (AA 21, 22) wurden Verweise auf die entsprechenden Faksimiles hinzugefügt (s.u., Abschnitt 2.3).

Wörter, die in der gedruckten Ausgabe durch einen Zeilenumbruch getrennt sind, werden in der elektronischen Edition nicht getrennt. Sie stehen vollständig in der Zeile, in der sie in der gedruckten Ausgabe beginnen.

Deutschsprachige Texte von Kant sind in der Akademie-Ausgabe in Fraktur gedruckt; für fremdsprachliche – z.B. lateinische oder französische – Texte und Textstellen werden lateinische Lettern (Antiqua) verwendet. In der elektronischen Edition werden auch die deutschen Texte in lateinischen Lettern dargestellt.

In der elektronischen Edition sind Eigennamen und Textstellen, die in der gedruckten Ausgabe kursiv oder in Antiqua gesetzt sind, kursiv gesetzt. Zu den Textstellen, die in der gedruckten Ausgabe in Antiqua gesetzt sind und die also in der elektronischen Ausgabe kursiv wiedergegeben werden, gehören u.a. alle fremdsprachlichen Stellen. Fett oder gesperrt sind solche Textstellen, die auch in der gedruckten Version fett oder gesperrt sind.

Texte Kants und an Kant gerichtete Briefe stehen in einer Serifen-Schrift (Times) in schwarz; Anmerkungen und Überschriften der Herausgeber stehen in einer Serifen-losen Schrift (Arial, Helvetica) in grau. Verknüpfungen sind gelb.

Die Inhaltsverzeichnisse der Bände AA 1-23 wurden nach HTML transferiert; darüber hinaus wurde ein Gesamtverzeichnis für die Bände erstellt: <http://www.ikp.uni-bonn.de/kant/verzeichnisse-gesamt.html>. Das Gesamtverzeichnis ist mit den Inhaltsverzeichnissen der Einzelbände verknüpft, und die Inhaltsverzeichnisse sind mit den Textseiten der jeweiligen Bände verknüpft, so dass es möglich ist, über das Gesamtverzeichnis und die Inhaltsverzeichnisse auf die Textseiten zuzugreifen. Jede Textseite ist außerdem mit der ihr vorhergehenden und der nachfolgenden Seite verknüpft, so dass gleichsam ‚geblättert‘ werden kann.

2.1.2 Alternative Anordnungen

Die Daten des Kant-Korpus müssen nicht ausschließlich seitenweise, ähnlich wie in der gedruckten Ausgabe dargestellt werden, sondern können auch in alternativen Anordnungen zugänglich gemacht werden:

- *Briefwechsel*: Kants Briefwechsel liegt elektronisch in zwei Varianten vor. Die erste Variante ist an der gedruckten Version der Akademie-

Ausgabe orientiert, wobei einer Druckseite eine HTML-Seite entspricht (s.o.). Ein Brief kann sich demnach über mehrere HTML-Seiten erstrecken; auf derselben HTML-Seite kann ein Brief enden und ein anderer anfangen. In der zweiten Variante wurde für jeden Brief genau eine HTML-Seite erzeugt. Diese Variante weicht von der gedruckten Ausgabe ab. Für sie wurde ein eigenes Gesamtverzeichnis des Briefwechsels erstellt: <http://www.ikp.uni-bonn.de/kant/briefe>.

Für die HTML-Kodierung der Einzelbriefe gelten die gleichen Grundsätze wie für die druckähnliche Kodierung. Seiten- und Zeilenkennungen sind allerdings nicht vorhanden; zum Zitieren muss auf die jeweils verknüpfte, druckähnliche Version rekurriert werden.

- *Sortierung der Notizen aus den Bänden AA 14-19 nach ihren Fundorten:* Die von Erich Adickes herausgegebenen Notizen Kants (AA 14-19) sind in der Akademie-Ausgabe thematisch angeordnet. Zusammen mit jeder Notiz ist ihre Fundstelle angegeben. Aufgrund dieser Angaben können die Notizen hinsichtlich ihrer Fundstellen sortiert werden. Beispielsweise können die Notizen, welche Kant in seinem Handexemplar von Baumgartens *Metaphysica* hinterlassen hat, zusammengefasst und in eine Teilordnung gebracht werden:¹⁸ So können Notizen, die Kant auf die Seite 131 seines Exemplars der *Metaphysica* gemacht hat, vor Notizen, die Kant auf die Seite 141 gemacht hat, gesetzt werden. Notizen allerdings, die auf der selben Seite des Handexemplars stehen – z.B. alle Notizen von Seite 131 – sind zumeist nicht in eine begründete Ordnung zu bringen. (Man bräuchte dazu mehr Information, als durch die Herausgeberanmerkungen gegeben ist.)

Die Sortierung nach Fundorten wurde begonnen, aber nicht abgeschlossen. Ein nach Fundorten sortiertes Verzeichnis der Notizen aus den Bänden 14-19 ist deshalb noch nicht verfügbar.

- *Chronologische Sortierung der Notizen aus den Bänden AA 14-19:* Die chronologische Sortierung der Notizen aus den Bänden AA 14-19 gemäß den von Adickes vergebenen Datierungen ist – anders als die Sortierung nach Fundorten – nicht praktikabel: Für die meisten der 8141 Notizen sind mehrere mögliche Datierungen angegeben, so dass die Anzahl der auf den alternativen Datierungen basierenden Anordnungen unüberschaubar groß wäre. Wir verfügen über keine Kriterien,

¹⁸ Fast die Hälfte, nämlich 4022 der 8141 von Erich Adickes herausgegebenen Notizen stammen aus Kants Handexemplar der *Metaphysica*. Die Notizen sind über die Bände AA 14-18 verteilt; nur in Band AA 19 ist keine Notiz aus der *Metaphysica* enthalten.

um die Menge der möglichen Anordnungen auf ein überschaubares Maß zu beschränken. Chronologischen Anordnungen wurden deshalb nicht erzeugt.

2.1.3 Rekonstruktion von Aufzeichnungen Kants

Aufgrund der Herausgeberanmerkungen zu den Nachlassdaten lassen sich die Aufbauten einzelner Aufzeichnungen Kants rekonstruieren. Versuchsweise wurde ein Foliobogen mit einer Reflexion zur Medizin (Notiz 1526, AA 15.2, 939-953) rekonstruiert. Die Rekonstruktion liegt als PDF- und als PS-Datei vor: <http://www.ikp.uni-bonn.de/kant/rekonstruktion>.

2.2 Durchsuchbarkeit

Der zweite Zweck des Kant-Korpus besteht darin, den Zugang zu den Daten zu optimieren und ihre effektive Durchsuchbarkeit herzustellen.

2.2.1 Suchmaschine

Zur Durchsuchung der Kanttexte wurde *webglimpse*¹⁹ installiert. Die Suchmaschine operiert auf den druckähnlichen HTML-Dateien der Bände AA 1-23. Sie erlaubt komplexe Suchausdrücke – ist also funktionsmächtig –, sie ist selbst erklärend und sehr einfach zu bedienen,²⁰ und sie funktioniert außerordentlich robust.

Ein weiterer Vorteil besteht darin, dass *webglimpse* äußerst einfach zu warten ist: Es ist leicht und ohne aufwändige Einarbeitung möglich, die Datenbasis der Suchmaschine zu modifizieren (z.B. durch Korrekturen oder Ergänzungen neuer Texte), den Suchindex zu aktualisieren und dadurch dauerhaft das richtige Funktionieren zu gewährleisten. Solange mit Änderungen der Datenbasis zu rechnen ist und außerdem diejenigen Personen, welche mit der Pflege der elektronischen Kanttexte beschäftigt sind, häufig wechseln, ist dieser Vorteil hoch zu schätzen.

Nichtsdestotrotz ist es sinnvoll, mittelfristig *webglimpse* durch eine Suchmaschine, die direkt auf den XML-Daten operiert, zu ersetzen. Daraus, dass die derzeitige Suchmaschine auf den druckähnlichen HTML-Dateien operiert, ergeben sich nämlich drei Nachteile: Erstens werden komplexe

¹⁹ <http://webglimpse.net>

²⁰ Eine englischsprachige Anleitung zum Gebrauch der Suchmaschine befindet sich unter: <http://webglimpse.net/docs/tips.html>.

Phrasen nur dann gefunden, wenn sie vollständig in einer Zeile stehen. Beispiel: In AA 1 finden wir auf der Seite 5 (Zeilen 5 f) die Phrase „Merkmale der Gütigkeit“; der Zeilenumbruch liegt vor „Gütigkeit“. Mithilfe der Suchmaschine finden wir zwar „Merkmale der“, aber nicht „Merkmale der Gütigkeit“. Um „Merkmale der Gütigkeit“ zu finden, müssen wir nach zwei Phrasen suchen, nämlich nach „Merkmale der“ (Zeile 5) und nach „Gütigkeit“ (Zeile 6). Die erfolgreiche Suchanfrage lautet: „Merkmale der AND Gütigkeit“. Es ist generell eine gute Suchstrategie, die zentralen Inhaltswörter einer langen Anfrage zu extrahieren und für die Suche mit dem Junktor „AND“ zu verknüpfen: „Merkmale AND Gütigkeit“.

Der zweite Nachteil der Operation auf den druckähnlichen HTML-Dateien besteht darin, dass Phrasen, die von einem Seitenumbruch unterbrochen werden, nicht gefunden werden. Beispiel: In AA 1 finden wir auf der Seite 7 f die Phrase „Mäßigung von dieser Art“; der Seitenumbruch ist nach „Mäßigung“. Der eben gemachte Vorschlag, die Anfrage „Mäßigung AND Art“ zu stellen, führt hier nicht zum Erfolg: „Mäßigung“ und „Art“ stehen auf verschiedenen Seiten; die Seite 7, auf der das Wort „Art“ nicht vorkommt, wird nicht gefunden. Wir können vorläufig zur Kompensation dieses Problems nur vorschlagen, kurze Suchanfragen zu stellen: Einzelwörter sind in den HTML-Daten nie getrennt (s.o., Abschnitt 2.1.1). Sie werden deshalb immer gefunden.

Der dritte Nachteil besteht darin, dass die HTML-Daten semantisch ärmer sind als die XML-Daten. Sehen wir uns als Beispiel die Seite 499 von Band AA 15 an. Dort stehen die folgenden beiden Zeilen, die so auch in der HTML-Version der Seite erscheinen:

```
Der Wille nach instincten ist das temperament,
```

```
-- -- -- Grundsätzen -- character.
```

Wenn man den Band AA 15 nach der Zeichenkette „Der Wille nach Grundsätzen ist character“ durchsucht, wird man kein Suchergebnis erhalten. Es wäre allerdings wünschenswert, die Seite 499 als Suchergebnis zu erhalten, da der gesuchte Text dort wenigstens implizit steht. Man erhielte die Seite als Suchergebnis, wenn bei der Bearbeitung der Suchanfrage Information darüber vorläge, durch welche Zeichenketten – nämlich „Der Wille nach“ und „ist“ – die Auslassungsstriche vor und nach „Grundsätze“ zu ersetzen sind. Diese Information ist in den XML-Daten vorhanden:

```

<zeile nr="02"/> Der Wille nach instincten ist das temperament,
<zeile nr="03"/> <wdh><bezug>Der Wille nach</bezug>
                  <text>&emdash; &emdash; &emdash;</text></wdh>
                  Grundsätzen
                  <wdh><bezug>ist</bezug><text>&emdash;</text></wdh>
                  character.

```

Die impliziten Wiederholungen von „Der Wille nach“ und „ist“ sind mit Angaben des Bezugstextes (als `bezug` markiert) und des Ersatztextes (als `text` markiert)²¹ kodiert. Man kann die Zeile 3 so darstellen, dass – wie in der druckähnlichen HTML-Version – nur die als `text` markierten Auslassungszeichen und nicht deren Bezugstext erscheinen. Gleichzeitig kann man zur Durchsuchung der Zeile auf den Bezugstext rekurrieren. Die Seite 499 kann also als Ergebnis der Suchanfrage „Der Wille nach Grundsätzen ist character“ gefunden werden.

Solange die Suchmaschine auf den druckähnlichen HTML-Dateien operiert und ausgelassener Text nicht durchsucht werden kann, empfiehlt es sich, die schon oben vorgeschlagene Strategie zu verfolgen und nur nach zentralen Inhaltswörtern zu suchen: Die Suchanfrage „Grundsatz AND Character“ (mit der Angabe, dass das Suchergebnis nur teilweise übereinzustimmen braucht) wird für das gesamte Kant-Korpus – d.h. die Bände AA 1-23 – durch Angabe von nur 23 Seiten beantwortet. Das Suchergebnis ist also leicht überschaubar, und es ist deshalb nicht zwingend notwendig, den ausgelassenen Text in der Suchanfrage zu berücksichtigen.²²

2.2.2 Indizes

Außer durch eine Suchmaschine kann das Kant-Korpus durch Indizes erschlossen werden:

- Der 1969 erschienene Personenindex 1. Stufe wurde digitalisiert und in

²¹ Die Gegenstandsreferenz „&emdash;“ referiert auf das Zeichen ‚–‘.

²² Wenn man das Wort „Wille“ in die Anfrage mit aufnimmt – „Wille AND Grundsatz AND Character“ –, dann sinkt die Anzahl der Treffer auf nur drei Seiten. Die Seite 499 von AA 15 wird gefunden, da ja alle drei Wörter auf dieser Seite vorkommen. Wenn man noch angibt, dass die Suchanfrage einen Rechtschreibfehler enthalten darf – damit außer „Character“ (mit „c“) auch „Charakter“ (mit „k“) gefunden werden kann –, dann steigt die Anzahl der Treffer auf 55 Seiten. Das Suchergebnis ist damit immer noch überschaubar. Es stellt sich die Frage, ob eine mächtigere Suche als die nach zentralen Inhaltswörtern überhaupt gebraucht wird.

HTML kodiert (s.o., Abschnitt 1.1). Bei dem Personenindex 1. Stufe handelt es sich um einen Index aller Personen, die Kant – nach Ausweis der Akademie-Ausgabe – erwähnt hat. Die Personennamen sind alphabetisch geordnet; mit ihnen sind alle Stellen, an denen die jeweiligen Personen genannt werden, angegeben. Die Stellenangaben sind wiederum durch Hyperlinks mit den entsprechenden Stellen in der druckähnlichen HTML-Version des Korpus verknüpft. Die elektronische Version des Personenindex 1. Stufe findet man unter: <http://www.ikp.uni-bonn.de/kant/persindex>.

- Der von Katharina Holger und Eduard Gerresheim bearbeitete Personenindex 2. Stufe ist ein bisher nur als Manuscript in 26 Heften vielfältigtes Register aller Stellen, an denen Kant eine Person erwähnt hat, wobei die jeweiligen Zitate, die zitierten Stellen aus den Originalwerken und bio- sowie bibliographische Angaben zu den zitierten Autoren ausgeführt sind. Mit Dr. Gerresheim wurde die Digitalisierung des Personenindex 2. Stufe und seine Aufnahme ins elektronische Kant-Korpus vereinbart. Die entsprechenden Arbeiten konnten allerdings noch nicht durchgeführt werden. Immerhin wurden einzelne Experimente zur Einbindung des Personenindex 2. Stufe in die elektronische Kant-Edition gemacht; Ergebnisse befinden sich unter: <http://www.ikp.uni-bonn.de/kant/pi2>.
- Schließlich wäre es attraktiv, über einen lemmatisierten Wortindex zu verfügen, so dass das Kant-Korpus nach Wörtern in verschiedenen Flexionsausprägungen und nicht nur nach spezifischen Wortformen durchsucht werden kann. (Für Suchanfragen mit *webglimpse* können in beschränkter Form reguläre Ausdrücke gebildet werden, so dass man sich bei der Suche nicht eindeutig auf bestimmte Wortformen festlegen muss. Die Vorzüge eines lemmatisierten Suchindizes werden dadurch aber nicht vollständig ausgeglichen.) Versuchsweise wurde von Dr. Gerd Willée mithilfe des Programms LEMMA4 eine lemmatisierte Form von Kants Schrift *Der Streit der Fakultäten* (AA 7) erzeugt. Das Ergebnis ist ein Vertikaltext, in dem in jeder Zeile je eine Wortform, ihre Grundform und ihre Wortklasse stehen; für Verben sind außerdem deren Flexionsausprägungen angegeben. Der lemmatisierte Text ist auf der Web-Seite der Kant-Edition verknüpft (<http://www.ikp.uni-bonn.de/kant/lemmatisiert>); das Format des Textes ist in [6] dokumentiert.

2.3 Verknüpfungen

Der dritte Zweck des Kant-Korpus besteht darin, die elektronische Kant-edition in einen Hypertext zu verwandeln, indem aufeinander Bezug nehmende Daten durch Hyperlinks verknüpft werden.

- *Verknüpfung thematisch verwandter Texte:* Texte mit direkten thematischen Bezügen können durch Hyperlinks verknüpft werden. Erstes Beispiel: In Band AA 7 befindet sich Kants Werk *Anthropologie in pragmatischer Hinsicht*, und in Band AA 15 befinden sich Kants handschriftliche Notizen zur Anthropologie. Die Notizen aus AA 15 sind in Anlehnung an das Inhaltsverzeichnis des Werkes aus AA 7 sortiert; sie können deshalb den Abschnitten des Werkes zugeordnet werden. Die Zuordnung wurde erreicht, indem die Notizen mit den ihnen entsprechenden Abschnitten des Werkes wechselseitig durch Hyperlinks verknüpft wurden. (Vgl. z.B. AA 15, Seite 55: <http://www.ikp.uni-bonn.de/kant/aa5/055.html>.)

Zweites Beispiel: In Band AA 20 befinden sich zusammenhängende Bemerkungen und lose Blätter zum Werk *Beobachtung über das Gefühl des Schönen und Erhabenen* (AA 2). Die entsprechenden Abschnitte aus AA 20 wurden mit dem Werk aus AA 2 verknüpft. (Vgl. AA 20, Seite 1: <http://www.ikp.uni-bonn.de/kant/aa20/001.html>.)

- *Verknüpfung handschriftlicher Aufzeichnungen mit ihren Bezugstexten:* Erläuterungen Kants zu anderen Texten können mit diesen Texten verknüpft werden. Beispiel: In Band AA 15 stehen Erläuterungen Kants zur *Psychologia empirica* in Baumgartens *Metaphysica*; rechts neben Kants Erläuterungen verweist ein Hyperlink auf den Bezugstext. (AA 15, Seite 3: <http://www.ikp.uni-bonn.de/aa15/003.html>.)²³
- *Verknüpfung handschriftlicher Aufzeichnungen und Briefe mit Bildern der Originalhandschriften respektive Originalbriefe:* Handschriften und Briefe können mit ihren Faksimiles verknüpft werden. Für das Opus postumum (AA 21 und 22) wurden solche Verknüpfungen bereits hergestellt: Die Originalhandschrift des Opus postumum wurde von der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW)

²³ Die Verknüpfungen der handschriftlichen Aufzeichnungen mit ihren Bezugstexten gilt es noch zu erweitern und zu optimieren. So sollen beispielsweise Erläuterungen Kants zu Baumgartens *Metaphysica* nicht nur mit dem Gesamtwerk *Metaphysica*, sondern präzise mit den Bezugsstellen innerhalb des Werks verknüpft werden.

vollständig eingescannt; die daraus resultierenden Bilder der einzelnen Manuscript-Seiten wurden als JPG-Dateien im WWW veröffentlicht. Die druckähnliche HTML-Version des Opus Postumum ist mit diesen Bildern verknüpft, so dass Nutzer des Kant-Korpus sich parallel zu Lehmanns Edition des Opus postumum (AA 21, 22) die entsprechenden Seiten des Original-Manuscripts anzeigen lassen können. (Z.B. AA 21, Seite 3: <http://www.ikp.uni-bonn.de/kant/aa21/003.html>.)

- *Verknüpfungen mit Anmerkungen neuer Editionen und mit Sekundärtexten:* Schließlich können die elektronischen Kanttexte mit externen Daten – etwa mit Anmerkungen und Ergänzungen, die im Zuge von Neueditionen gemacht werden – verknüpft werden. Für solche Verknüpfungen liegt noch kein Beispiel vor. Es wurden allerdings Vereinbarungen mit Dr. Tanya Gloyna und Jacqueline Karl von der BBAW darüber getroffen, dass Anmerkungen und Ergänzungen, die im Zuge der Neueditionen der kritischen Hauptwerke und des Opus postumum gemacht werden, ins elektronische Kant-Korpus integriert werden. (Vgl. diesbezüglich auch den folgenden Abschnitt 2.4.)

2.4 Bereitstellung und Sammlung von Daten

Der vierte Zweck des elektronischen Kant-Korpus besteht schließlich darin, als zentrale Datenbasis für die Kantforschung zu dienen. Einerseits werden Daten auf individuelle Anfragen hin in verschiedenen Formaten zur Verfügung gestellt, andererseits sollen neue Daten in das Korpus aufgenommen werden, d.h. das Korpus soll sukzessive ergänzt werden. So wurden für die Neueditionen der kritischen Hauptwerke und des Opus postumum Dr. Tanya Gloyna und Jacqueline Karl von der BBAW die entsprechenden Texte der Bände AA 3-5 und AA 21, 22 zur Verfügung gestellt. Es wurde vereinbart, dass Korrekturen und Ergänzungen, die im Zuge der Neueditionen anfallen, in das Korpus aufgenommen werden. Prof. Werner Stark hat für die Edition der Vorlesung über Moralphilosophie (2004 bei de Gruyter erschienen) die Daten der handschriftlichen Aufzeichnungen über Moralphilosophie (AA 19) erhalten. Die Daten wurden so formatiert, dass sie mit dem Programm *WordCruncher* verarbeitet werden konnte. Darüber hinaus wurden verschiedene kleinere Anfragen bedient, u.a. wurden dem Goethe-Institut (InterNationes) Kants Schrift *Zum Ewigen Frieden* (AA 8) für den Druck eines Posters überlassen.

Das elektronische Kant-Korpus kann in Zukunft sinnvollerweise durch Briefe, die nicht in die gedruckte Akademie-Ausgabe aufgenommen wurden, und durch neue Fundstücke des handschriftlichen Nachlasses ergänzt

werden. Ferner können Faksimiles aller noch verfügbaren Briefe und Handschriften eingefügt werden. Endlich wird die Integration bereits vorhandener, aber relativ schwer zugänglicher editorischer Vorarbeiten – z.B. des Personenindex 2.Stufe – angestrebt.

3 Fazit

Im Rahmen des von der DFG geförderten Projekts zur *Bereitstellung und Pflege der Werke Immanuel Kants in elektronischer Form* wurde das Bonner elektronische Kant-Korpus ergänzt, weitgehend korrigiert und in ein standardisiertes XML-Format überführt. Ausgehend vom Korpus wurde eine elektronische Edition der Schriften Kants hergestellt und über eine WWW-Schnittstelle allgemein zugänglich gemacht. Die elektronischen Kanttexte können nun *online* gelesen und durchsucht werden.

Die elektronische Edition wurde frühzeitig bekannt gemacht. Es wurden zwei Symposien zur Kant-Edition veranstaltet;²⁴ außerdem wurde die Edition zum Anlass von Kants 200stem Todestag breit annociert. Die Symposien und die Annoncen erzeugten starke Resonanz. Die elektronische Edition wird seitdem intensiv genutzt. Die Nutzer haben Anregungen gegeben, die zu einem großen Teil bei der (Weiter-) Entwicklung der Edition berücksichtigt wurden; die elektronische Edition konnte damit schon im Zuge ihrer Entwicklung praktisch evaluiert werden.²⁵

Die Arbeit an der elektronischen Kant-Edition ist noch nicht abgeschlossen respektive sollte noch nicht abgeschlossen sein. Es gilt u.a. noch die folgenden Aufgaben zu erledigen:

- Das Kant-Korpus ist nicht fehlerfrei; Korrekturen müssen deshalb fortlaufend eingearbeitet werden. (Vgl. Abschnitt 1.1.) Ferner sollen Daten, die im Zuge von Neueditionen und sonstigen Bearbeitungen der Kanttexte anfallen, in das Kant-Korpus integriert werden. (Vgl.

²⁴ Beide Syposioen fanden an der Universität Bonn statt. Am 21.7.2004 wurde ein Bonn-internes Symposion abgehalten. Am 22.4.2005 fand dann in nationales Symposion zur elektronischen Edition der Werke Kants statt. Am zweiten Symposion nahmen Vertreter aller wichtigen Kant-Forschungsstellen in Deutschland teil.

²⁵ Wir bedanken uns insbesondere bei Prof. Werner Stark, der uns zahlreiche Korrekturen zukommen ließ. Außerdem danken wir für Anregungen von Prof. Willaschek, der zu den Herausgebern eines voraussichtlich 2008 bei de Gruyter erscheinenden Kant-Lexikons gehört (<http://www.kant-lexikon.tk>). Die Autoren dieses Lexikons gehören zu den Nutzern der elektronischen Kant-Edition.

Abschnitt 2.4.) Eine vollständig korrigierte Version des Korpus kann als Datengrundlage für eine neu zu erzeugende, gedruckte Edition der Schriften Kants dienen.

- Der Personenindex 2. Stufe kann digitalisiert und in die elektronische Edition integriert werden. (Vgl. Abschnitt 2.2.2.)
- Wir haben bereits in Abschnitt 2.2.1 Hinweise zur Verbesserung der Suchmaschine gegeben. Folgender Hinweis sei ergänzt: Bisher werden Suchergebnisse (a) durch die Angabe des Bände und Seiten, auf denen die Suchausdrücke vorkommen und (b) durch die Angabe des Zeilenausdrucks, in dem der – dann fettgedruckte – Suchausdruck steht, angegeben. Die Angabe der Fundstelle (Band, Seite) ist mit der entsprechenden HTML-Seite verknüpft. Es wäre wünschenswert – so die Rückmeldung von Nutzern der elektronischen Edition – wenn der Suchausdruck auch auf der verknüpften HTML-Seite durch Fettdruck hervorgehoben wäre. Diese Hervorhebung kann realisiert werden, ohne dass *webglimpse* durch eine andere Suchmaschine ersetzt wird.
- Die Sortierung der Notizen aus AA 14-15 hinsichtlich ihrer Fundorte kann mit relativ geringem Aufwand abgeschlossen werden. (Vgl. Abschnitt 2.1.2.)
- Die elektronische Version von Achenwalls *Juris naturalis* kann mithilfe des Fine Reader XIX vervollständigt werden. (Vgl. Abschnitt 1.1.)
- Es sollten weitere Experimente zur Erzeugung und zum Gebrauch lemmatisierter Wortindizes gemacht werden. Dabei wäre insbesondere zu untersuchen, wie die Disambiguierung gewährleistet werden kann und wie adäquaterweise mit verschiedenen Schreibvarianten – gerade im handschriftliche Nachlass kommen Wörter in verschiedenen Schreibungen vor – umzugehen ist. (Vgl. Abschnitt 2.2.2.) Die Lemmatisierung ist ein wichtiger Schritt in der Erschließung von Kants Werken für die computergestützte und computerunterstützte wissenschaftliche Untersuchung.²⁶
- Schließlich können die Verknüpfungen der Texte und Textstellen ergänzt und optimiert werden. Dies betrifft erstens die Verknüpfungen von Notizen mit ihren Bezugstexten und zweitens die Verknüpfungen

²⁶ Die Erschließung von Kants Werken für die computergestützte und computerunterstützte wissenschaftliche Untersuchung wurden von den DFG-Projektgutachtern als ein besonders relevantes und interessantes Anliegen hervorgehoben.

von Briefen und Aufzeichnungen mit Bildern der Original-Handschriften. Drittens können die Kanttexte untereinander aufgrund inhaltlicher Kriterien verknüpft werden. Solche Verknüpfungen fallen u.U. bereits in den Bereich philosophischer Textinterpretationen. (Vgl. Abschnitt 2.3.)

A Datierungs-codes und Abkürzungen

Der von Erich Adickes und seinem Assistenten Friedrich Berger herausgegebene Teil des handschriftlichen Nachlasses ist in 8141 Einzelnotizen aufgeteilt. Jede Notiz ist mit einem Notizkopf versehen. In den Notizköpfen stehen erstens die jeweilige Notiznummer, zweitens ein Code für mögliche Datierungen, wann die Notiz verfasst wurde, drittens der Fundort und ggf. viertens weitere Anmerkungen. Adickes schlüsselt im Vorwort von AA 14 die Datierungs-codes und die – vor allem zur Angabe der Fundorte – verwendeten Abkürzungen auf. Der Schlüssel für die Datierungs-codes und das Abkürzungsverzeichnis sind im Folgenden komprimiert wiedergegeben:

A.1 Datierungs-codes

- α^1 : Etwa 1753 - 1754.
- α^2 : Etwa 1754 - 1755.
- β^1 : 1752 - W.S. 1755/56.
- β^2 : 1758 - 1759.
- γ : 1760 - 1764.
- δ : Um 1762 - 1763.
- ϵ : Sicher vor ζ , Verhältnis zu δ nicht sicher bestimmbar.
- ζ : Um 1764 - 1766.
- η : 1764 - 1768.
- ϑ : Etwa 1766 - 1768, sicher nach ζ , vor κ .
- ι : Etwa 1766 - 1768, sicher nach ζ , vor κ .
- κ : 1769.
- λ : Ende 1769 - Herbst 1770.
- μ : Etwa 1770 - 1771, sicher später als κ , λ , früher als ν , ξ , o .

- ν : Etwa 1771, sicher später als κ , λ , μ , früher als ξ , o .
- ξ : Etwa 1772, sicher später als $\kappa - \nu$, früher als v , φ .
- o : Sicher früher als v und φ , später als $\kappa - \nu$ und in vielen Fällen auch als ξ ; anderwo mögen ξ und o gleichzeitig sein.
- π : Wahrscheinlich zwischen den Phasen ξ und ϱ .
- ϱ : Um 1773 - 1775.
- σ : Etwa 1775 - 1777. σ ist oftmals gegen χ nur schwer oder gar nicht abzugrenzen, und auch nach rückwärts mag σ sich noch bis in das Jahr 1774 hinein erstrecken.
- τ : Um 1775 - 1776, sicher nach κ , μ , vor v , φ , ψ
- v und φ : um 1776 - 1778, sicher später als $\kappa - \tau$, früher als ψ
- χ : 1778 - 1779. Bei manchen Bemerkungen ist die Abgrenzung nach σ , bei anderen die nach ψ hin schwer oder gar unmöglich.
- ψ^1 : Etwa 1780 - 1783.
- ψ^2 : Etwa 1783 - 1784.
- ψ^3 : Etwa 1785 - 1788.
- ψ^4 : Etwa 1788 - 1789.
- ω^1 : 1790 - 1791.
- ω^2 : 1792 - 1794 (1. Drittel).
- ω^3 : 1794 - 1795.
- ω^4 : 1796 - 1798.
- ω^5 : Sommer 1798 - 1804.

Wenn zwei nicht unmittelbar aufeinander folgende griechische Buchstaben durch einen Bindestrich verbunden sind (Bsp.: $\kappa - \xi$), dann ist jede der durch diese Buchstaben und die zwischen ihnen liegenden Buchstaben bezeichneten Phasen eine mögliche Datierung der jeweiligen Notiz. Wenn eine Datierung nicht zweifelsfrei möglich ist, dann werden die gleich wahrscheinlichen Phasen hintereinander mit je einem Fragezeichen versehen angeführt (Bsp.: $\mu? \nu?$). Mit zwei Fragezeichen versehene Datierungen kommen weniger stark in Betracht (Bsp.: $\psi??$). Schließlich können Datierungen nach ihrer Wahrscheinlichkeit angeordnet werden. Weniger wahrscheinliche Datierungen werden rechts von den wahrscheinlicheren in Klammern gesetzt (Bsp.: $\mu? \nu? (\kappa? \varrho?)$).

A.2 Abkürzungen

- *A.M.* = Altpreußische Monatsschrift.
- *B* = Kants Handexemplar der „Beobachtungen über das Gefühl des Schönen und Erhabenen“.
- *E* = B. Erdmann: „Reflexionen Kants zur kritischen Philosophie“.
- *Hb.* = Kants „Sämtliche Werke“, in chronologischer Reihenfolge hrsgg. von G. Hartenstein, 1867 - 1868.
- *Hg.* = Herausgeber.
- *J* = Kants Handexemplar von Achenwalls „Juris naturalis“ pars posterior.
- *Ki.* = Kantausgabe der „Philosophischen Bibliothek“, ursprünglich hrsgg. von J. H. von Kirchmann; die beigetzten lateinischen Ziffern bezeichnen die Band-Zahlen der letzteren.
- *L* = Kants Handexemplar von Meiers „Auszug aus der Vernunftlehre“.
- *L Bl.* = Loses Blatt.
- *M* = Kants Handexemplar von Baumgartens „Metaphysica“.
- *Ms.* = Manuscript.
- *Pr* = Kants Handexemplar von Baumgartens „Initia philosophiae practicae primae“.
- *R* = R. Reike: „Lose Blätter aus Kants Nachlass“.
- *R.-Sch.* = Kants „Sämtliche Werke“ hrsgg. von K. Rosenkranz und F. W. Schubert.
- *R V* = Kants Handexemplar seiner „Kritik der reinen Vernunft“.
- *Th* = Kants Handexemplar von Eberhards „Vorbereitungen zur natürlichen Theologie“.

Literatur

- [1] Goldfarb, C., 1991. *The SGML Handbook*, Oxford: Oxford University Press.
- [2] Lenders, W., 1982. Der allgemeine Kantindex. Vom Stellenindex zum Informationssystem. *Kant-Studien* 73, 440-451.
- [3] Schmitz, H.-C. (Hg.), 2003. *Bereitstellung und Pflege der Werke Immanuel Kants in elektronischer Form*, IKP-Arbeitsbericht NF 03, Universität Bonn.

- [4] Stark, W., 1993. *Nachforschungen zu Briefen und Handschriften Immanuel Kants*, Berlin: Akademie-Verlag.
- [5] Weber, N., 1983. Textkodierung im maschinenlesbaren Kant-Korpus. In *Methoden der quantitativen Semantik und der Erstellung lemmatisierter Indizes in Anwendung auf die Werke I. Kants*, Büchel, G., N. Weber, G. Wirtz (Hgg.), IKP-Arbeitsbericht 3, Universität Bonn.
- [6] Willée, G., H. Elsen, 2006. *Funktionsbeschreibung von LEMMA4, einem Wortklassentagger und Lemmatisierer für unrestringierte deutsche Texte*, IKP-Arbeitsbericht NF 16, Universität Bonn.