

Datenaustausch mit Disketten

von Andreas Herberger

Der Einsatz von Microcomputern ist heute bereits so weit verbreitet, daß in vielen Fällen auf die Übersendung von Daten in gedruckter Form verzichtet werden könnte. Dennoch wird erst in seltenen Fällen von dieser Möglichkeit Gebrauch gemacht. Dies liegt zum einen an der mangelnden Information der Beteiligten über die Möglichkeiten, Daten mit Hilfe von Disketten auszutauschen. Ein weiteres nicht zu unterschätzendes Problem ist aber auch, daß es heute mehrere hundert verschiedene Aufzeichnungsarten gibt, die einen Diskettenaustausch nicht immer ohne weiteres ermöglichen. Schließlich scheidet diese Art von Datenaustausch überall dort aus, wo Urkundensicherheit zu gewährleisten ist.

1. Ausgangssituation

Zunächst sollen verschiedene Anwendungsbereiche, in denen ein Datenaustausch sinnvoll ist, skizziert werden. Grundsätzlich ist dabei zu bedenken, daß ein Austausch von maschinenlesbaren Daten immer nur dann zu empfehlen ist, wenn eine Informationsmenge bestimmten Umfangs an eine Stelle übergeben werden soll, die diese Informationen weiterverarbeitet. Es wäre sicher nicht sinnvoll, anstelle eines Briefes, der vom Empfänger gelesen und nach Bearbeitung abgelegt wird, eine Diskette zu versenden. In diesem Falle ist das Ziel der Information der Empfänger, der die übermittelten Daten gewissermaßen als „Endstation“ erhält.

Anders ist die Situation schon im Falle eines Anwalts, der für einen Kollegen, der ihn bei einem Gericht an einem anderen Ort vertritt, einen Schriftsatz vorbereitet. In diesem Falle kann es sinnvoll sein, den Schriftsatz auf einer Diskette zu transportieren. Der Inhalt kann dann vom empfangenden Anwalt aufgrund kurzfristiger Veränderungen der Situation eventuell ergänzt und auf eigenem Briefpapier ausgedruckt werden.

Noch geeigneter für eine Datenübermittlung ist die Zusammenarbeit „Autor – Verlag – Druckerei“. In dem meisten Fällen erstellt der Autor heute ein Manuskript mit der Schreibmaschine (oder auch schon mit einem Textsystem). Der mit der Schreibmaschine (oder dem Drucker) ausgegebene Text wird dann an den Verlag weitergeleitet, der ihn an die Druckerei zum Setzen gibt. Hier wird der Text erneut erfaßt, mit den Steuerzeichen, die zum Satz notwendig sind versehen und dann in Form einer Druckfahne als Probeabzug zu Papier gebracht. Den Probeabzug erhält der Autor zum Korrekturlesen, und im Anschluß daran werden die entstandenen Satzfehler korrigiert. Würde hier ein Datenaustausch über Disketten stattfinden, so könnten

verschiedene Arbeitsgänge rationeller gestaltet werden. In der Druckerei entfällt die Neueingabe des Textes, da dieser direkt von der Diskette in das Satzsystem eingelesen wird. Der Setzer fügt lediglich die Satzsteuerzeichen hinzu und veranlaßt die Probebelichtung. Da der Text nicht neu erfaßt wurde, fallen Eingabefehler weg, die bei der Texterfassung unweigerlich auftreten. Zusätzlich wird die Durchlaufzeit in der Setzerei reduziert, da die Erfassungszeit eingespart wird. Als weiterer zeitintensiver und anstrengender Arbeitsgang entfällt für den Autor das Korrekturlesen der Probeabzüge. Er muß die Probeabzüge nur noch daraufhin durchsehen, ob alle Satzsteuerzeichen richtig eingefügt worden sind (z.B. Fettdruck, Kursivschrift, Fußnoten etc.). Auch dieser Schritt kann bei entsprechender Absprache und Fachkenntnis des Autors entfallen, da durch die Vereinbarung von Abkürzungen praktisch alle Satzfunktionen bereits im Manuskript enthalten sein können.

Eine weitere Notwendigkeit für die Datenübernahme mittels Disketten ist der Systemwechsel eines Computeranwenders. In diesem Falle sind häufig größere Datenmengen von dem vorher benutzten System auf das Nachfolgesystem zu übernehmen, was am einfachsten durch das Einlesen der Daten von Disketten zu realisieren ist.

Es sind sicherlich noch viele weitere Anwendungsmöglichkeiten für die skizzierte Form des Datenaustausches denkbar. In allen Fällen wird man aber das Muster erkennen können, das auch in den beiden Beispielen enthalten ist:

1. Der Text oder die Daten werden bereits auf einem Computer erstellt
2. Der Text oder die Daten werden einem Empfänger zugesellt, der die Daten weiterhe- oder -verarbeitet.

2. Technische Probleme beim Diskettenaustausch

Schon nach kurzer Zeit stellt jeder, der sich mit der Materie des Diskettenaustauschs befaßt, fest, daß es verschiedene Hindernisse gibt, die es zu überwinden gilt, bevor eine Diskette die auf dem Computer X erstellt wurde von dem Computer Y gelesen werden kann (Selbstverständlich vorausgesetzt: X ungleich Y).

Zunächst gibt es rein mechanische Probleme. Es sind nämlich nicht nur verschiedene Aufzeichnungsarten sondern auch verschiedene Abmessungen der Disketten gebräuchlich. Die Disketten sind normalerweise quadratisch und werden nach ihrer Kantenlänge bezeichnet

Anfang der siebziger Jahre gab es nur Disketten im Format 8 Zoll. Kurz darauf wurden die ersten 5 1/4 Zoll Disketten eingeführt und in den letzten Jahren beginnen sich die 3 1/2 Zoll Disketten Marktanteile zu sichern. Andere Formate sind zwar ebenfalls vorhanden, konnten sich jedoch nicht durchsetzen.

Die erste Schwierigkeit beim Diskettenaustausch ist also die äußere Abmessung. Es leuchtet ein, daß ein Diskettenlaufwerk, das für den Betrieb von 5 1/4 Zoll Disketten ausgelegt ist keine Disketten aufnehmen kann, die eine Kantenlänge von 8 Zoll haben und umgekehrt. Auch ein Adapter, der zumindest für den Betrieb der kleineren 5 1/4 Zoll Disketten in einem 8 Zoll Laufwerk zu verwenden wäre, ist aus technischen Gründen nicht denkbar.

Erste Bedingung für den Datenaustausch ist also, daß Quell- und Zielgerät über Diskettenlaufwerke für Disketten mit den gleichen Abmessungen verfügen.

Doch diese Voraussetzung alleine genügt noch nicht. Das nächste Problem ist, daß unterschiedliche Aufzeichnungsverfahren existieren, wie die Daten in elektromagnetischer Form auf der Diskettenoberfläche abgelegt werden.

Die Parameter, die für die Aufzeichnungsart kennzeichnend sind, stellt die abgedruckte Tabelle - ohne Anspruch auf Vollständigkeit - zusammen.

Neben der unterschiedlichen Definition dieser Parameter haben die Hersteller von Computersystemen zusätzlich noch unterschiedliche Wege gewählt, die Daten auf der Diskette zu erschließen. Es leuchtet ein, daß es nicht sinnvoll ist, die Daten einfach als Strom von Zeichen beginnend bei der ersten Spur auf die Diskette zu schreiben. Es wäre dann sehr schwierig einen bestimmten Teil der Daten aufzufinden und es müßte bei der Suche jeweils die gesamte Diskette bearbeitet werden. Zugegebenermaßen wäre dieses Verfahren jedoch für einen reinen Datenaustausch recht geeignet und ist auch in modifizierter Form im Einsatz.

In der Praxis ist es jedoch so, daß ein bestimmter Bereich der Diskette reserviert ist und ein Inhaltsverzeichnis (auch „Directory“ genannt) enthält. Das Inhaltsverzeichnis besteht aus Informationen darüber, welche Daten sich wo auf der Diskette befinden. Der Aufbau und die Position des Inhaltsverzeichnisses wird vom Betriebssystem des Rechners bestimmt. Das Betriebssystem verwaltet auch die darin enthaltenen Einträge. Anlagen mit unterschiedlichen Betriebssystemen können daher auch nur selten Disketten miteinander austauschen, selbst wenn die übrigen Aufzeichnungsparameter identisch sind.

3. Vorschläge zur Realisierung eines Datenaustausches über Disketten

3.1 8-Zoll Formate

Obwohl der Datenaustausch mittels Disketten angesichts von über 100 Diskettenformaten ein scheinbar aussichtsloses Unterfangen zu sein scheint, ist dies praktisch trotzdem nicht so.

Es haben sich nämlich bestimmt durch die Gegebenheiten des Marktes mehrere sogenannte Standard-

formate herausgebildet. Die sogenannten hardsektorierten Disketten, die im Fotosatzbereich weit verbreitet sind, haben sich nicht durchgesetzt. Es gibt auch keine Möglichkeit, mit einem System, das Laufwerke für hardsektorierte Disketten enthält, sogenannte softsektorierte Disketten, wie sie nachfolgende beschrieben werden zu verarbeiten. Aus diesem Grunde und wegen der relativ geringen Verbreitung wird auf diesen Diskettentyp nicht näher eingegangen.

Eines der Standardformate, das im 8 Zoll Bereich auch heute noch von fast jedem System, das 8 Zoll Disketten Lesen oder Schreiben kann beherrscht wird, ist das sogenannte IBM 3740-Format. Dieses Format ist - wie bereits aus dem Namen ersichtlich ist - von der Firma IBM entwickelt und eingeführt worden. Es basiert auf einer 8-Zoll-Diskette, die mit einer Sektorlänge von 128 Zeichen einseitig mit einfacher Schreibdichte beschrieben wird. Dies ergibt eine nutzbare Menge von etwa 250.000 Zeichen pro Diskette (entsprechend etwa 100 Schreibmaschinen Seiten zu 2500 Zeichen).

Im Laufe der Zeit wurde das Format in den Bereich der doppelten Schreibdichte und der doppelseitigen Verwendung der Disketten erweitert. Diese Erweiterungen konnten sich aber nicht in dem Maße durchsetzen wie das ursprüngliche 3740 Format.

Praktisch alle Systeme, die mit 8-Zoll Diskettenlaufwerken ausgestattet sind, können Disketten mit diesem Format verwenden.

Das 3740-Format ist insbesondere beim Austausch von Daten mit Groß-EDV Anlagen von Bedeutung, da hier normalerweise nur dieses Format bearbeitet werden kann.

Soweit zum physikalischen Aufbau des IBM 3740 Formates. Bezüglich der logischen Struktur (d.h. Lage und Aufbau des Inhaltsverzeichnisses) sind zur Zeit zwei Verfahren gebräuchlich. Verfahren 1 ist das von IBM angewandte Verfahren, das auf der Spur 0 ein Inhaltsverzeichnis mit Platz für maximal 26 Dateien anlegt. Zu jeder Datei wird dann gespeichert auf welcher Spur sie beginnt und auf welcher Spur sie aufhört. Mit diesem Verfahren sind nur zusammenhängende Dateien realisierbar und nach dem Löschen einer Datei verbleibt auf der Diskette eine Lücke, die nur mit einer kleineren Datei oder mit einer Datei gleicher Größe gefüllt werden kann.

Diese Nachteile führten dazu, daß eines der im Zusammenhang mit 8 Zoll Diskettenlaufwerken am weitesten verbreiteten Betriebssysteme - das Betriebssystem CP/M von Digital Research - zwar die physikalische Aufzeichnungsart (einseitig, einfache Schreibdichte) übernommen hat, aber ein besser organisiertes Directory verwaltet. CP/M unterteilt die Diskette in 1024 Zeichen große Blöcke (insgesamt 243). Der erste dieser Datenblöcke ist für das Inhaltsverzeichnis reserviert und die ersten beiden Spuren einer Diskette werden für Benutzerdaten gar nicht verwendet sondern sind für das Betriebssystem reserviert. Dieses Format kann von fast jedem System, das CP/M verwendet und mit 8 Zoll Laufwerken ausgerüstet ist, verarbeitet werden, auch dann, wenn das normalerweise vom System

verwendete Diskettenformat davon abweicht. Dies hat seinen Grund darin, daß alle Software, die für CP/M auf dem Markt erhältlich und auch CP/M selbst in eben diesem Format geliefert wird, so daß es notwendig ist, dieses Format auf jedem Fall im System implementiert zu haben, um Software hinzukaufen zu können.

Für CP/M existieren auch Programme, die es ermöglichen, CP/M-formatierte Disketten in IBM-3740 Disketten (bzgl. des Inhaltsverzeichnisses) umzusetzen (z.B. Reformatter von Micro Tech Exports).

3.2 5 1/4-Zoll Formate

Bei den 5 1/4 Zoll Formaten fehlte es lange Zeit an einem brauchbaren Standard. Es existierte eine ECMA-Norm, die z.B. von der Datev verwendet wird, sich jedoch nicht durchsetzen konnte. Insgesamt gibt es weit über 100 Diskettenformate im 5 1/4 Zoll Bereich. Erst durch die Einführung des IBM Personal Computers, der in vielen Bereichen einen gewissen Standard gesetzt hat, entstand auch bezüglich des Diskettenformates ein Quasi-Standard. Anders als bei den 8 Zoll Formaten, wo IBM das erste Format eingeführt hatte und alle anderen Anbieter, die später Systeme herstellten, dieses Format zumindest beibehielten, gab es bereits 5 1/4 Zoll Systeme, bevor IBM ein solches System vorstellte. Daher existierten bereits vor dem Bekanntwerden des „IBM-Standards“ Systeme mit unterschiedlichen Diskettenformaten. So kommt es, daß auch heute noch immer nicht alle Systeme, die mit 5 1/4-Zoll Disketten arbeiten, Disketten im IBM-PC Format verarbeiten können. Erst neuere System sind, wenn sie nicht ohnehin zu den sogenannten „Kompatiblen“ gehören, in der Lage, normalerweise auch Disketten im IBM-PC Format zu verwenden.

Die IBM-PC Disketten sind 5 1/4 Zoll Disketten, die beidseitig mit doppelter Schreiddichte und 48 Spuren pro inch beschrieben werden. Jede Spur hat 9 Sektoren, was eine Gesamtkapazität von etwa 350.000 Zeichen pro Diskette ergibt (entsprechend etwa 140 Schreibmaschinenseiten). Auch hier stellt sich wieder die Frage nach der logischen Aufteilung der Diskette und der Lage und Verwendung des Inhaltsverzeichnisses. Anders als bei dem 3740-Format hat IBM hier bereits ein komfortableres System gewählt. Das von IBM eingesetzte Betriebssystem PC-DOS (abgeleitet vom Betriebssystem MS-DOS der Firma Microsoft) legt am Anfang der Diskette ein Inhaltsverzeichnis an und trägt hier ein, in welchen Blöcken auf der Diskette die zu einer Datei gehörenden Daten enthalten sind.

Dieses Verfahren ist ähnlich dem des Betriebssystems CP/M, das auch für den IBM-PC angeboten wird. Aufgrund unterschiedlichen internen Aufbaus beider Betriebssysteme kann aber eine mit CP/M auf dem IBM-PC erstellte Diskette nicht mit PC-DOS bearbeitet werden und umgekehrt. Zur Charakterisierung einer Diskette genügt es also nicht, festzustellen, daß es sich um eine Diskette im IBM-PC Format handelt. Es muß vielmehr immer noch angegeben werden, mit

welchem Betriebssystem diese Diskette beschrieben worden ist.

In neuerer Zeit ist eine Version von CP/M für den IBM-PC erhältlich (Concurrent PC-DOS), die beide Diskettenformate lesen und schreiben kann und damit einen erfreulichen Beitrag zur Vereinfachung des Diskettenaustausch leistet.

Zur Übernahme von Daten von nicht IBM-kompatiblen Disketten bieten sich verschiedene Möglichkeiten an. Eine Möglichkeit ist es, mit einem vom Hersteller des Systems angebotenen Hilfsprogramm IBM-PC formatierte Disketten bearbeiten zu können. Wenn ein solches Programm nicht erhältlich ist, kann man auf dem IBM-PC mit Hilfe von Diskettenkonvertierprogrammen (z.B. das Programm Crossdata von Award Software) fremde Diskettenformate in das IBM-PC Format umsetzen. Über diese Möglichkeit verfügen heute schon viele Setzereien, die je nach Größe des Betriebes auch über spezielle Diskettenkonvertiersysteme verfügen, die viele der beschriebenen Probleme überwinden helfen.

4. Weitere Probleme

Ist die Diskettenübernahme auf das Zielsystem sichergestellt, können immer noch einige Schwierigkeiten auftauchen, die aber sehr leicht behoben werden können, wenn man damit rechnet.

Bestimmte Textverarbeitungsprogramme verwenden Dateiformate, die Steuerzeichen für den Ausdruck und die Formatierung (z.B. Fettdruck oder Absatzende) enthalten. Hier empfiehlt es sich eventuell bereits bei der Erstellung des Manuskriptes ganz auf diese Funktionen zu verzichten und einen Absatz immer durch eine Leerzeile zu kennzeichnen. Handelt es sich um so weitverbreitete Programme wie z.B. Wordstar von Micropro, kann die Druckerei in der Regel aber auch die Drucksteuerzeichen verarbeiten und z. B. in Fontwechsel umsetzen. Bei anderen Programmen gibt es spezielle Umsetzungsprogramme, die eine Datei im sogenannten „System Data Format (SDF)“ erzeugen, die nur noch die reine Textinformation enthält.

Weiterhin ist an die Verwendung von Umlauten und Sonderzeichen zu denken. Hierbei muß man der Druckerei zumindest mitteilen, ob das System auf dem der Text erstellt worden ist, den 7-Bit ASCII Code mit nationalen Umlauten oder den erweiterten 8-Bit ASCII Code, der z.B. auf dem IBM PC eingesetzt wird, verwendet. Auch für die Umsetzung dieser beiden Codes ineinander existieren Programme. Der im Gross-EDV Bereich häufig verwendete EBCDIC-Code spielt im Micro- und Fotosatzbereich, falls dort nicht Magnetbänder verwendet werden, eher eine untergeordnete Rolle.