

5. Der Umgang mit Dateien

Die Dateien sind das A und O des PCs: Alle zusammengehörenden Informationen sind in einer bestimmten Art und Weise auf dem Datenträger abgelegt, sie bilden eine Datei.

Ob Sie mit einem Programm eine Konstruktionszeichnung erstellen oder ein Logo für Ihre Briefbögen entwerfen, ob Sie einen Brief mit der Textverarbeitung schreiben oder eine Berechnung mit der Tabellenkalkulation anstellen, alle Ergebnisse Ihrer Arbeit werden als zusammengehöriges Ganzes, als Datei abgespeichert. Der Vergleich mit einem Vorgang, der zusammengehörende Daten enthält, ist sicher nicht weit hergeholt.

5.1 Die Arten von Dateien

Man unterscheidet zwei große Gruppen von Dateien:

Da sind zum einen die Datendateien. Sie sind aus einer Anwendung heraus entstanden, eben jener Brief aus der Textverarbeitung. Diese Dateien sind vollkommen passiv. Allerdings werden sie in einem ganz bestimmten Format abgespeichert, das für das jeweilige Anwendungsprogramm, aus dem sie entstanden sind, typisch ist. Die Daten aus einem Datenbankprogramm werden z. B. in einer Struktur auf dem Datenträger abgelegt, die in der Zwischenzeit zu einem Standard für nahezu alle Datenbanken geworden ist. Dies hat den unbestreitbaren Vorteil, daß die Daten des einen Programms in das andere eingelesen werden können.

Eine Sonderform der Datendateien sind die reinen Textdateien oder auch ASCII-Dateien, die, anders als die Datendateien aus Anwendungsprogrammen, nur Text, also nur Buchstaben und Satzzeichen enthalten können, nicht aber Steuerzeichen, wie sie zur Codierung der Strukturen einer Datendatei benutzt werden.

Der Name ASCII leitet sich ab aus "American Standard Code of Information Interchange". Das ist ein normierter Code für die 256 Zeichen, die der Computer darstellen kann. Diese Zeichen sind von 32 - 255 durchnummeriert. Die Zeichen 0 bis 127 sind auf allen Computern und in allen Softwarepaketen der Welt gleich - der sogenannte erweiterte ASCII-Code von 128 bis 255 sind entweder Grafikzeichen oder Zeichen anderer Sprachen.

Die reinen Textdateien oder ASCII-Dateien enthalten nur die Zeichen zwischen ASCII 32 und 255 die Zeichen 0 bis 31 werden als Steuerzeichen reserviert und dürfen in dieser Art Datei nicht enthalten sein.

Die CONFIG.SYS und die AUTOEXEC.BAT müssen Textdateien sein. Daher darf für ihre Erstellung oder Änderung (Editierung) nur ein Editor verwendet werden, der solche ASCII-Dateien erstellt, wie der bei DOS mitgelieferte Editor EDIT.

Wegen des verschiedenen Dateiformates gelingt auch der Transfer von Daten zwischen verschiedenen Programmen manchmal nur über den (unkomfortablen) Umweg, aus dem einen Programm eine reine Textdatei zu erstellen und diese dann in das andere Programm einzulesen.

Die andere Gruppe sind die Programmdateien. Diese Dateien enthalten einen von einem Programmierer entwickelten Code, der zwar auch zuerst einmal passiv auf der Platte "herumliegt", der jedoch als Programm gestartet werden kann. Die in dieser Datei gespeicherten Steuerzeichen erwecken den PC zum Leben: Eröffnungsbilder werden gezeigt, Daten können eingegeben und verarbeitet werden.

Da diese Dateien im Gegensatz zu den völlig passiven Datendateien ausgeführt werden können, nennt man sie auch ausführbare Dateien.

Eine Sonderform der Programmdateien sind die sogenannten Batch-Dateien (Batch engl. = Stapel). Diese sind eigentlich ein Zwitter: Einerseits sind es reine Textdateien, enthalten also keine Steuerzeichen, andererseits sind es in der Tat Programme, die wie jedes andere Programm ausgeführt werden können.

5.2 Der Dateiname

Eine sehr wichtige Rolle spielt für das Betriebssystem und Sie als Anwender der Dateiname, anhand dessen eine Datei im System eindeutig bestimmt und identifiziert werden kann. Daher ist die erste Regel einleuchtend: Eine Datei unter DOS muß einen Namen haben. Eine Datei ohne Namen wäre nicht mehr identifizier- und damit nicht mehr auffindbar. Dieser Name kann bis zu acht Zeichen umfassen, mehr Zeichen werden nicht zugelassen (in Anwendungsprogrammen) bzw. einfach ignoriert (von MS-DOS).

Hinweis

Jedes Zeichen, das die höchstzulässige Anzahl überschreitet, wird von MS-DOS nicht zurückgewiesen. Wenn Sie z. B. einer Datei den Namen

LANGERNAME

zuordnen, werden die letzten zwei Zeichen ignoriert und Sie finden im Inhaltsverzeichnis eine Datei

LANGERNA

vor. Das kann - wie man sich denken kann - zu großen Problemen führen! Andererseits steht nirgendwo geschrieben, daß etwa

A

kein gültiger Dateiname ist - natürlich ist er das. Sie sollten jedoch derartig kurze Dateinamen vermeiden, da zum einen nach kurzer Zeit niemand mehr weiß, welchen Inhalt diese Datei haben könnte und zum anderen eine Verwechslung mit einem Laufwerknamen nur allzu leicht möglich ist.

Übrigens: Es ist Ihnen inzwischen sicher längst klar, warum die Laufwerknamen immer mit Doppelpunkt angegeben werden, denn ohne den Doppelpunkt wird in einem Befehl aus dem Laufwerk A: schnell die Datei A - mit möglicherweise schlimmen Folgen.

Nun steht jeder Anwender immer wieder vor der schwierigen Aufgabe, für eine Datei einen Namen und dabei die Quadratur des Kreises zu erfinden: Einerseits muß er einen möglichst deskriptiven Namen vergeben, der ihm auch noch in Monaten etwas sagt, andererseits darf man nicht mehr als acht Zeichen verwenden.

Eine allgemeingültige Regel für die Vergabe eines Dateinamens kann nicht gegeben werden, doch dürfte es sicher einleuchtend sein, daß der Name einen Bezug zum Inhalt der Datei haben sollte. Wenn mehrere Personen auf die gleichen Dateien zugreifen, sollte man präzise Regeln erstellen, damit die Namen für alle aussagekräftig genug sind.

Die Erweiterung

Es wird Ihnen schon mehrfach aufgefallen sein, daß der Name der Datei durch drei Zeichen ergänzt wird, etwa AUTOEXEC.BAT oder COMMAND.COM - diese sogenannte Erweiterung kennzeichnet den Typ der Datei, um welche Art von Datei es sich also handelt. So kennzeichnet die Erweiterung BAT den Typ Batch-Datei, die Erweiterung TXT eine Datei aus einer Textverarbeitung etc. Diese Erweiterung kann bis zu drei Zeichen haben und wird vom Dateinamen mit einem Punkt abgetrennt. Vor oder nach dem Punkt ist keine Leertaste erlaubt:

NAME.ERW

Zwei Erweiterungen wollen wir hier näher beleuchten, nämlich die Erweiterungen COM und EXE.

Diese beiden Erweiterungen (und nur diese!) kennzeichnen Programme und bilden zusammen mit den Batch-Dateien (Erweiterung BAT) die ausführbaren Dateien, also jene Dateien, die Sie auf Ihrem PC ausführen können: Wenn Sie die Datei WORD.EXE beispielsweise starten, indem Sie

WORD

eintippen und mit «Enter» an den Kommandoprozessor übergeben, wird der gesamte Programmcode des Textprogramms, der in dieser Datei enthalten ist, gestartet. Um eine ausführbare Datei zu starten, müssen wir die Erweiterung also nicht angeben.

Falls Sie nun meinen, mit der Erweiterung sei das Dilemma mit der etwas eingeeengten Namensgebung gelöst, so müssen wir Sie enttäuschen:

Die Erweiterung darf nur in den seltensten Fällen für eine weitere Kennzeichnung der Datei von Ihnen frei vergeben werden, sondern sie wird von den Anwendungsprogrammen vergeben. So sucht das (und nicht nur das) Textprogramm MS-Word - wenn Sie nichts anderes befehlen - nach einer Datei mit der Erweiterung TXT. Aus diesem Grunde steht Ihnen für die Arbeit mit Anwendungsprogrammen also lediglich der Dateiname selbst zur Verfügung.

Übrigens: Man könnte meinen, daß eine Erweiterung drei Zeichen haben muß, doch das stimmt nicht, es sind auch zwei Zeichen oder gar nur eines erlaubt.

Die Zeichen im Dateinamen und der Erweiterung

Es dürfen durchaus nicht alle Zeichen im Dateinamen oder der Erweiterung verwendet werden. Zugelassen sind alle Buchstaben von A bis Z sowie alle Zahlen, wobei der Name durchaus mit einer Zahl oder einem zugelassenen Sonderzeichen beginnen darf, der Dateiname

3_VON_7.TXT

ist genauso zulässig wie

(HELGA).DOC

Folgende Sonderzeichen sind erlaubt:

% \$ () & - _ ~ { } ! @ ^ ° ' `

Dagegen auf keinen Fall erlaubt sind die Leertaste sowie all jene Sonderzeichen, die unter DOS eine eigene Bedeutung haben, nämlich

+ : . , \ / = > < []

Die Zeichen

Ä Ö Ü ß

dürfen zwar verwendet werden, doch sind sie mit Vorsicht zu genießen, denn da es diese Zeichen in den USA nicht gibt, viele Programme jedoch aus Amerika kommen, kann es passieren, daß der Programmierer dort schlicht vergessen hat, diese Möglichkeit vorzusehen. In einem solchen Fall kann es vorkommen, daß Ihr Anwendungsprogramm eine Datei mit einem solchen Zeichen im Dateinamen nicht findet, obwohl die Datei physisch vorhanden ist.

Die letzte Regel, die Sie kennen müssen, ist diese: Wie bei der Eingabe eines Befehls, ist auch bei den Dateinamen die Schreibweise mit Groß- und Kleinbuchstaben zugelassen.

gross.txt

ist identisch mit

GROSS.TXT