

Bernd Wenzel

Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften,  
München

## EINIGE BEGRIFFLICHE FESTLEGUNGEN ÜBER DIE SPEICHERUNG UND WIEDERGEGENWANNUNG VON INFORMATIONEN

Zur Zeit wird die Fachwelt im Bereich der Speicherung und Wiedergewinnung von Information von einer nahezu babylonischen Begriffsverwirrung heimgesucht. Das hat verständlicherweise zu großer Unsicherheit, Beunruhigung und auch Ablehnung auf Seiten der Anwender geführt. Deshalb sollen hier einige wichtige Begriffe einerseits so pragmatisch definiert werden, wie es für die praktische Arbeit von heute sinnvoll erscheint, andererseits sollen die Definitionen aber halbwegs konsistent sein mit dem Datenbank-Jargon. Das muß an einigen Stellen notwendig zu relativ weitgefaßten Begriffen führen, es wird jedoch stets darauf geachtet, daß der Auflösungsgrad der Definitionen für den Anwendungsbe-  
reich voll ausreicht.

Grundlegende Begriffe für unsere folgende Diskussion, deren Bedeutung man sich üblicherweise nicht klarmacht, sind Information und Daten. *Information* ist (gemäß DIN 44300) die Kenntnis über Sachverhalte und/oder Vorgänge.

*Daten* (engl. data) sind (in Anlehnung an DIN 44300) Zeichen (oder kontinuierliche Funktionen), die zum Zweck der Verarbeitung Information aufgrund bekannter oder unterstellter Abmachungen darstellen. Dieser Unterschied zwischen Daten und Information verdient es herausgestellt zu werden. Norbert Wiener sagte einmal: "Information ist weder Materie noch Energie." Daten dagegen sind Abbilder von Information gerade im Bereich der Physik, Abbilder vermöge der zuvor als "Abmachung" bezeichneten Funktion.

Die Begriffe Datensatz (häufiger nur Satz) und Datei sind auch jedem Gelegenheitsanwender der Informationsverarbeitung vertraut. Trotzdem oder vielmehr gerade deshalb hier die Definitionen:

Ein *Datensatz* (engl. record) ist eine Zusammenfassung von Daten zu einer Einheit unter der Annahme, daß eine gemeinsame Verarbeitung der dargestellten Information sinnvoll ist, also eine verarbeitungslogische Dateneinheit.

Eine *Datei* (im Sinne des englischen data set) ist eine Menge von gleichstrukturierten Datensätzen, auf die nach einer Methode zugegriffen wird.

Eine *Datei* (im Sinne des englischen file) ist die Einheit, in der ein Betriebssystem Datenmengen verwaltet. Diese Verwaltungseinheit hat im allgemeinen einen Namen.

Die beiden verschiedenen Bedeutungen des Begriffs Datei werden leider im Deutschen vermengt, doch legt sich kaum jemand darüber Rechenschaft ab.

Dafür streiten die Anwender im Bereich der Begriffe Datenbank, Datenbanksystem, Informationssystem ... oft um das berühmte Jota.

Doch blockieren solche Grundsatzdiskussionen in den allermeisten Fällen nur den Weg zu einem sachlichen Abwägen der Vor- und Nachteile der verschiedenen Problemlösungen. Die folgenden Definitionen sollen das vermeiden helfen.

Eine *Datenbank* (engl. data base) - seltener findet man Informationsbank - ist eine Menge von Daten zusammen mit der Information, was diese Daten bedeuten und wer auf welche Weise auf diese Daten zugreifen (d.h. möglicherweise auch sie verändern) kann.

Ein *Informationssystem* (engl. information system, manchmal auch data bank) ist eine Datenbank zusammen mit Werkzeugen (wie etwa bestimmte

Zugriffstechniken) und Verfahren (wie etwa Programme), die die Information der Datenbank verarbeiten und/oder zur Verfügung stellen.

Will man besonders hervorheben, daß eine Datenbank von einer Rechenanlage verwaltet wird, so sollte man von einer *rechnergestützten Datenbank*, bzw. analog von einem *rechnergestützten Informationssystem* sprechen.

Man mag an dieser Stelle einwenden, daß eine Pappschachtel voll loser Blätter im Sinne obiger Definition eine Datenbank ist, vorausgesetzt irgend jemand ist in der Lage, diese Blätter zu interpretieren. Dieser Mensch bildet dann zusammen mit der Schachtel ein Informationssystem, wenn er Zugang zu ihr hat. Dieses Beispiel ist bewußt extrem formuliert. Es zeigt deutlich, wie weit diese beiden Definitionen gefaßt sind. Man sieht jedoch ebenso deutlich, daß die Definitionen die grundlegenden Anforderungen abdecken ohne sich durch anwendungsspezifische Anforderungen (wie z.B. Dialogfähigkeit) in irgendeiner Richtung einengen zu lassen. Sie bieten also eine solide Basis für sachliche, anwendungsbezogene Diskussionen.

Weniger strittig ist wohl der folgende Spezialfall:

Ein *Dokumentationssystem* (engl. documentation system) ist ein spezielles Informationssystem, das im allgemeinen wissenschaftliche Literaturzitate zu einem gewünschten Themenkreis zur Verfügung stellt.

Wir haben uns bisher mit Begriffen befaßt, die für die Anwender von Informationssystemen von Interesse sind. Jetzt wollen wir uns Begriffen zuwenden, die aus dem Bereich der Implementierung von Informationssystemen stammen, die jedoch häufig mit den uns bereits bekannten Begriffen verwechselt werden.

Ein *Datenbanksystem* (engl. data base system) ist ein Softwarepaket, das die Erzeugung und Wartung von rechnergestützten Datenbanken und Informationssystemen erleichtert und unterstützt.

*Vorsicht:* Der Begriff Datenbanksystem wird häufig in der Bedeutung des Begriffs Informationssystem verwendet!

Eine sehr wichtige Aufgabe eines Datenbanksystems ist die Erzeugung des *Datenbankbetriebssystems* (engl. data base operating system) für eine konkrete Datenbank. Darunter versteht man alle Programme, die zusammen mit Eigenschaften des zugrundeliegenden Betriebssystems und der zugrundeliegenden Rechenanlage die Grundlage der möglichen Zugriffe auf eine Datenbank und die Abwicklung von Programmen, die Informationen aus dieser Datenbank verarbeiten, steuern und überwachen. (Analog zur Definition von Betriebssystem in DIN 44300.)

Das Datenbankbetriebssystem ist Bestandteil des entsprechenden Informationssystems.

Sehr häufig findet man schließlich den Begriff *Datenbankverwaltungssystem* (engl. data base management system). Das ist eine Kombination aus Datenbanksystem und Datenbankbetriebssystem, bei der entweder jede Kopie des Datenbanksystems nur ein Datenbankbetriebssystem und damit nur eine Datenbank erzeugen kann, oder aber das Datenbanksystem enthält ein allgemeines und rudimentäres Datenbankbetriebssystem für alle Datenbanken.

Damit ist die Menge der einschlägigen Begriffe noch lange nicht erschöpft, doch hofft der Autor, daß die hier definierten Begriffe für sachliche Diskussionen über Informationssysteme voll ausreichen.

## LITERATUR

DIN 44300: Informationsverarbeitung, Begriffe, DIN Taschenbuch 25, Beuth-Verlag Berlin-Köln-Frankfurt 1975.

ROUETTE, Leo: Begriffliche Festlegungen über Datenstrukturen und Speicherorganisationen, DIN-Mitteilungen 58/1979, Beuth-Verlag.