

# Glossar

Das Glossar enthält sowohl die im Text eingeführten Grundbegriffe als auch nicht eingeführte Begriffe angrenzender Bereiche. Ein Klick auf ein mit › eingeleitetes Wort führt zur Definition.

**1:1-Beziehung:** ›Beziehung

**1:n-Beziehung:** ›Beziehung

**3GL** (engl. 3rd Generation Language): ›Drittgenerationssprache

**4GL** (4th Generation Language): ›Viertgenerationssprache

**ABC-Analyse:** Verfahren zur Bildung von Schwerpunkten, zum Beispiel zur Unterscheidung in “wichtige” (A), “weniger wichtige” (B) und “unwichtige” (C) Massnahmen. Anwendungen finden sich zum Beispiel bei der Unterscheidung von Beschaffungs- und Bereitstellungsmaßnahmen, in der Organisationsanalyse und im Zeitmanagement.

**Abfrage** (engl. query): Man kann lesend oder schreibend auf  
›Datenbanken zugreifen: Eine *Auswahl*abfrage (Abfrage im engeren Sinn) ist eine Anweisung, die Daten sucht, aber nicht ändert. Eine *Aktions*abfrage ist hingegen eine Anweisung, die Daten fort-schreibt (Anfüge-, Lösch-, Aktualisierungs- oder Tabellenerstellungsabfragen).

**Abfragesprache** (engl. query language): Sprache zur Suche oder Modifikation von Daten. *Prozedurale* Abfragesprachen sind Programmiersprachen, die vom Benutzer die Angabe von Datenstrukturen, Anweisungssequenzen und Kontrollstrukturen in einer bestimmten Reihenfolge verlangen. Ein Beispiel ist die eingebaute Programmiersprache von MS Excel und ›MS Access (›Visual Basic for Applications). In einer *nichtprozeduralen* oder deklarativen Abfragesprache genügt hingegen meist eine einzige (verbale oder

grafische) Anweisung. Sie spezifiziert nur das gewünschte Ergebnis, ohne darauf einzugehen, wie es am effizientesten erzielt werden kann. Beispiele nichtprozeduraler Abfragesprachen sind ›QBE und ›OLAP.

**Abhängigkeit, funktionale** (engl. dependency): Begriff aus der ›Normalisierung. Seien A und B ›Attribute einer ›Relation R: Wenn zu jedem Zeitpunkt t jeder Wert von A eindeutig einen Wert von B =  $f(A, t)$  identifiziert, so sagt man B sei funktional abhängig von A und schreibt  $A \rightarrow B$ .

**ActiveX**: Komponententechnologie von Microsoft. Eine ActiveX-Komponente ist ein Programmmodul (›DLL), das klein genug für die Übertragung auf dem ›Internet ist. Im Gegensatz zu einem ›Applet sind ActiveX-Komponenten Binärcode-Module und für die Betriebssysteme von Microsoft optimiert.

**ad hoc**: ungeplant, “aus dem Stegreif”. Der Begriff wird meist für ›Abfragen verwendet, die erst als Reaktion auf Ergebnisse der gleichen Abfragesitzung formuliert werden.

**ad hoc berechnete** Daten: ›vorberechnete Daten

**Agent**: Prozedur, die nach einem bestimmten Ereignis automatisch eine Aktion auslöst. Auf einem ›OLAP-Server kann zum Beispiel ein Agent nach dem Überschreiten eines bestimmten Schwellenwertes automatisch den Anwender benachrichtigen. Agenten können auch beim Vorliegen benutzerdefinierter Bedingungen Aktionen wie Bestellungen per E-Mail oder Fax in die Wege leiten.  
›Trigger

**Aggregation** (Konsolidierung, engl. auch roll-up): Zusammenfassung innerhalb einer ›Kategorienhierarchie. Zum Beispiel werden die Jahreserträge aller Unternehmungen einer Region zum regionalen Jahresertrag addiert. Allgemeiner ist eine Aggregation eine Zusammenfassung mit einer Rechenfunktion, insbesondere einer

›SQL-Aggregatsfunktion (zum Beispiel der Summen-, Durchschnitts- oder Minimumsfunktion). Zusammenfassungen oder Aggregate werden ›ad hoc ausgeführt oder vom ›Data Warehouse - Verwalter bzw. erfahrenen Benutzern ›vorberechnet. Die Aggregation ist ein Sonderfall der ›Vorberechnung.

**AHP:** ›Anytischer Hierarchieprozess

**AIX** (engl. Advanced Interactive eXecutive): UNIX-Derivat von IBM

**Aktion:** ›Regel

**Aktionsabfrage:** ›Abfrage

**algorithmisch:** Ein algorithmisches Verfahren ist eine umgangs- oder programmiersprachlich formulierte Regelfolge, die nach endlich vielen eindeutigen Schritten Probleme eines meist engen Bereichs beweisbar optimal löst. Beispiel: Simplex-Algorithmus zur Lösung eines Problems der linearen ›Optimierung. ›heuristisch

**Analyse:** entscheidungsvorbereitende Auswertung von Daten, die in der Regel Ergebnisse von ›Abfragen sind

**Analytischer Hierarchieprozess** (AHP, engl. Alytic Hierarchy Process): Verfahren der ›Nutzwertanalyse, welches die Entscheidungselemente “Kriterien” und “Alternativen” an einem einzigen Oberziel misst. Im Unterschied zu anderen nutzwertanalytischen Verfahren vergleicht AHP die Entscheidungselemente auch sukzessiv (paarweise) statt nur simultan (alle zusammen).

**Anomalie:** Unregelmässigkeit beim Datenentwurf, die das Einfügen, Löschen oder Fortschreiben von Tabelleneinträgen erschwert (Einfüge-, Lösch- und Fortschreibungsanomalien).

Die ›Normalisierung versucht Anomalien zu vermeiden, um die Wahrscheinlichkeit widersprüchlicher Daten bereits während des Datenentwurfs zu verringern.

**ANSI** (engl. American National Standards Institute): US-Gremium, das unter anderem informationsverarbeitende Hardware und Software standardisiert. ANSI hat zum Beispiel den ›ANSI-Code und ›SQL standardisiert. ›Industriestandard, ›ISO, ›ASCII, ›EBCDIC

**ANSI-Code:** vom ›ANSI standardisierter Zeichensatz unter dem Betriebssystem MS Windows. Die ersten 128 Zeichen des ANSI-Code entsprechen dem 7bit-›ASCII-Code. Die restlichen 128 Zeichen stellen Spezialzeichen dar, zum Beispiel Elemente nationaler Zeichensätze und Währungssymbole.

**Antwortzeit** (engl. response time): Zeit zwischen dem Versand einer ›Abfrage und dem Erhalt des *Gesamtergebnisses*. Im Gegensatz zu ›OLTP-Abfragen kann der Unterschied zwischen dem Erhalt des ersten Teilergebnisses einer Data Warehouse-Abfrage und dem letzten Teil beträchtlich sein.

**Anwendung:** ›Applikation

**API** (engl. Application Programmer Interface): Schnittstelle einer Programmbibliothek, die dem Programmierer erlaubt, seine ›Applikation mit einem Drittprogramm, zum Beispiel einem Datenbank- oder Betriebssystem, zu verbinden. Ein API beschreibt vor allem die Namen und Funktionen der Unterprogramme sowie die Zahl und den Typ ihrer Parameter und Rückgabewerte. Das Betriebssystem MS Windows besteht zum Beispiel aus ›DLLs, die auch API-Unterprogramme enthalten, die der Programmierer aufrufen kann. Weitere Beispiele sind ›ODBC und das MD-API des ›OLAP Council.

**Applet:** plattformunabhängiges kleines ›Java-Modul, das über das Web verteilt werden kann und vom Web Browser interpretiert wird

**Applikation** (dt. Anwendung): Programmsystem für eine bestimmte Gruppe von Anwendungen und Benutzern. Eine Applikation besteht in der Regel aus einem prozeduralen Teil (Programmlogik)

und einem deklarativen Teil (Datenstrukturen, ›Formulare, ›Berichte).

**Applikationsgenerator:** Programm, das auf Grund einer ›deklarativen Spezifikation des Benutzers einen Prototyp erstellt, der oft programmiersprachlich verfeinert werden kann

**Architektur, neuronale:** ›Modell, neuronales

**ASCII-Code** (engl. American Standard Code for Information Interchange): Code, der ausgewählte Zeichen eines Alphabets auf Bitfolgen abbildet. Die meisten dieser Zeichen lassen sich mit Ausgabegeräten wie Bildschirm und Drucker sichtbar machen, die übrigen codieren nicht direkt sichtbare Operationen. Der 7bit-ASCII-Code ist identisch mit den ersten 128 Zeichen des ›ANSI-Code. Unter DOS/Windows-Rechnern wird der ASCII-Code auch DOS- oder OEM-Code genannt. Wenn die direkte Konversion zwischen Daten verschiedener Anwendungen - zum Beispiel zwischen Datenbank- und Tabellenkalkulationsprogrammen - nicht möglich ist, wird oft ASCII-Text ausgetauscht. ›ANSI, ›EBCDIC

**Assoziationsanalyse:** ›Data Mining-Methode zur Entdeckung von Beziehungen (›Assoziationsregeln). Eine typische Anwendung ist die ›Marktkorbanalyse.

**Assoziationsregel** (engl. association rule): Regel, die eine Verbindung zwischen Datenbankeinträgen herstellt. Beispiel: WENN ein Kunde Bier kauft, DANN ist die Wahrscheinlichkeit, dass er auch Chips kauft 0.8.

**Assoziierung:** ›Assoziationsanalyse

**Attribut** (Spalte, Feld): (1) Eigenschaft eines Objekts oder (2) Element eines ›Satzes bzw. einer Tabellenzeile

**Aufgabe** (engl. task): Einheit der ›Parallelverarbeitung. Eine Aufgabe ist je nach Zusammenhang ein Teil einer Anweisung, eine gan-

ze Anweisung, ein Prozess aus mehreren Anweisungen, ein Job aus mehreren Prozessen oder sogar ein unabhängiges Programm.

**Ausreisser** (engl. outlier): sehr stark abweichender ›Attributwert, der manche Lage- und Verteilungsmasse unerwünscht beeinflusst. Man versucht deshalb mit statistischen Verfahren Ausreisser zu identifizieren und auszusondern.

**Auswahlabfrage:** ›Abfrage

**B-Baum:** Datenstruktur auf einem Externspeicher, die einen schnellen Zugriff auf sich ändernde Daten erlaubt. Ein B-Baum ermöglicht die Binärsuche, ohne dass neu eingefügte Daten zu einer unausgeglichenen Baumstruktur führen.

**Backend:** Programm (zum Beispiel eine ›Datenbankmaschine), das vor dem Endbenutzer verborgen bleibt. ›Frontend

**Bandbreite** (engl. bandwidth): Kapazität eines Kommunikationsmediums. Die Bandbreite wird oft in Dateneinheiten pro Zeiteinheit gemessen (zum Beispiel Bit pro Sekunde oder ›Gigabyte pro Stunde).

**Basistabelle** (physische Tabelle): permanente ›Tabelle, die - im Gegensatz zu einer ›Ergebnis- oder ›Viewtabelle - samt ihrer Beschreibung länger als eine Sitzung gespeichert bleibt

**bedingte Wahrscheinlichkeit:** ›Wahrscheinlichkeit, bedingte

**Befehlsschaltfläche** (engl. command button): ›Steuerelement, das unter einer grafischen Benutzeroberfläche auf einen Mausklick reagiert und eine ›Ereignisprozedur startet. Ein Klick auf eine Schaltfläche kann zum Beispiel ein ›Formular öffnen oder eine ›SQL-Anweisung ausführen.

**Benutzersicht** (engl. view): Teil einer ›Datenbank, der für einen bestimmten Benutzer(kreis) von Bedeutung ist



**Bericht** (engl. report): strukturierte und zusammenfassende Abfrage, deren Ergebnis sortiert und formatiert ausgegeben wird. Ein Bericht kann schriftlich oder bildschirmbasiert, ›ad hoc oder ›vorberechnet sein. Er unterscheidet sich von einer ›Abfrage i.e.S. durch die Breite der Frage, den Umfang, die Formatierung des Ergebnisses und meist durch die regelmässige Anwendung.

**Berichtsdatei** (engl. report file): Datei, die zur Erstellung eines ›Berichts aufbereitet wird

**Berichtsgenerator** (engl. report generator, report designer): Programm, das aus einer Spezifikation des Benutzers eine ›Berichtsdatei erstellt

**Beta-Koeffizient**: Verhältnis zwischen der Rendite eines Marktportefeuille und der Rendite einer einzelnen Aktie

**betriebliche Kennzahl**: ›Kennzahl

**Beziehung** (engl. relationship): Verbindung zwischen mehreren ›Objekttypen. Im ›Relationenmodell werden Beziehungen zwischen zwei ›Tabellen über ein ›Verbundattribut hergestellt, das in beiden Tabellen (eventuell unter verschiedenen Namen) vorkommt. Man unterscheidet zwischen drei Beziehungstypen:

Eine *1:1-Beziehung* ist eine Beziehung zwischen zwei Tabellen, in denen der ›Primärschlüsselwert jeder Zeile der einen Tabelle dem ›Fremdschlüsselwert *genau einer* Zeile der anderen Tabelle entspricht.

Eine *1:n-Beziehung* (auch 1:m-Beziehung genannt) ist eine Beziehung zwischen einer Haupttabelle (1:-Tabelle) und einer Detailtabelle (:n-Tabelle), wobei der ›Primärschlüsselwert jeder Zeile der Haupttabelle dem ›Fremdschlüsselwert *einer oder mehrerer* Zeilen der Detailtabelle entspricht.

Eine *m:n-Beziehung* ist eine Beziehung zwischen zwei Tabellen, wobei der ›Primärschlüsselwert jeder Zeile der einen Tabelle dem ›Fremdschlüsselwert mehrerer Zeilen der anderen Tabelle entspricht und umgekehrt. In ›relationalen Datenbanken muss eine dritte Tabelle, die *Verbindungstabelle*, die beiden Tabellen verknüpfen. Sie muss mindestens die Primärschlüssel der Ausgangstabellen enthalten.

**Beziehungsintegrität:** (engl. referential integrity):

- ›Integritätsbedingung, die verlangt, dass zu jedem
- ›Fremdschlüsselwert der einen ›Tabelle ein passender
- ›Primärschlüsselwert der verbundenen Tabelle existiert.
- ›Entitätsintegrität, 1:n-›Beziehung

**binärer Entscheidungsbaum:** Entscheidungsbaum, der nach einem Testattribut auf eine von *zwei* Antworten (von zwei Baumästen) verzweigt. ›n-ärer Entscheidungsbaum

**Bitmap-Index** (Bitmuster-Index, engl. bitmapped index): Codierung

- ›diskreter Attribute als Bitfolgen, die einen schnellen Zugriff auf Attributwerte ermöglichen

**Bitmuster-Index:** ›Bitmap-Index

**Bonitätsprüfung:** Prüfung der Kreditwürdigkeit (Bonität) einer natürlichen oder juristischen Person. Zur Bonitätsbeurteilung werden neben ›Kennzahlen auch ›Data Mining-Verfahren eingesetzt.

**Browsing:** benutzerfreundliche Abfrage mit Operationen wie

- ›Filtern, ›Slicing and Dicing, ›Drill Down and Up, ›Drill Through und
- ›Drill Across

**Business Process Reengineering (BPR):** Analyse von

- ›Geschäftsprozessen (engl. business processes) mit dem Ziel ihrer Verbesserung

**Cache:** ›Puffer



**CART** (engl. Classification And Regression Trees):

- ›Data Mining-Verfahren, das binäre ›Entscheidungsbäume erstellt.
- ›CHAID, ›ID3

**Cascade Correlation Neural Net:** ›CCN

**Cash Flow** (dt. Kapitalfluss): Kennzahl zur Beurteilung der (vergangenen) Ertrags-, Investitions- und Schuldendeckungskraft einer Unternehmung. Vereinfachend die Summe von Nettogewinn und Abschreibungen

**CCN** (Cascade Correlation Neural net): Netzmodell, das je nach Aufgabenstellung keine oder mehrere verborgene Schichten inkrementell aufbaut. CCN eignet sich wegen seiner Flexibilität gut für ›generische neuronale Werkzeuge.

**CGI:** ›Common Gateway Interface

**CHAID** (engl. Chi Square Automatic Interaction Detector):

- ›Data Mining-Verfahren, das mit ›Kontingenztafeln und dem ›Chi Quadrat-Test ›Entscheidungsbäume der Ordnung = 2 erstellt.
- ›CART, ›ID3

**Chi Quadrat-Test** (engl. chi square test): Goodness-of-fit-Test, der beobachtete und erwartete Häufigkeiten vergleicht, um zu testen, ob alle Häufigkeiten gleich gross sind bzw. wie gut sie auf ein vorgegebenes Muster passen

**CISC** (engl. Complex Instruction Set Computing): ›RISC

**Client:** ›Client/Server-System

**Client/Server-System** (C/S-System): Verbund von Hardware und Software, der Aufträge von Clients (Arbeitsplatzrechnern) durch Server (Dienstleistungsrechner) ausführt

**Client/Server-System, dreistufiges:** Rechnersystem, das Datenhaltung und -verarbeitung auf unterschiedliche, dedizierte Applikationsserver und die Interaktion auf Clients verteilt. ›zweistufiges Client/Server-System, ›Host-System, ›Master/Slave-System, ›Fileserver-System

**Client/Server-System, zweistufiges:** Rechnersystem, das die Datenhaltung und -verarbeitung auf einen Server und die Clients verteilt. Die Präsentation ist Aufgabe der Clients. ›dreistufiges Client/Server-System, ›Host-System, ›Master/Slave-System, ›Fileserver-System

**Clusteranalyse** (engl. cluster analysis): ›multivariates statistisches Verfahren, das beobachtete Objekte in möglichst homogene Gruppen (engl. clusters) einteilt. Dabei wird die Ähnlichkeit der Objekte innerhalb eines Clusters minimiert und die Unähnlichkeit zwischen Objekten verschiedener Cluster maximiert.

**Clustering:** (1) lose Koppelung von Einzelrechnern oder ›SMP-Systemen, welche die ›parallele Verarbeitung von ›Aufgaben erlaubt, ohne dass auch auf gemeinsame Internspeicher und Peripheriegeräte zugegriffen werden muss. (2) ›Clusteranalyse

**COM** (engl. Component Object Model): Objektmodell von Microsoft

**Common Gateway Interface (CGI):** Schnittstelle zwischen einem Web Server und einem Programm, das auf einen Datenbank- oder anderen ›Server zugreift. Ein CGI-Programm kann in einer beliebigen serverunterstützten Programmiersprache geschrieben sein (zum Beispiel C oder Visual Basic). CGI-Programme gehören zu den serverseitigen Erweiterungen von ›HTML.

**Cookie:** Nachricht eines Web Servers, die ein Browser in einer Textdatei speichert. Der Browser sendet diese Datei jedesmal an den Server, wenn er von ihm eine Webseite verlangt. Der Server kann so Webseiten an den Benutzer anpassen. Statt einer unper-

sönlichen Seite erhält der Benutzer zum Beispiel eine an ihn persönlich adressierte Seite.

**CPU** (engl. Central Processing Unit): Zentraleinheit, Prozessor

**Data Dictionary** (DD, Katalog, Datenlexikon, ․metadata repository, encyclopedia): Verzeichnis, das ․Metadaten zur Form und zum Inhalt von Anwendungsobjekten, insbesondere ․Datenbankobjekten, enthält (Datentyp, Grösse, Aufbau, Verwendung, etc.)

**Data Dictionary-System:** Informationssystem, das ․Metadaten in einem Data Dictionary so verwaltet, dass es möglichst viele Tätigkeiten der Softwareentwicklung und -wartung koordinieren kann. Ein *integriertes* Data Dictionary-System benutzt im Gegensatz zu einem *stand-alone* Data Dictionary-System ein allgemeines Datenbanksystem, zum Beispiel ein ․RDBMS.

**Data Mart** (DM): lokales ․Data Warehouse, das auf die Daten eines ․Funktionsbereichs, einer Abteilung, einer Arbeitsgruppe oder einer einzelnen Person beschränkt bleibt

**Data Mining:** nichttriviales “Schürfen” nach Wissen in Massendaten. Nichttrivial ist es, weil komplexe Methoden aus der mathematischen Statistik oder der ․Künstlichen Intelligenz genutzt werden. Idealerweise ist die Datenquelle ein ․Data Warehouse.

**Data Mining-Werkzeug:** Software, die Daten so vorbereitet, dass sie Ergebnisse mit ․Data Mining-Methoden berechnen und präsentieren kann

**Data Warehouse:** analytische ․Datenbank, die strategische ․Entscheidungen unterstützt, indem sie umfangreiche und regelmässige Auszüge aus ․Produktionsdatenbanken periodenbezogen und oft ․aggregiert Endbenutzern zur ․ad hoc-Analyse bereit stellt. Nach ihrem Umfang unterscheidet man zwischen ․Enterprise Data Warehouses und ․Data Marts.

**Data Warehouse, zentrales:** ›Data Warehouse, das  
›Produktionsdaten sammelt, transformiert und dann koordiniert an  
›Data Marts verteilt

**Datenbank:** Sammlung von ›Datenbankobjekten, zwischen denen  
›Beziehungen bestehen. Oft teilt man die Daten in Benutzer- und  
Systemdaten ein. Die Systemdaten (›Metadaten) enthalten Defini-  
tionen von ›Datenbankobjekten wie ›Basistabellen, ›Viewtabellen,  
Synonymen, ›Wertebereichen, ›Indizes, Benutzern und Benutzer-  
gruppen.

**Datenbankmaschine** (engl. data base engine): Kern des  
›Datenbanksystems, der nur Datenbankaufgaben im engeren Sinn  
erledigt. Er enthält keine Benutzerschnittstelle und keine Entwick-  
lungs- und Dienstprogramme.

**Datenbankobjekt:** Teil einer ›Datenbank, den der Benutzer ändern  
kann. Beispiele von Datenbankobjekten sind ›Tabellen, ›Abfragen,  
›Formulare, ›Berichte, Programme, ›Integritätsregeln, Synonyme,  
›Datentypen und ›Indizes.

**Datenbankstrukturdiagramm:** Graph, der ›Beziehungen zwischen  
den ›Tabellen einer ›Datenbank ›visualisiert. Eine verbreitete Form  
von Datenbankstrukturdiagrammen sind ›Entity-Relationship-Dia-  
gramme.

**Datenbanksystem** (engl. Data Base Management System,  
DBMS): Gruppe von Programmen zur Verwaltung von  
›Datenbanken

**Datenbankverwaltungssystem:** ›Datenbanksystem

**datengetriebene Methode:** rechenintensive Data Mining-Methode,  
die *ohne* gezielte Hypothese oder einschränkendes Modell eine  
Stichprobe aus einer grossen Datenmenge zieht und Ergebnisse

ableitet. Beispiele: ›Regelinduktion, ›neuronale Netze. Gegensatz:  
›modellgetriebene Methoden. ›explorativ

**Datenmodell, hierarchisches:** Beschreibung hierarchischer Beziehungen zwischen Datensätzen. Grundlage älterer  
›Datenbanksysteme wie ›IMS. Nachteilig wirkt sich die geringe Flexibilität bei Datenbankabfragen und -änderungen (Datenabhängigkeit) aus.

**Datenmodell, konzeptionelles** (engl. conceptual data model): Vorstufe ›logischer und ›physischer Datenmodelle. Im Gegensatz zum physischen Datenmodell bietet das konzeptionelle eine benutzerorientierte Sicht einer Datenbank. Oft fällt ein konzeptionelles Datenmodell - im Gegensatz zum logischen - auch keine detaillierten Entscheidungen zum Datenentwurf (zum Beispiel, ob relational oder normalisiert). Der Text unterscheidet nicht zwischen konzeptionellen und logischen Datenbeschreibungen.

**Datenmodell, logisches:** abstrakte Beschreibung der Daten von ›Geschäftsprozessen, die im Gegensatz zu einem ›physischen Datenmodell möglichst wenig Rücksicht auf die Art und Weise der Implementierung nimmt. Eine verbreitete Notation zur Beschreibung logischer Datenmodelle ist das ›Entity-Relationship-Diagramm.

**Datenmodell, physisches:** konkrete Datenbeschreibung, die Rücksicht auf die ›physische Implementierung der Datenstrukturen auf Externspeichern (und damit auf konkrete Software) nimmt. Das physische Datenmodell spielte eine wichtige Rolle für das Leistungsverhalten (engl. performance) eines DBMS.

**Datentyp:** Paar aus ›Wertebereich und Zugriffsoperationen, die sich auf dem Wertebereich ausführen lassen

**DB:** ›Datenbank

**DB2:** ›RDBMS von IBM

**dBASE:** ältestes PC-Datenbanksystem. Das Dateiformat und die Programmiersprache von dBASE haben sich zum Quasistandard XBase entwickelt, den viele Datenbankprodukte mehr oder weniger streng befolgen.

**DBMS** (engl. data base management system): ›Datenbanksystem

**DCL** (engl. Data Control Language): Teilmenge der Anweisungen einer Datenbanksprache, welche die Sicherheit und den Schutz der Daten gewährleistet. In ›SQL gehören dazu die Anweisungen “lock”, “commit”, “rollback”, “grant “ und “revoke”. ›Transaktion, ›DDL, ›DML

**DD:** ›Data Dictionary

**DDL** (engl. Data Definition Language): Datenbeschreibungsanweisungen einer Datenbanksprache. In ›SQL gehören dazu die Anweisungen “create”, “alter”, “drop table”, “view “ und “index”. ›DCL, ›DML

**deklarativ** (nichtprozedural): beschreibend. ›Abfragesprache

**deklarative Abfragesprache:** ›Abfragesprache

**Denormalisierung:** teilweise oder völlige Rücknahme der ›Normalisierung aus Gründen der Abfrageeffizienz oder Benutzerfreundlichkeit

**deskriptive Entscheidungstheorie:** ›Entscheidungstheorie

**deskriptive Statistik:** ›Inferenzstatistik

**deterministische Entscheidung:** ›Entscheidung

**Dimension** (engl. dimension): meist ›symbolisches und ›diskretes ›Attribut, das die Auswahl, Zusammenfassung und Navigation eines ›Indikators erlaubt. Beispiele: Region, Produkt, Periode



**Direct Mailing:** Form der ›Direktwerbung, die auf der Basis eigener oder gekaufter Adressen eine Zielgruppe auswählt und deren Mitglieder einzeln anschreibt

**Direktwerbung:** Werbemaßnahmen, die den Empfänger mit einem selbständigen Werbemittel direkt (ohne vermittelnde Partner) ansprechen. Beispiele: Werbebrief, Prospekt oder Katalog, Reaktionskarte

**diskretes Merkmal:** ›Attribut mit aufzählbar vielen Werten. Gegensatz kontinuierliches (stetiges) Merkmal

**Diskriminanzanalyse** (engl. discriminant analysis): ›multivariates statistisches Verfahren zur ›Klassifikation von Objekten in zwei oder mehr Klassen. Die Klassifikation erfolgt mit einer Diskriminanzfunktion, in welche die beobachteten Werte mehrerer diskriminierender Variablen eingehen.

**DLL** (engl. Dynamic Link Library): Unterprogrammbibliothek, deren Komponenten erst zur Laufzeit gebunden werden (engl. dynamically linked). Anders als ein statisch gebundenes Unterprogramm ist ein dynamisches nicht permanenter Teil des Anwendungscode, sondern wird erst bei Bedarf geladen. Weil der Anwendungscode nur einen Verweis auf das DLL-Unterprogramm enthält, müssen Unterprogramme auch dann nur einmal in den Hauptspeicher gehalten werden, wenn sie von verschiedenen Anwendungen aufgerufen werden. Wichtige Beispiele von DLLs sind ›ActiveX-Komponenten.

**DM:** ›Data Mart

**DML** (engl. Data Manipulation Language): Such- und Änderungsanweisungen einer Datenbanksprache. In ›SQL gehören die Anweisungen “select”, “insert”, “delete “ und “update” zur DML-Teilsprache. ›DCL, ›DDL

**DOLAP** (engl. Desktop OLAP): ›MOLAP- oder ›ROLAP-System, das ›mehrdimensionale Daten auf einem ›Client (in der Regel einem PC) speichert

**dreistufiges Client/Server-System:** ›Client/Server-System, dreistufiges

**Drill Across:** ›Drill Down and Up über mehrere ›Data Marts hinweg

**Drill Down and Up:** Zusammenfassung und Detaillierung nach einer ›Dimension

**Drill Through:** Zugriff auf ›Transaktionsdetails aus einem ›OLAP-Werkzeug heraus. Der Zugriff erfolgt aus der gewohnten OLAP-Umgebung (ohne dass der Endbenutzer vordefinierte ›SQL-Anweisungen zu Gesicht bekommt).

**Drittgenerationssprache:** ›Viertgenerationssprache

**DSS** (engl. Decision Support System):  
›entscheidungsunterstützendes System

**dünn besetzte Matrix** (auch dünn besiedelte Matrix): Matrix mit vielen ›Nullwerten. Eine dünn besetzte Matrix entsteht zum Beispiel dann, wenn für einen ›Indikator *Zahl der Unternehmungen* die ›Dimensionen *High Tech-Branche* und *Land* verglichen werden. Bei allen Ländern ohne High-Tech-Branche erscheinen dann Nullwerte. Die effiziente Speicherung dünn besetzter Matrizen ist eine wichtige Aufgabe ›mehrdimensionaler ›DBMS.

**Durchsatz:** Mass für die Leistung eines Rechensystems oder einer Systemkomponente. Der Durchsatz kann als Zahl der möglichen Aufträge (engl. jobs) oder ›Transaktionen pro Zeiteinheit gemessen werden.

**durchschnittlicher absoluter Fehler:** ›Fehler, durchschnittlicher absoluter

**DW:** ›Data Warehouse

**e:** Basis des natürlichen Logarithmus  $\ln$ .  $e$  hat den Näherungswert 2,71828. Statt  $e$  schreibt man auch  $\exp(1)$  und statt  $e^x$  auch  $\exp(x)$ .  $e$  ist definiert als Grenzwert von  $(1 + 1/n)^n$ .

**EBCDIC** (engl. Extended Binary Coded Decimal Interchange Code): 8-bit-Code für ältere IBM- und kompatible Rechner. ›ASCII, ›ANSI

**Ebene** (engl. layer): Zusatztabelle für jeden Wert einer vierten Dimension eines Hyperwürfels in einem OLAP-Werkzeug

**EDW:** ›Enterprise Data Warehouse

**Eigenvektor:** ›Eigenwert

**Eigenwert:** Zahl  $\lambda$ , für welche die Gleichung  $\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$  nicht triviale Lösungen ( $\mathbf{x} \neq 0$ ) hat. Eine Lösung für  $\mathbf{x}$  heisst Eigenvektor zum Eigenwert  $\lambda$ .

**eingebettetes SQL:** ›SQL, eingebettetes

**einstufiges neuronales Netz:** ›neuronales Netz, einstufiges

**Eigenkapitalrentabilität** (engl. Return On Equity, ROE): Gewinn / Eigenkapital

**Endbenutzerwerkzeug:** Programm, das auch von einem Anwender ohne Programmierkenntnisse einfach bedient werden kann. Beispiele sind Textverarbeitungs- und Tabellenkalkulationspakete. Der Übergang von Endbenutzerwerkzeugen zu Programmierwerkzeugen ist fließend, weil die meisten Endbenutzerwerkzeuge eine eingebaute Makro- oder Programmiersprache - zum Beispiel ›Visual Basic für Applikationen - enthalten.

**Enterprise Data Warehouse (EDW):** ›zentrales Data Warehouse oder Verbund von ›Data Marts, das/der auf die ganze Unternehmung zugreift

**Enterprise Data Warehouse, hierarchisches:** ›zentrales Data Warehouse, das hierarchisch gegliederten ›Data Marts aufbereitete Daten bereit stellt

**Enterprise Data Warehouse, koordiniertes:** virtuelles ›Data Warehouse aus den koordinierten ›Data Marts der ganzen Unternehmung

**Entität** (engl. entity): konkretes oder abstraktes Objekt eines Informationssystems. Im ›Entity-Relationship-Datenmodell werden Objekte zu Klassen (Entitätstypen) zusammengefasst und zusammen mit ihren Beziehungen analysiert.

**Entitätsintegrität** (engl. entity integrity): ›Integritätsbedingung, die verlangt, dass jede Tabellenzeile durch einen Primärschlüssel eindeutig bestimmt ist

**Entity-Relationship-Datenmodell** (engl. entity relationship data model): ›konzeptionelles bzw. ›logisches Datenmodell zur formalisierten Darstellung der realen Welt. Es besteht aus Objekten (Entitäten) und Beziehungen (engl. relationships). Eigenschaften beschreiben Objekte und ›Beziehungen. Zum ursprünglichen Vorschlag von Chen existiert eine Vielzahl von Varianten.

›Entity-Relationship-Diagramm

**Entity-Relationship-Diagramm:** Notation zur Beschreibung ›konzeptioneller bzw. ›logischer Datenmodelle. In einem Entity-Relationship-Diagramm werden Objekte oft durch Rechtecke und Beziehungen durch Pfeile oder Rauten dargestellt.

›Entity-Relationship-Datenmodell, ›Datenbankstrukturdiagramm

**Entropie** ("Unordnung"): in der ›Regelinduktion ein Mass für die Information, die eine Klassifikation ohne Testattribute benötigen würde. ›ID3, ›Informationsgehalt, ›Informationstheorie

**Entscheidung**, betriebliche: rationale Wahl, die ein Entscheidungsträger aus betrieblichen Alternativen in einer gegebenen Umwelt trifft. ›entscheidungsunterstützendes System

**Entscheidungsbaum** (engl. decision tree): hierarchische Darstellung mehrstufiger ›Entscheidungen. Varianten des Begriffs finden sich in der statistischen ›Entscheidungstheorie, der Unternehmensforschung (›Operations Research) und der Künstlichen Intelligenz (›KI, vgl. auch ›Expertensysteme und ›Induktion). Für die KI ist ein Entscheidungsbaum eine Hierarchie mit den folgenden Eigenschaften: Der Wurzel und jedem Zwischenknoten ist ein Test zugeordnet, der aufgrund der Knoteneingabe (der relevanten Umweltbedingungen) eine der ausgehenden Kanten (Alternativen, Aktionen) wählt. Jedem Endknoten (Blatt) ist eine Entscheidung (Folgerung, Klasse) zugeordnet. Die Folgerung kann deterministisch ("Wetter ändert") oder probabilistisch ("Wetter ändert mit einer Wahrscheinlichkeit von 0.7") sein.

**Entscheidungstheorie**: (1) Die *deskriptive* Entscheidungstheorie untersucht, wie Menschen ›Entscheidungen tatsächlich treffen. (2) Die *präskriptive* Entscheidungstheorie untersucht hingegen, wie Menschen oder Maschinen vorgehen müssen, um zu optimalen (rationalen) Entscheidungen zu kommen.

**entscheidungsunterstützendes System** (EUS, engl. decision support system): computergestütztes Planungs- und Informationssystem, das Endbenutzern aller Entscheidungsebenen beim Lösen komplexer Probleme hilft, indem es Information verdichtet und darstellt. ›Entscheidungsunterstützende Systeme lassen sich nach den folgenden Kriterien einteilen:

(1) *Entscheidungsträger*: Eine Entscheidung kann von einem einzelnen oder einer Gruppe gefällt werden (Individual- oder Kollektiventscheidung).

(2) *Reichweite*: Entscheidungen können ›operativ oder ›strategisch sein.

(3) *Sicherheit*: Deterministische Entscheidungssituationen führen unter einem Ziel sowie bekannten und sicher eintreffenden Umweltbedingungen zu einer sicheren Entscheidung. Objektiv bzw. subjektiv probabilistische Entscheidungssituationen ordnen hingegen Umweltbedingungen objektive (empirische) bzw. subjektive Wahrscheinlichkeiten zu.

(4) *Komplexität*: Wohlstrukturierte Entscheidungen gehen im Gegensatz zu schlecht strukturierten von einer überblickbaren Zahl bekannter Entscheidungsvariablen und Beziehungen aus.

(5) *Programmierbarkeit* (Automatisierungsgrad): Programmierbare Entscheidungsprozesse lassen sich im Gegensatz zu den nichtprogrammierbaren ›algorithmisch oder ›heuristisch darstellen. Oft synonym gebraucht werden die Termini *Management-Informationssystem* (MIS), *Führungsinformationssystem* (FIS) bzw. *Executive Information System* (EIS).

**Entwurfsanomalie**: ›Anomalie

**Ereignisprozedur**: Prozedur, die durch ein Ereignis - zum Beispiel einen Klick auf eine ›Schaltfläche - aufgerufen wird

**Erfolgsrechnung**: neben der Bilanz der wichtigste Teil des ›Jahresabschlusses. Die Erfolgsrechnung ermittelt aus der Gegenüberstellung von Aufwand und Ertrag den Erfolg (Gewinn oder Verlust).

**Ergebnistabelle** (engl. answer table): temporäre Tabelle, welche die Ergebnisse einer ›Abfrage enthält



**Erklärung:** ›Wissensbasierte Systeme, insbesondere ›regelbasierte Systeme unterscheiden folgende Erklärungen:

(1) Eine *Wie*-Erklärung begründet, weshalb eine Antwort gilt. Sie ist oft die Ausgabe eines Teils des Lösungsbaums (des Lösungsprotokolls) zwischen der Benutzerfrage und der Antwort. Sie präsentiert also top down (vom Ausgangsziel bis zur Lösung) Regeln und Fakten, die zur Antwort führen.

(2) Eine *Warum*-Erklärung begründet, weshalb eine bestimmte Frage an den Benutzer gestellt wird. Sie ist oft die Ausgabe eines Teils des Suchbaums zwischen der Frage an den Benutzer und dem Ziel der Konsultation. Die Erklärung präsentiert also bottom up (von der Frage bis zum Ausgangsziel zurück) Regeln und Fakten, die zur Frage geführt haben.

(3) Eine *Was-Wenn*-Erklärung präsentiert hypothetische Lösungen unter veränderten Bedingungskonstellationen (engl. hypothetical reasoning).

(4) Eine *Warum-nicht*-Erklärung begründet, weshalb eine bestimmte Regelbedingung nicht zutrifft oder eine Frage abschlägig beantwortet wurde. Das ›Expertensystem gibt zum Beispiel falsifizierte Regeln aus.

(5) Eine *Wozu*-Erklärung begründet, zu welchem Zweck ein bestimmtes Vorgehen gewählt wird.

**Erwartungswert:** (1) in der Statistik ein Mass für die Verteilung einer Zufallsvariablen  $X$ . Sind  $x_i$  die Ausprägungen einer *diskreten* Zufallsvariablen  $X$  und  $f(x_i)$  die jeweiligen Wahrscheinlichkeiten, so ist  $E(X) = \sum x_i f(x_i)$  der Erwartungswert von  $X$ . Für eine *kontinuierliche* Zufallsvariable gilt die entsprechende Integral-Definition. (2) in der ›Regelinduktion (›ID3) ein Mass für die Information, die eine Klassifizierung mit einem bestimmten Testattribut benötigt.

**EUS:** ›Entscheidungsunterstützendes System

**Executive Information System (EIS):** ›entscheidungsunterstützendes System

**Expertensystem:** Software, die Methoden der ›Künstlichen Intelligenz anwendet, indem sie (1) Wissen auf einem eng begrenzten Gebiet problemangepasst, änderungsfreundlich und verarbeitungseffizient darstellt, (2) ›algorithmisch oder ›heuristisch Schlüsse daraus zieht und (3) diese Schlüsse unter Bezug auf Falldaten im Dialog mit dem Benutzer erklärt. Expertensysteme sind meist ›regelbasiert.

**explorativ** (engl. exploratory, entdeckend): Explorative Verfahren sind im Gegensatz zu ›hypothesentestenden Methoden datengetrieben. Ausgangspunkt ist nicht eine Hypothese, die aus einem Modell abgeleitet worden ist sondern eine Datengesamtheit, deren Strukturen und Abweichungen beschrieben und allenfalls verallgemeinert werden sollen. ›explorative Datenanalyse

**explorative Datenanalyse** (engl. exploratory data analysis): deskriptive statistische Verfahren zur Aufdeckung von Datenstrukturen und von Abweichungen von einer Referenzgesamtheit. Die Modellannahmen der ›explorativen Datenanalyse sind weniger streng als jene der schliessenden Statistik. Sie vermeidet zum Beispiel einschränkende Anforderungen an das ›Skalenniveau (zum Beispiel ›verhältnisskalierte Daten) oder die Verteilung der Daten (zum Beispiel die Normalverteilungsannahme). Verbreitet sind vor allem graphische Verfahren, Variablentransformationen und ›robuste Kenngrößen (zum Beispiel der Median statt des arithmetischen Mittels).

**F-Test:** Test der Nullhypothese, dass die Aufnahme einer zusätzlichen unabhängigen Variable nicht eine ›signifikante Zunahme von  $R^2$  bewirkt ( $R$  ist der Korrelationskoeffizient einer ›Mehrfachregression)

**Fakt** (engl.fact): ›Indikator

**Faktorenanalyse** (engl. factor analysis): ›multivariates statistisches Verfahren, das versucht, Faktoren zu finden, die  
›Korrelationsmuster beobachteter Variablen erklären. Ziel sind wenige Faktoren, die einen möglichst grossen Anteil der ›Varianz einer viel grösseren Menge von Variablen erklären. Nachteile sind einschränkende Annahmen und die schwierige Interpretation der gefundenen Faktoren. Anwendungen finden sich etwa in der Psychologie, der Soziologie und im Marketing

**Fehler, durchschnittlicher absoluter**: ›Validitätsmass für  
›neuronale Netze, das die durchschnittliche absolute Abweichung von tatsächlichen und berechneten Ausgaben ermittelt

**Fehlertoleranz** (engl. fault tolerance): Fähigkeit eines Systems, auch nach einem teilweisen Ausfall seiner Komponenten die verlangte Leistung ganz oder teilweise zu erfüllen. Dies kann vor allem durch den Einsatz redundanter Komponenten erreicht werden (zum Beispiel durch Stand by-›Server, ›Plattenspiegelung oder ›RAID).

**Feld** (engl. field, item): ›Attribut

**Fileserver-System**: Rechnersystem, das Arbeitsplatzrechner so verbindet, dass ein Dienstleistungsrechner (Server) Drucker-, Speicher- und Dateiaustausch-Funktionen, aber keine höheren Verarbeitungsfunktionen übernimmt. ›Master/Slave-System, ›Host, ›zweistufiges Client/Server-System, ›dreistufiges Client/Server - System

**Filtern**: Auswahloperation (›Selektion) in ›Abfragesprachen und ›OLAP-Systemen

**flache** Datei (engl. flat file): Datei, die nicht in ein ›DBMS oder ein anderes strukturiertes Zugriffssystem integriert ist

**Formular** (Bildschirmformular, engl. form): Bildschirmfenster, das  
›Steuerelemente enthält, mit denen der Benutzer Operationen wie  
Anzeigen und Editieren durchführen kann. Wichtige Steuerelemen-  
te von Datenbank Anwendungen sind Navigationsschaltflächen, mit  
denen der Benutzer die Position der laufenden Zeile oder des lau-  
fenden ›Felds ändert.

**Fremdschlüssel** (engl. foreign key): Attribut, das in einer anderen  
›Tabelle ›Identifikationsschlüssel ist. Ein Fremdschlüssel ermöglicht  
›Beziehungen zu anderen Tabellen.

**FTP**: engl. File Transfer Protocol

**Frontend**: Vom ›Backend (der ›Datenbankmaschine und der Daten-  
bank) benutzernaher Anwendungsteil eines Datenbanksystems. Das  
Frontend übernimmt in der Regel die Eingabe und Ausgabe.

**funktionale Abhängigkeit**: ›Abhängigkeit, funktionale

**Funktionsbereich**: Eine Unternehmung wird oft nach den wichtig-  
sten Unternehmungsfunktionen in die Bereiche “Forschung und  
Entwicklung”, “Beschaffung”, “Produktion” und “Absatz “ eingeteilt  
(Funktionalorganisation).

**Gap-Analyse** (dt. Lückenanalyse): Methode des strategischen Ma-  
nagements, die Abweichungen zwischen zukünftigen Entwicklungs-  
verläufen des Geschäfts grafisch darstellt

**Gbyte**: Abkürzung für Gigabyte (1024 Megabyte bzw.  $10^9$  Bytes)

**Genauigkeit** (engl. accuracy): ›Trefferquote

**Generator**: ›Applikationsgenerator, ›Berichtsgenerator

**generisch**: allgemein, Gegenteil von dediziert

**Geschäftsprozess** (Geschäftsvorgang, engl. business process):  
Folge von geschäftlichen Vorgängen, die in die Entwicklung eines

betrieblichen Informationssystemen oder in die Simulation einer betrieblichen Aktivität eingeht. Ein Beispiel ist die Abwicklung eines Kundenauftrags von der Bestellung bis zur Lieferung des Produkts. Wichtige Klassen von Geschäftsprozessen sind Beschaffungsprozesse, Produktionsprozesse, Verteilungsprozesse und Informations- und Kommunikationsprozesse. ▶ Funktionsbereich

**gespeicherte Prozedur:** ▶ Prozedur, gespeicherte

**GIF** (engl. Graphics Interchange Format): komprimiertes Bitmap-Format, insbesondere für den Datenaustausch auf ▶ WANs (zum Beispiel dem ▶ World Wide Web)

**Gigabyte:** ▶ GByte

**Granularität** (engl. granularity): Grad der Detaillierung bzw. Zusammenfassung in einem Data Warehouse. Eine hohe Granularität bedeutet einen hohen Speicheraufwand.

**Grossrechner:** ▶ Mainframe

**Grundgesamtheit** (engl. population): Menge aller Elemente, auf die ein statistisches Ergebnis verallgemeinert werden soll. Wichtig ist, dass die Grundgesamtheit sachlich, räumlich und zeitlich genau definiert wird. Die ▶ Inferenzstatistik überträgt die Ergebnisse einer ▶ Stichprobe auf die Grundgesamtheit.

**Gültigkeit:** ▶ Validität

**GUI** (engl. Graphical User Interface): graphische Benutzerschnittstelle

**Hashfunktion** (Adressberechnungsfunktion): Funktion, die breit gestreute ▶ Schlüsselwerte auf einen schmalen Bereich von Zellen abbildet. Die Hashfunktion ist eine effiziente Zugriffsmethode für grosse Datenbestände.

**heuristisch:** Ein heuristisches Verfahren ist eine umgangs- oder programmiersprachlich formulierte Regelfolge, die nach endlich vielen Schritten Probleme eines meist breiteren Bereichs suboptimal, aber im allgemeinen brauchbar löst. ›algorithmisch

**hierarchisches Data Warehouse:** ›Data Warehouse, hierarchisches

**hierarchisches Datenmodell:** Datenmodell, hierarchisches

**Host:** ›Mainframe

**Host-System:** Rechnersystem aus einem “intelligenten” Grossrechner (›Mainframe) und “dummen” Terminals.  
›Master/Slave-System

›**HTML:** ›Hypertext Markup Language

**HTTP:** ›Hypertext Transport Protocol

**Hypertext Markup Language (HTML):** Standard zur Erstellung von Dokumenten für das ›World Wide Web (WWW). Der HTML-Zeichenstrom wird sowohl vom Web Server als auch vom Browser (›Client) decodiert. HTML regelt in erster Linie die Form einer Webseite. Erweiterungen wie ›XML spezifizieren auch inhaltliche Aspekte.

**Hypertext Transport Protocol (HTTP):** Standard zur ›HTML-Kommunikation zwischen Web Server und ›Client

**hypothesentestend** (engl. verification driven): ›modellgetriebene statistische Methode

**ID3** (engl. Iterative Dichotomiser 3): eines von mehreren  
›Data Mining-Verfahren, die ›Entscheidungsbäume induzieren.  
›CHAID, ›CART

**Identifikationsschlüssel:** ›Primärschlüssel



**importierte Daten:** Gegenteil von ›vorberechneten Daten

**IMS** (engl. Information Management System): weit verbreitetes, älteres Datenbanksystem von IBM mit einem ›hierarchischen Datenmodell

**Indikator** (Fakt, engl. business measure): ›aggregierbares (in der Regel numerisches und ›kontinuierliches) ›Attribut, das die ›mehrdimensionale Messung eines betrieblichen Erfolgskriteriums erlaubt. Beispiele: Umsatz, ›Cash Flow

**Induktion:** automatische Ableitung eines ›Entscheidungsbaums bzw. von ›Regeln aus Attributwerttupeln. Bei der Entwicklung von Expertensystemen ist die Induktion oft eine Alternative zur Expertenbefragung.

**induktive Statistik:** ›Inferenzstatistik

**Industriestandard** (engl. de facto standard): Ein *formaler* Standard wird von einem nationalen oder internationalen Standardisierungsgremium, zum Beispiel ›ANSI oder ›ISO, erlassen. Ein *Industriestandard* entsteht hingegen, wenn ein Produkt oder eine Regelung so populär wird, dass es/sie nicht nur vom Hersteller sondern auch von vielen anderen Anbietern unterstützt oder sogar kopiert wird. Beispiele sind der IBM-PC, ›TCP/IP und ›dBASE.

**Inferenzstatistik** (induktive Statistik): Gesamtheit mathematisch-statistischer Fragestellungen, welche die Ergebnisse einer ›Zufallsstichprobe auf ihre Übereinstimmung mit der ›Grundgesamtheit prüfen

**Informationsgehalt:** domänenunabhängiges syntaktisches Mass für den Gehalt einer Information. Der Informationsgehalt wird operationalisiert als dualer Logarithmus der Zahl der Binärentscheidungen, die zur Darstellung der Information erforderlich sind.  
›Informationstheorie, ›Entropie

**Informationstheorie:** wahrscheinlichkeitstheoretische Theorie der Speicherung, Umformung und Übermittlung von Information.  
▸ Entropie, ▸ Informationsgehalt

**Index:** Tabelle, die für jeden vorkommenden Wert eines ▸ Attributs, zum Beispiel des Geschlechtsnamens eines Mitarbeiters, die Adresse seiner Tabellenzeile angibt. Ein Index ist in der Regel für den Endbenutzer unsichtbar und beschleunigt vor allem das Suchen und Sortieren von Datensätzen.

**Inkonsistenzverhältnis** (engl. inconsistency ratio): ▸ Kennzahl in ▸ AHP, welche das Ausmass der Inkonsistenz mehrerer Paarvergleiche schätzt

**Integrität:** Richtigkeit und Vollständigkeit eines Datenbestandes. Die Integrität kann zum Beispiel durch die Eingabe von Werten ausserhalb des ▸ Wertebereichs eines Attributs (zum Beispiel 13 für eine Monatszahl) oder die fehlende Fortschreibung geänderter Werte (zum Beispiel des Wechsel des Zivilstands) verletzt werden. Ein Teil der Integritätsverletzungen lässt sich durch einen sorgfältigen Datenentwurf, insbesondere eine ▸ Normalisierung, vermeiden. Die restlichen Integritätsbedingungen sollten nach Möglichkeit zentral (auf der Ebene des ▸ DBMS) statt lokal (in jeder ▸ Applikation) geprüft werden. ▸ Anomalie, ▸ Entitätsintegrität, Beziehungsintegrität

**Integritätsbedingung** (Integritätsregel): Bedingung, welche hilft, die ▸ Integrität einer Datenbank zu gewährleisten

**Internet:** weltweites öffentliches Netz aus Tausenden von Teilnetzen und Millionen von Teilnehmern, die fehlertolerant kommunizieren und Dienste wie elektronische Post und ▸ World Wide Web teilen

**Internet-Dienste:** Dazu gehören vor allem die elektronische Post (E-Mail), Publikationsdienste (vor allem ▸ World Wide Web), Diskussionsdienste (News Groups), Datentransfer (zum Beispiel ▸ FTP) so-

wie Bestellung und Bezahlung von Waren (zum Beispiel Pay-Per-View Publishing).

**Internet Protocol (IP):** Teil des ›TCP/IP-Protokolls

**Internet Server Application Programming Interface (ISAPI):**  
›proprietäre ›API-Schnittstelle zwischen einem Microsoft-Web Server und einer Applikation, die auf einen Datenbankserver oder eine andere ›Serverapplikation zugreift

**Inervallskala:** ›Skalenniveau

**intervenierende Variable:** Variable, die sich einer direkten Beobachtung entzieht, aber indirekt gemessen werden kann

**IP:** ›Internet Protocol

**ISAPI:** ›Internet Server Application Programming Interface

**ISDN** (Integrated Services Digital Network): öffentliches Netz, das mehrere Telecom-Dienste integriert (vor allem Telefon, Fax und den Austausch von Computerdaten)

**ISO** (engl. International Standards Organisation): internationales Gremium, das unter anderem informationsverarbeitende Hardware und Software standardisiert. ›Industriestandard, ›ANSI

**IT:** engl. Information Technology

**Jahresabschluss:** Bericht über den Erfolgs und das Vermögen einer Unternehmung im abgeschlossenen Geschäftsjahr. Der Jahresabschluss besteht in der Regel aus der Bilanz (Vermögensrechnung), der ›Erfolgsrechnung und der ›Kapitalflussrechnung.

**Java:** objektorientierte Programmiersprache von SUN Microsystems, welche unter anderem die Kommunikation auf dem ›World Wide Web unterstützt

**JavaScript:** von ›Java abgeleitete Skriptsprache, die in ›HTML eingebettet werden kann und die Interaktion mit WWW-Browsern unterstützt. ›VBScript

**JCL:** ›Job Control Language

**Job Control Language:** Skriptsprache der ›Mainframes von IBM zur Steuerung von Aufgaben (engl. jobs) des Betriebssystems. ›MVS

**Kapitalflussrechnung:** Neben der Bilanz und der ›Erfolgsrechnung kann der ›Jahresabschluss eine Kapitalflussrechnung enthalten. Während die traditionellen Teilrechnungen vor allem der Kontrolle und Planung von Gewinnzielen dienen, zeigt die gebräuchlichste Form der Kapitalflussrechnung die Ursachen einer bestimmten Stufe der Liquidität (flüssige Mittel).

**Kategorie:** einer von mehreren ›Wertebereichen einer ›Dimension. Eine Dimension kann mehrere flache oder hierarchische Kategorien haben. Zur Dimension “Produkt” gehören zum Beispiel die flachen Kategorien Name, Farbe und Grösse. Hierarchische Kategorien der Dimension “Zeit” sind zum Beispiel Jahr, Quartal, Monat und Woche.

**kardinal:** ›Skalenniveau

**Kardinalität:** Zahl der Elemente einer Menge, insbesondere Zahl der möglichen Werte eines ›diskreten ›Attributs, einer ›Kategorie oder Zahl der an einer ›Beziehung beteiligten Entitäten

**Kennzahl** (engl. ratio, dt. auch Kenn”ziffer”): Messwert für betriebliche Bestandes- und Bewegungsgrössen. Kennzahlen erhalten ihre Aussagekraft vor allem durch den Vergleich verschiedener Abrechnungsperioden und Betriebe (Perioden- und Betriebsvergleich, innerbetriebliche und zwischenbetriebliche Kennzahlen).

**KI:** ›Künstliche Intelligenz

**Klassifikation:** Einteilung von Individuen in bekannte ›diskrete Klassen. Verbreitete Verfahren sind ›Entscheidungsbäume, ›Diskriminanzanalyse und ›neuronale Netze. ›Vorhersage

**Klassifikationsgewinn:** Gewinn, den eine Klassifizierung mit einem bestimmten Testattribut gegenüber einer Klassifizierung ohne Testattribut erzielt

**Kohonen-Netz:** sich selbst organisierendes ›unüberwacht lernendes ›neuronales Netz, das vor allem für das ›Clustering von Eingabedaten verwendet wird

**Kompression:** Verdichtung von Daten. Ein Kompressionsverfahren konvertiert nach einer Abbildungsvorschrift die Quelldaten in ein Komprimat, das wesentlich weniger Speicherplatz erfordert. Kompressionsverfahren werden vor allem von Betriebssystemen, von Datenbanksystemen und von multimedialen Anwendungen bereit gestellt.

**Kontingenztafel** (engl. contingency table): ›Tabelle, welche die Kombinationshäufigkeiten der Ausprägungen zweier oder mehrerer ›Attribute aufzeichnet

**kontinuierliches Merkmal:** ›diskretes Merkmal

**konzeptionelles Datenmodell:** ›Datenmodell, konzeptionelles

**Korrelation:** Quantifizierung des Zusammenhangs zwischen ›verhältnis- oder ›ordinalskalierten ›Attributen. Der Korrelationskoeffizient zwischen zwei Variablen ist positiv, wenn zu einem hohen Wert des einen Attributs tendenziell auch ein hoher Wert des zweiten Attributs gehört. Korrelationskoeffizienten sind so normiert, dass sie zwischen -1 (oder 0) und +1 liegen. Bei ›nominalskalierten Attributen spricht von Kontingenz (›Assoziation) statt von Korrelation.

**Künstliche Intelligenz** (KI, engl. Artificial Intelligence oder AI): Simulation kognitiver Fähigkeiten des Menschen auf dem Computer. Gegenstand der KI sind vor allem Darstellung und Manipulation von Wissen. Künstlich intelligente Systeme werden deshalb auch ›wissensbasierte Systeme genannt. Wichtige Anwendungsbereiche sind Sprachübersetzung, Bildverarbeitung, ›Expertensysteme und Intelligente Tutorielle Systeme. Zu den Methoden gehören vor allem die Methoden ›regelbasierter Systeme (einschliesslich der Regelinduktion) und ›neuronale Netze.

**LAN:** ›Local Area Network

**Legacy-System:** älteres Informationssystem, das nicht mehr dem State of the Art entspricht. Meist versteht man unter einem Legacy-System einen Grossrechner (›Mainframe) mit ›3GL-Anwendungen und einem ›Datenbanksystem alter Technologie (zum Beispiel mit einem hierarchischen ›Datenmodell oder gar einem blossen Dateisystem).

**Lernalgorithmus:** ›Modell, neuronales

**Lernen, neuronales:** Anpassung der Gewichte ›neuronaler Netze durch Analyse historischer Daten mit neuronalen ›Modellen

**Lernen, überwachtes:** Methode des maschinellen Lernens, die aus einer Menge von Eingaben und *vorgegebenen* korrekten Ausgaben (Musterlösungen) lernt. Typische Beispiele sind die ›Regelinduktion und ›Perzeptrons.

**Lernen, unüberwachtes:** Methode des maschinellen Lernens, die *nur* aus der Beobachtung der Eingaben (ohne vorgegebenen Musterlösungen) lernt. Typische Beispiele sind ›Clusteranalysen und ›Kohonen-Netze.



**Lernmenge:** Stichprobe, aus der die Parameter eines ›Modells (zum Beispiel die Gewichte eines ›neuronalen Netzes) berechnet werden

**Listenfeld:** ›Steuerelement, das eine Liste von Auswahlmöglichkeiten anbietet. Die Auswahl kann eine Einfach- oder eine Mehrfachauswahl sein.

**Local Area Network** (LAN, lokales Rechnernetz): Übertragungsnetz für Computerdaten mit räumlich eng begrenzter Ausdehnung. Die Übertragung läuft im Gegensatz zum ›Wide Area Network in der Regel über private Leitungen.

**logisch:** ›physisch

**logisches Datenmodell:** ›Datenmodell, logisches

**lokales Netz:** ›Local Area Network

**m:n-Beziehung:** ›Beziehung

**Mainframe** (Grossrechner, engl. auch host): Einzelrechner mit vielen ›Terminals und einem ›proprietären Betriebssystem. Meist identifiziert man Mainframes mit IBM-Grossrechnern unter dem Betriebssystem ›MVS und stellt sie ›Client/Server-Systemen unter UNIX gegenüber. Mainframes waren bis in die 80er Jahre die dominierenden Rechner für kommerzielle Anwendungen. Heute übertreffen einzelne ›Client/Server-Systeme, vor allem ›Parallelrechner, die Leistungsfähigkeit von Mainframes. Vorteile von Mainframes sind ihre Fähigkeit, ›Legacy-Anwendungen zu unterstützen und ihre Zuverlässigkeit. Hauptnachteile sind ihre geringe Flexibilität und Kompatibilität.

**Marktkorbanalyse** (engl. market basket analysis):

›Data Mining-Methode, die aufdeckt, welche Produkte zusammen mit anderen Produkten im “Warenkorb” der Kunden zu finden sind. Die Ergebnisse von Marktkorbanalysen können zum Beispiel zu ei-

ner veränderten Produktpräsentation (etwa einer anderen Gestellanordnung) führen. ›Assoziationsanalyse

**massiv paralleles Mehrprozessorsystem:** (MPP, engl. Massively Parallel Processing System): ›Mehrprozessorsystem, das viele CPUs über einen Hochgeschwindigkeitsbus lose koppelt und dessen Prozessoren (im Gegensatz zum einem ›symmetrischen Mehrprozessorsystem) den Internspeicher und periphere Geräte getrennt nutzen

**Master/Slave-System:** Rechnersystem, in dem ein Grossrechner die Verarbeitung und die Steuerung der Interaktion mit dem Benutzer übernimmt, die Präsentation aber durch intelligente Terminals übernommen wird. ›Host-System, ›Fileserver-System, ›zweistufiges Client/Server-System, ›dreistufiges Client/Server-System

**Median:** statistischer Lageparameter, der den Mittelpunkt von  $n$  ›ordinalskalierten Werten berechnet. Bei ungeradem  $n$  entspricht der Median dem  $(n+1)/2$ -ten Wert, bei geradem dem arithmetischen Mittel aus dem  $n/2$ -ten und dem  $(n/2 + 1)$ -ten Wert.

**mehrdimensional:** gleichzeitig auf mehrere ›Dimensionen bezogen

**mehrdimensionale Datenbank** (engl. multidimensional data base): Eine mehrdimensionale Datenbank erleichtert die ›Abfrage von ›Fakten nach ihren ›Dimensionen. Sie verbindet entweder zweidimensionale ›relationale ›Tabellen (multirelationale Datenbank) oder benutzt die ›proprietären Datenstrukturen eines mehrdimensionalen Datenbankverwaltungssystems.

**Mehrfachregression** (engl. multiple regression): ›multivariate Regressionsanalyse, die im Gegensatz zur einfachen ›Regression eine abhängige Variable durch *mehrere* unabhängige Variablen erklärt

**Mehrprozessorsystem** (MPS, engl. Multiprocessor System): Rechner, der mehrere Prozessoren eng oder lose koppelt, um mehrere ›Aufgaben gleichzeitig (›parallel) zu bearbeiten. Wichtige Arten von Mehrprozessorsystemen sind ›symmetrische und ›massiv parallele MPS.

**mehrstufiges neuronales Netz:** ›neuronales Netz, mehrstufiges

**Messniveau:** ›Skalenniveau

**Metadaten:** Daten, welche die eigentlichen Daten und Prozesse (*Objektdaten* und -prozesse) beschreiben. Metadaten werden meist in einem Metadatenbanksystem gesammelt. Beispiele sind Beschreibungen von ›Attributen, ›Datentypen, ›Integritätsregeln und Prozessen. Eine Metadatenbank heisst auch ›Data Dictionary.

**Mining:** ›Data Mining

**MIS:** engl. Management Information System. ›DSS

**Modell, neuronales:** Ein neuronales Modell besteht aus einer Architektur, einer oder mehreren Transferfunktionen und einem Lernalgorithmus. Die *Architektur* definiert Eingabe-, Ausgabe- und Verarbeitungsneuronen und ihre Beziehungen. Eine *Transferfunktion*  $f$  leitet die Eingaben an andere Neuronen oder die Umwelt weiter. Der *Lernalgorithmus* berechnet die Gewichte von  $f$ .

**modellgetriebene Methode:** Methode, die von einer Hypothese und einem einschränkenden Modell ausgeht, meist aus einer kleinen ›Grundgesamtheit Daten auswählt und rechnerisch weniger aufwendig ist als eine ›datengetriebene Methode. Beispiele: ›Optimierung, ›Expertensysteme. ›explorativ

**Modus:** statistischer Lageparameter, der den häufigsten Wert einer Variable angibt. Bei einer ›kontinuierlichen Verteilung ist der Modus der Wert der Zufallsvariablen mit der größten Wahrscheinlichkeit.

**MOLAP** (engl. Multidimensional OLAP): ›On Line Analytical Processing ›mehrdimensionaler ›Objektdaten in einem ›proprietären ›mehrdimensionalen ›Server-Datenbanksystem

**monotone** Funktion: Funktion, die entweder nur “wächst” oder nur “fällt”, das heisst: wenn  $x_1 \leq x_2$ , dann  $f(x_1) \leq f(x_2)$  bzw. wenn  $x_1 \geq x_2$ , dann  $f(x_1) \geq f(x_2)$ .

**Monte Carlo-Technik:** vor allem von der Simulationstechnik und der ›Inferenzstatistik angewendetes Verfahren zur Erzeugung von Zufallsvariablen

**MPP:** ›massiv paralleles Mehrprozessorsystem

**MPS:** ›Mehrprozessorsystem

**MS Access:** PC-Datenbanksystem von Microsoft, das den Schwerpunkt auf die endbenutzernahe Verwaltung lokaler Daten legt

**multirelationale Datenbank:** ›RDBMS mit ›mehrdimensionalem ›Datenmodell

**multivariat:** mehrere *unabhängige* Variablen betreffend

**MVS** (engl. Multiple Virtual Storage): Betriebssystem der ›Mainframes von IBM

**na:** engl. not available. ›Nullwert

**n-ärer Entscheidungsbaum:** Entscheidungsbaum, der nach einem Testattribut auf eine von  $n$  Antworten (auf einen von mehreren Baumästen) verzweigt. ›binärer Entscheidungsbaum

**naive Vorhersage:** ›Vorhersage

**NASDAQ** (National Association of Security Dealers Automated Qotation System): Bildschirmkommunikationssystem für den Han-

del mit ausserbörslich kotierten Wertpapieren (vor allem von High Technology-Werten) von etwa 5000 Unternehmungen.

**Netscape Server Application Interface (NSAPI):** ›proprietäre  
›API-Schnittstelle zwischen einem Netscape-Web Server und einem Programm, das auf einen Datenbankserver oder einen anderen Anwendungsserver zugreift

**Netzwerk-Datenbanken:** Datenbanken mit einem ›Netzwerkmodell

**Netzwerkmodell:** älteres ›Datenmodell, das stärker als das  
›Relationenmodell auf die ›physische Implementation Rücksicht nimmt. Es ist deshalb für den Benutzer komplexer und weniger flexibel, kann aber bei der Verwaltung von Massendaten mit stabiler Struktur effizienter sein.

**Neuron:** lernfähige Verarbeitungseinheit (Prozessor) eines  
›neuronalen Netzes, die Eingaben aus anderen Neuronen oder der Umwelt verarbeitet und an weitere Neuronen oder die Umwelt ausgibt

**neuronaies Lernen:** ›Lernen, neuronaies

**neuronaies Modell:** ›Modell, neuronaies

**neuronaies Netz** (engl. neural net): System von ›Neuronen, von denen jedes aus mehreren Eingaben eine einzige Ausgabe berechnet und an andere Neuronen oder die Umwelt weiterleitet. Die Architektur neuronaler Netze ist anders als jene biologischer Netze mit Millionen und Milliarden von Nervenzellen sehr einfach.

›Modell, neuronaies

**neuronaies Netz, einstufiges:** ›neuronaies Netz aus Eingabe- und Ausgabeneuronen, die nur über *eine* Gewichtungsstufe verbunden sind. ›neuronaies Netz, mehrstufiges

**neuronales Netz, mehrstufiges:** › neuronales Netz aus Eingabeneuronen, verborgenen Neuronen und Ausgabeneuronen, die durch *mehr als eine* Gewichtungsstufe verbunden sind. Das bekannteste mehrstufige Netz ist das mehrstufige Perzeptron mit Fehlerrückführung (engl. backpropagation). › neuronales Netz, einstufiges

**nichtprozedurale Abfragesprache:** › Abfragesprache

**nominalskaliert:** › Skalenniveau

**Normalisierungstheorie:** › Normalisierung

**Normalisierung:** schrittweises Aufspalten einer › Tabelle in kleinere Tabellen mit dem Ziel der Redundanzminimierung. Eine wichtige Rolle spielen die funktionalen › Abhängigkeiten zwischen Tabellenattributen.

**NSAPI:** › Netscape Server Application Programming Interface

**Nullwert:** Wert, der eine fehlende Information darstellt. Im › relationalen Modell dürfen › Primärschlüssel keine Nullwerte enthalten. Im › mehrdimensionalen Modell führen zu viele Nullwerte zu › dünn besetzten Matrizen.

**Nutzwertanalyse** (engl. multicriteria analysis): entscheidungsunterstützende Methode, die qualitative und quantitative Entscheidungselemente (Alternativen und Kriterien) auf einer gemeinsamen Punkteskala bewertet. Ein einzelnes oder mehrere Ziele werden in Unterziele (Kriterien) aufgeteilt und gewichtet. Die gewichteten Kriterien gehen in eine Entscheidungsregel ein, welche die Alternativen ordnet und eine oder mehrere auswählt.

Die Nutzwertanalyse eignet sich vor allem für ungewisse, schlecht strukturierte und innovative › Entscheidungen. Ein verbreitetes Anwendungsbeispiel ist die Beurteilung von Investitionsalternativen. › AHP ist ein mathematisch gut begründetes Beispiel eines nutz-



wertanalytischen Verfahrens. *Nachteile* der Nutzwertanalyse sind: (a) Die Synthese von Gewichten (›Prioritäten) setzt voraus, dass sich die Entscheidungselemente (Ziele, Kriterien und Alternativen) ›kardinal und unabhängig ›messen lassen. (b) Die Gewichtung lässt sich auch durch Verfahren wie ›AHP nur unzulänglich objektivieren. Dies führt vor allem bei Mehrpersonenentscheidungen zu Konflikten. *Vorteile* sind: (a) Die Nutzwertanalyse ist als ›heuristische Methode leicht nachvollziehbar und (b) oft die einzige Alternative bei qualitativen Entscheidungselementen.

**Objektdaten:** ›Metadaten

**objektive Wahrscheinlichkeit:** ›Wahrscheinlichkeit, subjektive

**objektiv probabilistische Entscheidung:** ›Entscheidung

**Objekttyp:** Klasse von Objekten mit gleichen Attributen. Im ›Relationenmodell wird ein Objekttyp durch ein ›Tabellengerüst (›Relationenschema) beschrieben.

**OCR:** engl. Optical Character Recognition

**ODBC** (engl. Open Data Base Connectivity): ›Industriestandard, der - basierend auf ›CLI - die Schnittstelle zwischen einem ›SQL-Datenbanksystem (›Backend) und einem Anwendersystem (›Frontend) oder einem anderen Datenbanksystem normiert

**OLAP:** ›Online Analytical Processing

**OLE** (Object Linking and EMBEDding): Technologie von Microsoft, welche die Kommunikation zwischen verschiedenen Anwendungen (z.B. der MS Office-Suite) erlaubt

**OLTP:** Online ›Transaction Processing. Gegensatz zu ›OLAP

**Online Analytical Processing (OLAP):** Abfragemethode, die Endbenutzern einen ›mehrdimensionalen, schnellen Zugriff und benut-

zerfreundliche ›Analysen (wie mit mehrdimensionalen Tabellenkalkulationspaketen) auf ›Data Marts ermöglicht

**Online Transaction Processing** (OLTP, Transaktionsverarbeitung): Verarbeitung in einer Transaktionsdatenbank (›Produktionsdatenbank). Die Verarbeitung detaillierter ›operativer Geschäftsvorgänge steht im Gegensatz zur analytischen Verarbeitung zusammenfassender Information zur Unterstützung managementkritischer ›Entscheidungen.

**Operations Research** (dt. Unternehmungsforschung, engl. auch Management Science): Erforschung mathematischer Methoden zur Unterstützung komplexer Entscheidungen mit dem Ziel der computergestützten Anwendung in der betrieblichen Praxis. Verbreitete Anwendungen sind ›Optimierungsverfahren und Netzwerktechniken.

**operativ** (engl. operational, dt. auch operational): Anders als ›strategische ›Entscheidungen betrifft eine operative Entscheidung konkretisierte und unmittelbar ausführbare Tätigkeiten. Operative Entscheidungen sind kurzfristig und gut strukturiert. Sie lassen sich deshalb gut automatisieren. Strategische Entscheidungen sind hingegen grundsätzlich, langfristig und kaum operationalisiert.

**Optimierung:** mathematisch begründete Lösungssuche (meist die Minimierung oder Maximierung eines Zielwerts) unter Nebenbedingungen

**Optimierung, lineare:** mathematisches Verfahren, das die Variablen einer linearen Zielfunktion ermittelt, indem es die Zielfunktion optimiert (minimiert oder maximiert), ohne dabei die Nebenbedingungen (Ungleichungen) zu verletzen. ›Was-Wenn-Analyse

**ordinalskaliert:** ›Skalenniveau

**Overfitting** (dt. Überlernen): Gefahr des neuronalen Lernens, die Gewichte eines ›neuronalen Netzes so lange anzupassen, bis sich alle Zufälligkeiten der

›Lernmenge in den Gewichten spiegeln. Overfitting erschwert die Verallgemeinerung von Stichprobenergebnissen.

**Parallelverarbeitung:** gleichzeitige Verarbeitung mehrerer  
›Aufgaben (engl. tasks) auf mehreren Prozessoren. ›MPS

**Partitionierung:** Teilung einer logisch zusammengehörenden Dateneinheit, zum Beispiel einer mehreren Dutzend ›GByte grossen Tabelle, in mehrere kleinere ›physische Untereinheiten, zum Beispiel in mehrere physische Tabellen oder Eingabe-/Ausgabeblocke. Partitionierungen verbessern die Effizienz der Verwaltung, ›Abfrage und Sicherung von Daten. ›Spaltenpartitionierung, ›Speicherpartitionierung

**PCI** (engl. Peripheral Component Interconnect): 64-bit-Bus, der durch Intel entwickelt wurde, aber prozessorunabhängig ist

**Perzeptron**, einstufiges: einstufiges ›neurales Netz, das einem Eingabevektor mit einer meist linearen oder Treppenfunktion eine Ausgabe, in der Regel eine Klasse, zuordnet

**physisch:** Oft entspricht die Benutzersicht (›logische Sicht) auf die Daten nicht der physischen Sicht. Zum Beispiel entspricht eine Tabellenzeile (›Satz) nicht der Dateneinheit, die in den ›Puffer geladen wird (Block), oder der Zugriff auf das Attribut einer Zeile (logische Adresse) entspricht nicht der physischen Adresse des Attributs auf dem Externspeicher. Physisch heisst deshalb meist “für die Maschine verständlich” und logisch “für den Benutzer verständlich”.

**physisches Datenmodell:** ›Datenmodell, physisches

**Pivoting** (engl. to pivot heisst drehen): Austausch von Zeilen und Spalten in einer Tabelle (zum Beispiel in einem Tabellenkalkulations- oder OLAP-System)

**Planerfolgsrechnung:** zukunftsgerichtete Erfolgsrechnung, die auf der Planerlösrechnung und der Plankostenrechnung aufbaut

**Plattenspiegelung** (engl. disk mirroring): Duplizierung aller Eingabe- und Ausgabeoperationen auf einen zusätzlichen Plattenspeicher

**Portfolio-Analyse:** ursprünglich ein Verfahren zur Zusammenstellung eines Wertpapierbündels (frz. portefeuille), das - nach bestimmten Kriterien bewertet - eine optimale Verzinsung erzielen soll. Seither wurde die Methode variiert und ist ein verbreitetes Instrument des strategischen Managements. Populär ist vor allem die Portfolio-Matrix der *Boston Consulting Group*: Sie vergleicht in einer Vierfelder-Matrix den relativen Marktanteil und das Marktwachstum eines Produkts während seines Lebenszyklus.

**POS:** engl. Point Of Sale

**Präprozessor:** Werkzeug, das ein Quellprogramm so umformt, dass es von einem Compiler in ablauffähigen Objektcode übersetzt werden kann

**präskriptive Entscheidungstheorie:** Entscheidungstheorie

**Primärschlüssel** (engl. primary key, Identifikationsschlüssel):  
Attributmenge, die einen Satz eindeutig identifiziert

**Priorität:** Zahl, die nach dem paarweisen Vergleich aller Kriterien bzw. Alternativen in AHP die Präferenzen des Anwenders ordnet. *Lokale* Prioritäten vergleichen die Alternativen bezüglich *eines* Kriteriums (lokale Alternativenprioritäten). *Globale* Alternativenprioritäten vergleichen hingegen die Alternativen bezüglich *aller* Kriterien des Entscheidungsmodells.

**Produktionsdatenbank** (operative Datenbank, Transaktionsdatenbank, engl. operational data base): Datenbank, die ›operative Geschäftsvorgänge aufzeichnet, indem sie Einzeldaten mit häufigen und detaillierten ›Transaktionen laufend, vollständig und redundanzarm fortschreibt.

**programmierbare Entscheidung:** ›Entscheidung

**proprietär** (engl. proprietary): durch einen einzelnen Hersteller definiert. Gegenteil von standardisiert

**Projektion:** relationale Grundoperation, die jene *Spalten* der Ausgangstabelle in eine Ergebnistabelle aufnimmt, welche die Projektionsbedingung erfüllen. ›Selektion, ›Verbund

**Prozedur, gespeicherte** (engl. stored procedure): Prozedur, die nicht ein Arbeitsplatzrechner (Client) sondern ein ›Server auf einer Datenbank ausführt. Eine gespeicherte Prozedur besteht aus ›SQL-Anweisungen und prozeduralen Erweiterungen, insbesondere Kontrollkonstrukten. Gespeicherte Prozeduren werden zur Entwicklungszeit oder erst vor ihrem ersten Aufruf geprüft, kompiliert und gebunden. Sie sind meist einige Male schneller als entsprechende dynamisch entstandene SQL-Anweisungen. ›Trigger, ›Agent

**prozedurale Abfragesprache:** ›Abfragesprache

**Pseudocode:** Entwurfssprache, die den Programmcode in einer programmiersprachen-unabhängigen Notation leserlich zusammenfasst

**Puffer** (engl. buffer, cache): Teil des Internspeichers, der Daten vor Schreib- oder nach Lesezugriffen auf dem Externspeicher möglichst effizient intern zwischenspeichert. Um Dateioperationen zu beschleunigen, werden zum Beispiel mehrere Sätze eines Extern-

speichers als Block in den Puffer gelesen, wo sie zur Weiterverarbeitung bereit stehen.

**QBE** (engl. Query By Example): benutzerfreundliche ›deklarative  
›Abfragesprache für ›relationale Datenbanken, in welcher der Benutzer die gewünschten Ergebnisspalten in Tabellengerüsten markiert und die Kriterien für die Auswahl der Zeilen und den Verbund von Tabellen einträgt. ›SQL

**RAID** (engl. Redundant Array of Inexpensive Disks): System aus vielen kleinen Plattenspeichern, das als ganzes an den Systembus angeschlossen ist. Wegen ihrer Zahl und ihrer geringen Grösse sind die einzelnen Festplatten sicherer und bei einem Ausfall ist nur ein Teil des gesamten Plattenspeichers betroffen. Ausserdem lassen sich ausgefallene Daten wegen der redundanten Speicherung einfach wieder herstellen.

**RBF-Netz** (engl. Radial Basis Function network): vorwärtsgerichtetes ›überwacht lernendes ›neuronales Netz mit einer einzigen verborgenen Schicht und einer nichtlinearen ›Transferfunktion

**RDBMS** (engl. Relational Database Management System, relationales Datenbanksystem): ›Datenbanksystem, das dem  
›Relationenmodell gehorcht

**Recovery**: Wiederherstellung des korrekten Zustands einer Datenbank nach einem Fehler, zum Beispiel dem Zusammenbruch des Betriebssystems. Der Wiederanlauf (Rollback) nach einer nicht abgeschlossenen ›Transaktion lässt sich als Sonderfall einer Recovery-Massnahme interpretieren.

**Regel**: Paar aus Bedingung und Folgerung, die beide einfach oder zusammengesetzt sein können. Im einfachsten Fall besteht eine Bedingung und eine Folgerung nur aus Attribut/Wert-Paaren. Attribut/Wert-Paare sind wahr oder falsch und lassen sich durch die Operatoren “und” bzw. “oder” verknüpfen. In einer Folgerung kann



statt eines Attribut/Wert-Paars auch eine Aktion stehen. Eine Aktion ist im Gegensatz zu einem ›Attributwert nicht wahr oder falsch, sondern führt eine Handlungsanweisung aus.

**Regelbasiertes System:** ›wissensbasiertes System, das Wissen als Regeln und Fakten darstellt

**Regelinduktion:** ›Induktion

**Regelmodul:** Teil eines Expertensystems, dessen ›Attribute und ›Regeln inhaltlich und formal zusammengehören

**Regressionsanalyse** (engl. regression analysis): statistisches Verfahren, das die Koeffizienten einer mathematischen Gleichung so schätzt, dass der Fehler zwischen den aktuellen und vorhergesagten Werten minimiert wird. ›Mehrfachregression

**Relation:** zweidimensionale ›Tabelle aus eindeutig identifizierbaren Zeilen, die ihrerseits aus Zellen (›Attributen) bestehen, die je höchstens einen Eintrag enthalten. ›Wiederholungsgruppen

**Relationenmodell:** ›Datenmodell, in dem alle Daten und ihre ›Beziehungen als ›Relationen dargestellt werden

**Reliabilität** (engl. reliability, dt. auch Zuverlässigkeit): Die Reliabilität einer Methode gibt an, wie gut ihre Ergebnisse bei wiederholter Anwendung der Methode auf Daten der gleichen ›Grundgesamtheit übereinstimmen. Sie wird oft als ›Korrelation zwischen zwei Reihen von Ergebniswerten berechnet. ›Validität

**Replikat** (engl. replica): Datenobjekt, das auf mehrere Rechner kopiert und regelmässig synchronisiert wird

**Replikation:** Erstellung und Synchronisierung von ›Replikaten

**Reverse Engineering:** Analyse eines Informationssystems mit dem Ziel, Objekte und Beziehungen zu erkennen und sie auf ein

anderes (meist höher entwickeltes) System abzubilden. Ein Beispiel ist die Analyse von Quelldaten des Ladeprozesses, um ihre Metadaten in das ›Data Dictionary eines ›Data Warehouse einzufügen.

**RISC** (engl. Reduced Instruction Set Computing): Im Gegensatz zu ›CISC-Rechnern (engl. Complex Instruction Set Computing) verwenden RISC-Rechner möglichst einfache Anweisungen, die nur je einen CPU-Zyklus erfordern.

**robust**: Ein Verfahren ist robust, wenn es ohne oder mit wenig einschränkenden Annahmen auskommt und wenig empfindlich auf deren Verletzung reagiert. Gegenteil von sensitiv.

›Sensitivitätsanalyse

**ROE** (engl. Return On Equity, dt. Eigenkapitalrendite): Verhältnis einer Erfolgsgröße, zum Beispiel des Nettogewinns zum eingesetztem Eigenkapital einer Rechnungsperiode. Beide Größen werden bilanziell gemessen. ›ROI

**ROI** (engl. Return On Investment, dt. Kapitalrendite): Verhältnis des investierten Kapitals oder des Umsatzes zum Gewinn. Möglicher Ausgangspunkt für Massnahmen der Unternehmenspolitik und Unternehmensplanung. ›ROE

**ROLAP** (engl. Relational OLAP): ›OLAP auf ›mehrdimensionalen ›Metadaten in einem ›relationalen ›Server-Datenbanksystem

**Satz** (engl. record): Gruppe von Datenelementen (›Feldern), die als eine Einheit behandelt wird. Im ›Relationenmodell entspricht einer Tabellenzeile ein Satz.

**Schaltfläche**: ›Befehlsschaltfläche

**Schema**: Gerüst einer (1) ›Tabelle (*Relationenschema*, Tabellenskelett) oder einer (2) Datenbank (*Datenbankschema*, Datenbankdefinition)

**schlecht strukturierte Entscheidung:** ›Entscheidung

**Schlüssel:** Menge von ›Attributen, die Sätze ›identifiziert, indiziert, sortiert oder verbindet. ›Primärschlüssel, ›Sekundärschlüssel

**schrittweise Regression** (engl. stepwise regression): Verfahren, das variablenweise (schrittweise) aufgrund eines Aufnahmekriteriums (zum Beispiel des ›F-Tests) entscheidet, wieviele und welche unabhängige Variablen in welcher Reihenfolge in eine ›Regressionsanalyse eingehen

**Schneeflockenschema** (engl. snow flake schema): um zusätzliche Faktentabellen oder Unterdimensionstabellen erweitertes ›Sternschema

**Schwellenkonstante:** Wert, den die Summe der gewichteten Eingaben eines ›Neurons kompensieren muss, damit das Neuron gerade noch feuert

**Script** (auch Skript, Makro oder Batch File): Befehlsliste, die von einer einfachen Skriptsprache interpretiert wird. Beispiele von Skriptsprachen für das WWW sind ›VBScript oder ›JavaScript.

**Sekundärschlüssel** (engl. alternate key): Zugriffsschlüssel, der im Gegensatz zum ›Primärschlüssel nicht alle Sätze eindeutig identifizieren muss

**Selektion:** ›relationale Grundoperation, die jene *Zeilen* der Ausgangstabelle in eine ›Ergebnistabelle aufnimmt, welche die Selektionsbedingung erfüllen. ›Filtern, ›Projektion, ›Verbund

**semantisch:** auf die Bedeutung - und nicht die Form - bezogen. Der Satz "Der Rhein fließt bunt" ist zwar syntaktisch (grammatikalisch) korrekt, ›semantisch aber bedeutungslos.

**Sensitivitätsanalyse** (engl. sensitivity analysis): Prüfverfahren, das die Empfindlichkeit (Sensitivität, ›Robustheit) eines Modells auf

Parameter- und Datenänderungen untersucht. Beispiel: “Wie lassen sich die Koeffizienten eines linearen ›Optimierungsmodells ändern, ohne dass die Lösung ihre Optimalität verliert?”

**Server:** ›Client/Server-Architektur

**Server-Datenbank:** Datenbank auf einem Dienstleistungsrechner (›Server), die im Gegensatz zu einer ›lokalen Datenbank mehreren Benutzern zur Verfügung steht

**Sicht:** ›Benutzersicht

**Signifikanzniveau** (dt. Irrtumswahrscheinlichkeit): Wahrscheinlichkeit der fälschlichen Ablehnung einer Hypothese

**Signifikanztest:** statistisches Verfahren, das prüft, ob eine Hypothese fälschlicherweise abgelehnt wird

**Skalenniveau** (Messniveau oder -qualität): Statistiker unterscheiden vier Messqualitäten:

(1) *Nominalskalen* klassifizieren (Bsp. Standort A, Standort B, ...). Die Analyse nominalskalierter Daten beschränkt sich auf Häufigkeitsanalysen (zum Beispiel die Berechnung des ›Modus).

(2) *Ordinal- oder Rangskalen* ordnen (Bsp. Standort A ist besser geeignet als Standort B), messen jedoch den Abstand zwischen den verglichenen ›Attributwerten nicht. Neben Häufigkeiten lassen sich auch Lageparameter berechnen (zum Beispiel der Median).

(3) *Intervallskalen* messen den Abstand zweier Zahlen, ohne dass ein eindeutiger Nullpunkt definiert ist (Bsp. Abstände in einem genormten Schulleistungstest). Sie erlauben die Berechnung des arithmetischen Mittelwerts.

(4) *Verhältnisskalen* (Ratioskalen) kennen zusätzlich einen eindeutigen Nullpunkt (Bsp. Unternehmungserfolg). Werte, die auf einer

Intervall- oder Verhältnisskala gemessen werden können, heissen zusammenfassend *kardinal*.

**skalierbar** (engl. scalable): Ein Software- oder Hardwaresystem ist skalierbar, wenn seine Kapazität inkrementell und ohne überdurchschnittlichen Aufwand vergrössert werden kann. Skalierbarkeit ist eine wichtige Anforderung an Data ›Warehouses, ›Mehrprozessorsysteme und ›Client/Serversysteme.

**Skript:** ›Script

**Slice and Dice** (auch Slicing and Dicing): Auswahl und Weglassen von ›Dimensionen in einem ›OLAP-Werkzeug

**SMP:** ›Symmetrisches Mehrprozessorsystem

**Spaltenpartitionierung:** spalten- statt zeilenweise Besetzung der ›physischen Eingabe-/Ausgabeseiten (engl. I/O pages) eines Datenbanksystems. Weil ›Data Warehouses Tabellen meist spaltenweise statt zeilenweise analysieren, speichern sie ihre Daten oft so, dass die Werte der gleichen *Attributspalte* statt der gleichen Satzzeile im gleichen I/O-Block stehen. ›Partitionierung

**Speicherpartitionierung:** Objektdaten werden so auf dem Externspeicher angeordnet, dass zusammengehörende Daten möglichst rasch in den Internspeicher gelangen

**SQL** (engl. Structured Query Language): nichtprozedurale, standardisierte ›Abfragesprache, die für den Endbenutzer entworfen wurde, heute aber vor allem für den programmierten Zugriff auf ›Server-Datenbanken, ›verteilte Datenbanken und zentrale ›Grossrechner-Datenbanken verwendet wird. Die meisten SQL-Produkte sind Erweiterungen des laufend angepassten ›ANSI-Standards. eingebettetes ›SQL, ›QBE, ›Abfragesprache

**SQL, eingebettetes** (engl. embedded SQL): ›SQL-Anweisungen, die in eine Gastsprache wie C oder COBOL eingebettet sind. Ein

›Präprozessor übersetzt die eingebetteten SQL-Befehle in Aufrufe aus einer Bibliothek von Unterprogrammen, die der Programmierer nicht zu kennen braucht. Nachteilig ist, dass der Quellcode nach jedem Wechsel der Plattform neu übersetzt werden muss. ›CLI

**Statistik:** ›Inferenzstatistik

**statistische Vorhersage:** ›Vorhersage

**Sternschema** (engl. star schema): ›logisches Datenbank›schema, das die ›Dimensionstabellen eines ›relationalen ›Data Warehouse abfragefreundlich um eine ›Faktentabelle ordnet

**Steuerelement** (engl. control): graphisches Objekt, zum Beispiel ein Textfeld oder eine ›Befehlsschaltfläche, in einem ›Formular oder ›Bericht. Der Programmierer kann die Eigenschaften von Steuerelementen ändern und so ihr Aussehen oder ihre Reaktionen auf Ereignisse ändern.

**Stichprobe** (engl. sample, dt. auch Teilerhebung): Erhebung, in die nur ein Teil der ›Grundgesamtheit eingeht. Eine Vollerhebung kommt meist aus Kosten- und Zeitgründen nicht in Frage. Je nach Auswahlverfahren ist die Repräsentativität von Stichprobenergebnissen mehr oder weniger gewährleistet. Die Modelle der Inferenzstatistik setzen Zufallsstichproben voraus. Der Einsatz von ›Data Mining-Software erlaubt oft die Nutzung des gesamten Datenvorkommens.

**strategisch:** ›operativ

**symbolisch:** nicht numerisch, auf Zeichenketten bezogen

**symmetrisches Mehrprozessorsystem** (SMP, engl. Symmetric Multiprocessor System): ›Mehrprozessorsystem, das wenige CPUs über den Systembus eng koppelt und dessen Prozessoren im Gegensatz zum ›massiv parallelen Mehrprozessorsystem den Internspeicher und periphere Geräte gemeinsam nutzen



**syntaktisch:** formbezogen. Der Satz “Der Rhein fließt bunt” ist zum Beispiel syntaktisch (grammatikalisch) korrekt, ›semantisch aber bedeutungslos.

›**Synthese:** in ›AHP Berechnung der Prioritäten von Kriterien und Alternativen durch Zusammenfassung von Paarvergleichen

**subjektive Wahrscheinlichkeit:** ›Wahrscheinlichkeit, subjektive

**subjektiv probabilistische Entscheidung:** ›Entscheidung

**t-Tests:** statistische Verfahren, die unter anderem testen, ob zwei Mittelwerte gleich gross sind oder ob ein Regressionskoeffizient 0 ist

**Tabelle:** grundlegende Datenstruktur ›relationaler ›Datenbanksysteme. Eine Tabelle besteht aus einer Menge eindeutiger Zeilen mit einer festen Zahl von ›Attributen (Spalten) ohne ›Wiederholungsgruppen.

**Task:** ›Aufgabe

**Tbyte:** Abkürzung für ›Terabyte

**TCP/IP** (Transmission Control Protocol/Internet Protocol):  
›Industriestandard, der die Protokolle der Transport- und Vermittlungsschicht des ›ISO-Referenzmodells definiert. TCP regelt das Verpacken und Entpacken von Nachrichten, während IP die eigentliche Übermittlung der Pakete bestimmt.

**Template:** Ersetzungstext (Textskelett), der Platzhalter enthält, die zur Laufzeit durch konkrete Werte ersetzt werden können

**Terabyte:** 1024 Gigabytes ( $10^{12}$  Bytes)

**Terminal:** Endstation, an der Daten ein- oder ausgegeben werden können. Ein “intelligentes” Terminal kann im Gegensatz zum “dum-

men” Terminal Eingabe-Ausgabe-Funktionen wie Cursorsteuerung, Pufferung und Datenübertragung selbst durchführen.

**Testmenge:** Stichprobe, die beurteilen hilft, ob sich ein ›Modell verallgemeinern lässt. ›Lernmenge

**Transaktion:** Folge von Datenbankbefehlen, die logisch zusammengehören und eine Datenbank von einem konsistenten Zustand in einen anderen überführen. Beispiele von Transaktionen sind Buchungsvorgänge an einem Bankschalter. ›Transaktionsdatenbank

**Transaktionsdatenbank:** ›Produktionsdatenbank

**Transferfunktion:** ›Modell, neuronales

**Trefferquote** (engl. accuracy): Quotient zwischen der Zahl der richtigen ›Vorhersagen und der Gesamtzahl der Vorhersagen

**Trigger:** Datenbankprozedur, die im Gegensatz zu einer gespeicherten ›Prozedur, die der Programmierer explizit aufruft, automatisch vor oder nach einer “insert”-, “update”- oder “delete”-Anweisung auf einer ›Basistabelle startet. ›Agent

**überwachtes Lernen:** ›Lernen

**Unternehmensdatenmodell** (engl. corporate data model, auch enterprise data model): ›Datenmodell, das möglichst viele der lokalen Datenmodelle einer Unternehmung (zum Beispiel von Abteilungen oder Arbeitsgruppen) integriert

**Unternehmensforschung:** ›Operations Research

**unüberwachtes Lernen:** ›Lernen

**URL** (Uniform Resource Locator): eindeutige Adresse eines Internetrechners bzw. einer Webseite. Im Beispiel [http:// www.wwz.unibas.ch/wi/home.html](http://www.wwz.unibas.ch/wi/home.html) folgt zum Beispiel auf das Kommunikationsprotokoll (›http) die Domäne (zum Beispiel WWZ der Universität

Basel, CH). Der restliche Teil der Adresse nennt das Verzeichnis auf dem Web Server (wi/ für Wirtschaftsinformatik) und den Namen der anzuspringenden Webseite (home.html).

**Validität:** Verallgemeinerungsfähigkeit. Die Validität *messen* heisst nach der Anwendung eines Verfahrens auf eine ›Stichprobe die berechneten mit den tatsächlichen Ergebnissen vergleichen und diesen Vergleich in einem Mass der Vorhersagegüte operationalisieren. ›Reliabilität

**Varianz:** Streuungsmass zur Quantifizierung der Verteilung einer Zufallsvariablen. Ein grosse Varianz bedeutet, dass die untersuchten Variablenwerte “weit voneinander entfernt” sind.

**Varianzanalyse** (engl. Analysis Of Variance, ANOVA): Test der Nullhypothese, dass die Mittelwerte mehrerer ›Stichproben in der ›Grundgesamtheit gleich gross sind. Dazu wird die geschätzte ›Varianz der Stichproben-Mittelwerte mit den geschätzten Varianzen innerhalb der Stichproben verglichen. ›t-Tests

**Verdichtung** (Datenreduktion): Rückführung von Daten auf wenige aussagekräftige Grössen. Beispiele von Verdichtungsmethoden sind Gruppenbildung, ›Kennzahlen (zum Beispiel Mittelwerte, Quoten, Indexzahlen) und ›multivariate statistische Methoden wie ›Regressions-, ›Diskriminanz-, ›Cluster- oder ›Faktorenanalyse.

**VBA:** ›Visual Basic for Applications

**VBScript:** von ›Visual Basic abgeleitete Skriptsprache von Microsoft, die in ›HTML eingebettet werden kann und die Interaktion mit dem Browser unterstützt. ›JavaScript

**Verbund** (engl. join): ›relationale Grundoperation, die genau zwei ›Tabellen verbindet. Eine gebräuchliche Art des Verbunds vergleicht jeden ›Satz der ersten Tabelle mit allen Sätzen der anderen Tabelle. Wenn die Verbundbedingung erfüllt ist, werden die beiden

vergleichenen Sätze zusammengehängt und als ein Satz in die  
›Ergebnistabelle geschrieben. ›Selektion, ›Verbund

**Verbundnetz:** ›Wide Area Network

**Vergleichsmatrix:** quadratische Matrix, die Entscheidungselemente in ›AHP (Kriterien oder Alternativen) vergleicht. Weil ein Entscheidungselement  $i$  mit sich selbst verglichen eine Vergleichszahl (›Priorität) von 1 ergibt, besteht die Hauptdiagonale der Vergleichsmatrix aus lauter *Einsen*. Einander entsprechende Vergleiche unter und über der Hauptdiagonale sind *reziprok*. Eine Vergleichsmatrix heisst *konsistent*, wenn sie nur konsistente Paarvergleiche darstellt.

**verhältnisskaliert:** ›Skalenniveau

**verteilte Datenbank** (engl. distributed database): Datenbank, die sich über mehrere geographisch entfernte Rechner erstreckt

**Viertgenerationssprache** (engl. 4th Generation Language, 4GL): Sprache, die auch ungeübten Programmierern erlaubt, in kurzer Zeit eine Anwendung zu entwickeln. Die meisten 4GL unterstützen eine intuitive ›GUI-Definition und erlauben einen einfachen Zugriff auf Datenbanken. Anders als ›Drittgenerationssprachen wie C und COBOL werden 4GL-Programme in der Regel interpretiert oder in ›3GL-Programme übersetzt (›Applikationsgeneratoren). Sie sind deshalb weniger laufzeit- und speichereffizient als 3GL-Sprachen. ›VBA

**View:** ›Benutzersicht

**Visual Basic for Applications (VBA):** interpretierter BASIC-Dialekt von Microsoft. VBA eignet sich zur applikationsübergreifenden Programmierung, zum Beispiel zur Verknüpfung von MS Office-Anwendungen wie ›MS Access und MS Excel.

**Visualisierung:** bildliche Darstellung komplexer Daten zur Entdeckung von ›Data Mining-Hypothesen oder Veranschaulichung von Ergebnissen

**vorberechnete Daten:** ›Data Warehouse-Inhalte, die im Gegensatz zu importierten Daten nicht fremdbezogen sind, sondern vom Datenbankverwalter definiert und nach jedem Ladevorgang berechnet (›aggregiert) und in das ›DW eingefügt werden. Vorberechnete Daten verringern die Abfragezeit, erfordern aber anders als ›ad hoc berechnete Daten Speicherplatz, vergrössern die Ladezeit und erschweren die Datenverwaltung. Dafür erübrigt sich eine Sicherung der vorberechneten Daten.

**Vordefinition:** Für Abfragen, die vor ihrer Durchführung bereits bekannt sind, können Indizes gezielt erstellt und Teiloperationen gezielt sequenziert werden

**Vorhersage:** Berechnung ›kontinuierlicher Zukunftswerte aus unabhängigen Variablen. *Naive* Vorhersagen suchen nicht nach Faktoren, welche die vorherzusagende Grösse beeinflussen könnten und berücksichtigen nur leicht verfügbare Vorinformation (a priori-Information). *Statistische* Vorhersagen (etwa Regressionen oder Entscheidungsbäume) schätzen hingegen die a posteriori-Wahrscheinlichkeit einer Vorhersagegrösse unter der Annahme, dass diese von bestimmten Faktoren beeinflusst wird. ›Klassifikation

**Warum-Erklärung:** ›Erklärung

**VSAM** (engl. Virtual Sequent Access Method): ›index-sequentielles Methode des Dateizugriffs von IBM

**Wahrscheinlichkeit, bedingte** (engl. conditional probability): Wahrscheinlichkeit eines zufälligen Ereignisses unter der Bedingung, dass ein anderes zufälliges Ereignis bereits eingetreten ist

**Wahrscheinlichkeit, objektive:** ›Wahrscheinlichkeit, subjektive

**Wahrscheinlichkeit, subjektive:** Quantifizierung der persönlichen Überzeugung eines Experten. Ein Modell muss sich dann mit *subjektiven* Wahrscheinlichkeiten begnügen, wenn keine relativen Häufigkeiten (*objektive* Wahrscheinlichkeitswerte) zur Verfügung stehen.

**WAN:** ›Wide Area Network

**Warum-nicht-Erklärung:** ›Erklärung

**Was-Wenn-Analyse** (engl. what if analysis): Sammelbegriff für Methoden, die untersuchen, wie sich eine Änderung einer oder mehrerer unabhängiger Variablen (Ausgangswerte) auf eine abhängige Variable (Zielwert) auswirkt. Beispiele: Mehrfachoperation, Szenario und Zielwertsuche in Tabellenkalkulationspaketen, mathematische ›Optimierungsverfahren

**Was-Wenn-Erklärung:** ›Erklärung

**Wertebereich** (engl. domain): Menge aller Werte, die ein ›Attribut annehmen kann

**Wide Area Network** (WAN): privates oder öffentliches Übertragungsnetz für Computerdaten mit grosser geographischer Ausdehnung (bis in andere Kontinente). Im Gegensatz zu ›LANs transportieren WANs ihre Daten in der Regel über öffentliche Leitungen.

**Wie-Erklärung:** ›Erklärung

**Wiederholungsgruppen:** Gruppe von Werten des gleichen Attributs im gleichen ›Satz

**Wissensbasiertes System** (engl. knowledge based system): (1) System, das Methoden der ›Künstlichen Intelligenz einsetzt. (2) Oft auch Synonym zu ›Expertensystem

**wohlstrukturierte Entscheidung:** ›Entscheidung



**World Wide Web (WWW):** benutzerfreundliche Oberfläche, auf der Endbenutzer hypermedial durch die wichtigsten ›Internetdienste navigieren

**Wozu-Erklärung:** ›Erklärung

**Würfel** (engl. cube, hypercube, multidimensional array):  
›mehrdimensionale Datenstruktur, welche die ›Analyse eines  
›Indikators (Fakts) nach mehreren ›Dimensionen ermöglicht. Ein dreidimensionaler Würfel besteht zum Beispiel aus dem Indikator “Unternehmungsertrag” und den Dimensionen “Produkt”, “Zeit” und “Gebiet”. Die Indikatorwerte kann man sich an Schnittpunkten der drei Dimensionen vorstellen. Zusätzliche Dimensionen lassen sich durch Farbe, Form, etc. von Objekten im dreidimensionalen Koordinatensystem darstellen. ›OLAP-Werkzeuge erlauben durch ›Slicing and Dicing und ›Ebenen einen benutzerfreundlichen Wechsel zwischen beliebigen Dimensionen.

**WWW:** ›World Wide Web

**WYSIWYG** (engl. What You See Is What You Get): Benutzeroberfläche, die einen Inhalt wirklichkeitsgetreu formatiert

**XML** (engl. eXtensible Markup Language): erweiterungsfähige deklarative Sprache, die nicht nur die Form sondern auch den Inhalt von Webseiten beschreiben kann. ›HTML

**Zeitreihenanalyse** (engl. time series analysis): Auswertung einer Folge numerischer Beobachtungen in Abhängigkeit von der Zeit. Ziel ist meist die Zerlegung des Verlaufs der Zeitreihe in ihre Komponenten. Zu den Komponenten gehören der *Trend* (längerfristige Bewegung, die den Grundverlauf bestimmt) sowie *zyklische* (zum Beispiel saisonale) und *zufällige* Schwankungen. Verbreitete Verfahren sind die Methode der gleitenden Durchschnitte, die Methode der kleinsten Quadrate (›Regressionsanalyse), die Spektralanalyse und ›neuronale Netze.

**zentrales Data Warehouse:** ›Data Warehouse, zentrales

**Zufallsstichprobe:** ›Stichprobe

**Zugriff** (engl. access): Positionieren des Lesegeräts mit anschließendem Lesen oder Schreiben von Daten auf einem meist externen Speicher

**Zugriffszeit** (engl. access time): ›Antwortzeit

**zweistufiges Client/Server-System:** ›Client/Server-System, zweistufiges