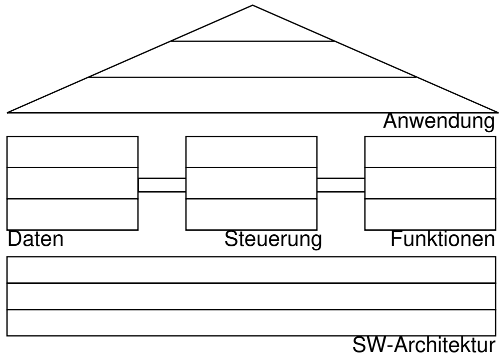
1. Einführung und Grundbegriffe

- **GBIS-Rahmen: Einordnung**



- **Miniwelt**[1]
  - modellhafte Abbildung eines Realitätsausschnitts
  - Transaktionskonzept – ACID-Eigenschaften

- **Information – was ist das?**

- **Aufgaben eines Informationssystems?**
  - Komponenten eines rechnergestützten Informationssystems
  - Aufgaben und Zielsetzungen eines betrieblichen Informationssystems
  - wichtige Anwendungsklassen

- **Daten in Informationssystemen**
  - strukturierte und unstrukturierte Daten
  - semi-strukturierte Daten (HTML, XML)
  - Multimedia-Daten (VITA-Daten)

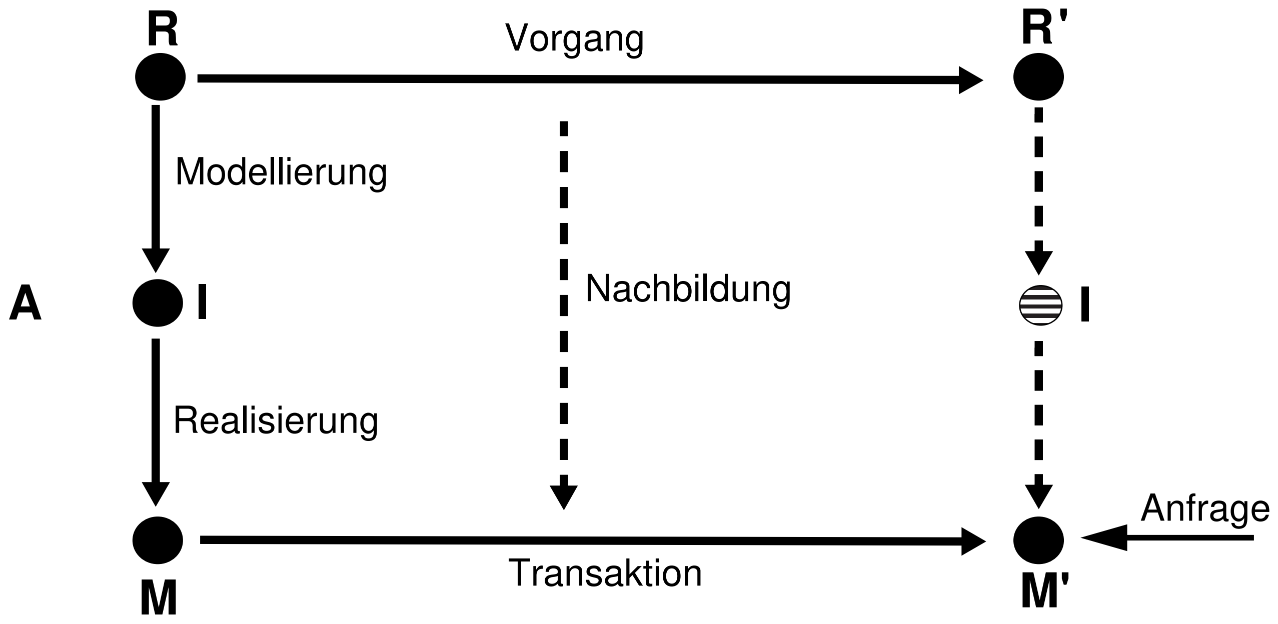
- **Beispiele für Informationssysteme**
  - Informationssystem einer Universität, eines Produktionsbetriebs, einer Bank . . .
  - Straßeninformationssystem

---

1. Ein Datenbanksystem verwaltet Daten einer realen oder gedanklichen Anwendungswelt. Diese Daten gehen aus Informationen hervor, die stets aus den Sachverhalten und Vorgängen dieser Anwendungswelt durch gedankliche Abstraktionen (Abbilder, Modelle) gewonnen werden. Sie beziehen sich nur auf solche Aspekte des betrachteten Weltausschnitts, die für den Zweck der Anwendung relevant sind. Ein solcher Weltausschnitt wird auch als *Miniwelt* (Diskurswelt) bezeichnet.

Miniwelt – modellhafte Abbildung

- **Grobe Zusammenhänge**



R: Realitätsausschnitt (Miniwelt)

I: Informationsmodell
(zur Analyse und Dokumentation der Miniwelt)

M: DB-Modell der Miniwelt
(beschrieben durch Objekt- und Beziehungsmengen
sowie Integritätsbedingungen usw.)

A: Abbildung aller relevanten Objekte und Beziehungen
➥ Abstraktionsvorgang

- **Transaktion (TA):**
  - bildet Vorgang in **R** im DBS nach und
  - garantiert ununterbrechbaren Übergang von **M** nach **M'**
    ➥ implementiert durch Folge von DB-Operationen
  - DB-Anfragen beziehen sich auf **M** bzw. **M'**

- **Integritätsbedingungen:**
  - Zusicherungen über **A**, **I** und **M**: A_1: R→I, A_2: I→M
    ➥ Ziel: möglichst gute Übereinstimmung von **R** und **M**
  - **Idealfall:** Die DB ist zu jeder Zeit ein Abbild (Modell) der gegebenen Miniwelt

Miniwelt – modellhafte Abbildung (2)

- **Transaktionskonzept**
  - führt ein neues Verarbeitungsparadigma ein
  - ist Voraussetzung für die Abwicklung betrieblicher Anwendungen (*mission-critical applications*)
  - erlaubt „Vertragsrecht“ in rechnergestützten IS zu implementieren

- **Welche Eigenschaften von Transaktionen sind zu garantieren? (ACID-Paradigma)**
  - **Atomicity (Atomarität)**
    - TA ist kleinste, nicht mehr weiter zerlegbare Einheit
    - Entweder werden alle Änderungen der TA festgeschrieben oder gar keine („alles-oder-nichts“-Prinzip)
  - **Consistency**
    - TA hinterläßt einen konsistenten DB-Zustand, sonst wird sie komplett (siehe Atomarität) zurückgesetzt
    - Zwischenzustände während der TA-Bearbeitung dürfen inkonsistent sein
    - Endzustand muß die Integritätsbedingungen des DB-Modells erfüllen
  - **Isolation**
    - Nebenläufig (parallel, gleichzeitig) ausgeführte TA dürfen sich nicht gegenseitig beeinflussen
    - Alle anderen parallel ausgeführten TA bzw. deren Effekte dürfen nicht sichtbar sein
  - **Durability (Dauerhaftigkeit)**
    - Wirkung einer erfolgreich abgeschlossenen TA bleibt dauerhaft in der DB erhalten
    - TA-Verwaltung muß sicherstellen, daß dies auch nach einem Systemfehler (HW- oder System-SW) gewährleistet ist
    - Wirkungen einer erfolgreich abgeschlossenen TA kann nur durch eine sog. kompensierende TA aufgehoben werden

Information – Was ist das?

- **Beobachtung**

  Die Praxis der Information und Kommunikation entwickelte sich rasant, ohne daß der Informationsbegriff von einer Theorie hinreichend geklärt wurde. Es gibt keine Theorie oder gar Philosophie der Information.

  ➥ Eine resignierende Schlußfolgerung[1]:
  Die Definition des Begriffes „Information“ ist nicht möglich.
  Jeder Versuch dazu setzt ähnliche Begriffe voraus, beispielsweise „Wissen“ oder „Kommunikation“. Diese Definition wäre damit zyklisch.

- **Erklärungsversuche (philosophisch, technisch, pragmatisch):**

  1. „Information ist neben Materie und Energie etwas Drittes“

  2. Information für Menschen über seine Umwelt:
     Information setzt den Menschen über seine Außenwelt in Kenntnis, ist also der „Stoff“, der Erkenntnis ermöglicht

  3. Informationstheorie nach Shannon:
     Statistischer Informationsbegriff (Entropie einer Nachrichtenquelle):

     $$H(p_1,\ldots,p_n) = -\sum_{i=1}^{n} p_i \cdot \log\ p_i$$

  4. Information und Nachricht
     Die übermittelte Nachricht ist dann von Bedeutung, wenn wir eine Abbildung kennen, die sie mittels einer Interpretationsvorschrift α auf eine Information abbildet:

     $$\mathbf{N} \xrightarrow{\alpha} \mathbf{I}$$

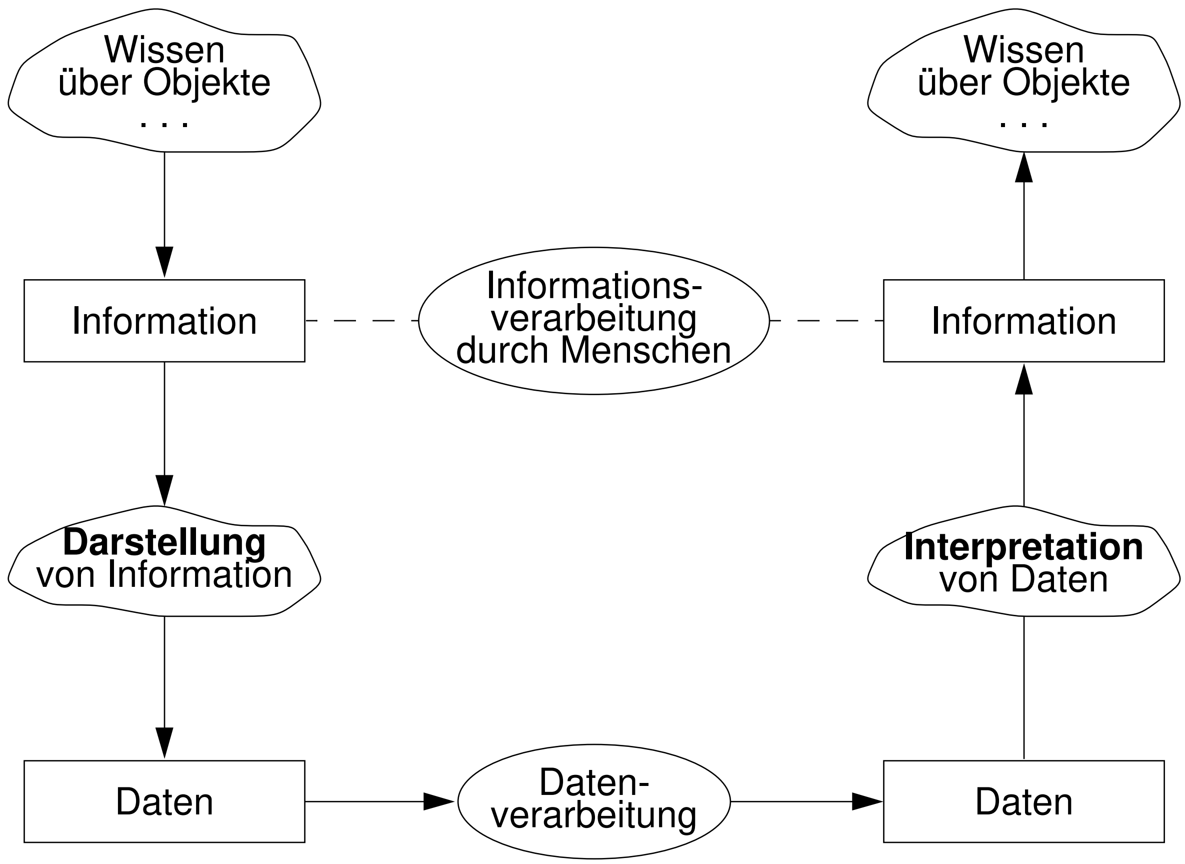
---

1. Bauer, F.L., Goos, G.: Informatik – Eine einführende Übersicht, 3. Auflage, Springer-Verlag, 1982

Information – Was ist das? (2)

- **Erklärungsversuche:**

5. Informationsbegriff nach DIN: **Erklärungsmodell**
   (auch für Nachrichtenaustausch zwischen Sender und Empfänger)



Information: subjektive Welt der bewerteten Daten

Daten: objektive Welt der nicht-interpretierten Daten

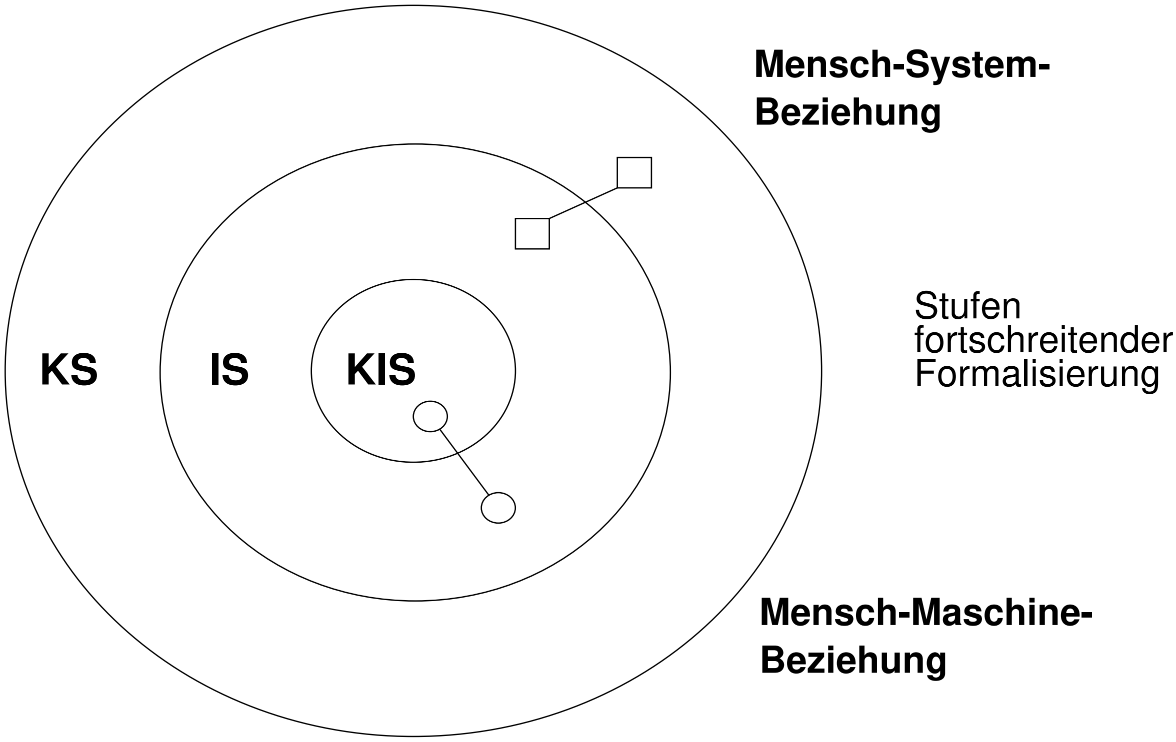
6. Pragmatische Festlegung in der BWL[1]
   Information: **Angaben über Sachverhalte und Vorgänge** (Hansen)
   - „Dabei kann man Information im Sinne von als zweckgerichtetes Wissen zur Vorbereitung und Durchführung von Handlungen verstehen".
   - „Eine andere Sichtweise ergibt sich aus der Betrachtung der Information als Produktionsfaktor".

1. Wirtschaftsinformatik-Lexikon, Gabler-Verlag, 1997

Was ist ein Informationssystem?

- **Charakterisierung eines IS nach Senko:**

“The purpose of an information system is to provide a *relatively exact, efficient, unambigous* model of the significant resources of a real world enterprise.”



Formalisiertes Kommunikationssystem = Informationssystem

- **(Vage) Definitionen:**

Ein Informationssystem[1] (IS) besteht aus Menschen und Maschinen, die Informationen erzeugen und/oder benutzen und die durch Kommunikationsbeziehungen miteinander verbunden sind.

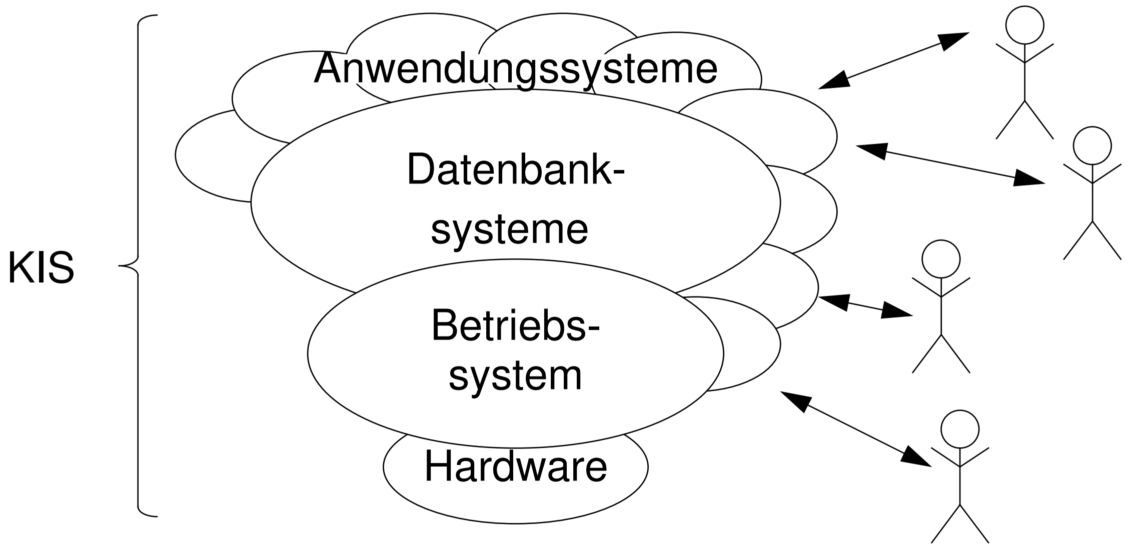
Ein betriebliches IS dient zur Abbildung der Leistungsprozesse und Austauschbeziehungen im Betrieb und zwischen dem Betrieb und seiner Umwelt.

Ein rechnergestütztes IS ist ein System, bei dem die Erfassung, Speicherung und/oder Transformation von Informationen durch den Einsatz von EDV teilweise automatisiert ist. In der betrieblichen Praxis besteht es typischerweise aus einer Menge unabhängiger Systeme, die zusammen die angestrebte Leistung erbringen (*KIS: kooperatives Informationssystem*).

1. Als „System im weiteren Sinne“ gilt (a) eine Menge von Elementen (Systembestandteilen), die (b) durch bestimmte Ordnungsbeziehungen miteinander verbunden und (c) durch klar definierte Grenzen von ihrer Umwelt geschieden sind; von „Systemen im engeren Sinne“ oder „technischem System“ spricht man, wenn sowohl die Außenwirkungen des Systems insgesamt wie auch seine Binnenstruktur (d. h. die Ordnungsbeziehungen der Systembestandteile) durch Zielfunktionen bestimmt sind (H. Wedekind).

Rechnergestützte Informationssysteme

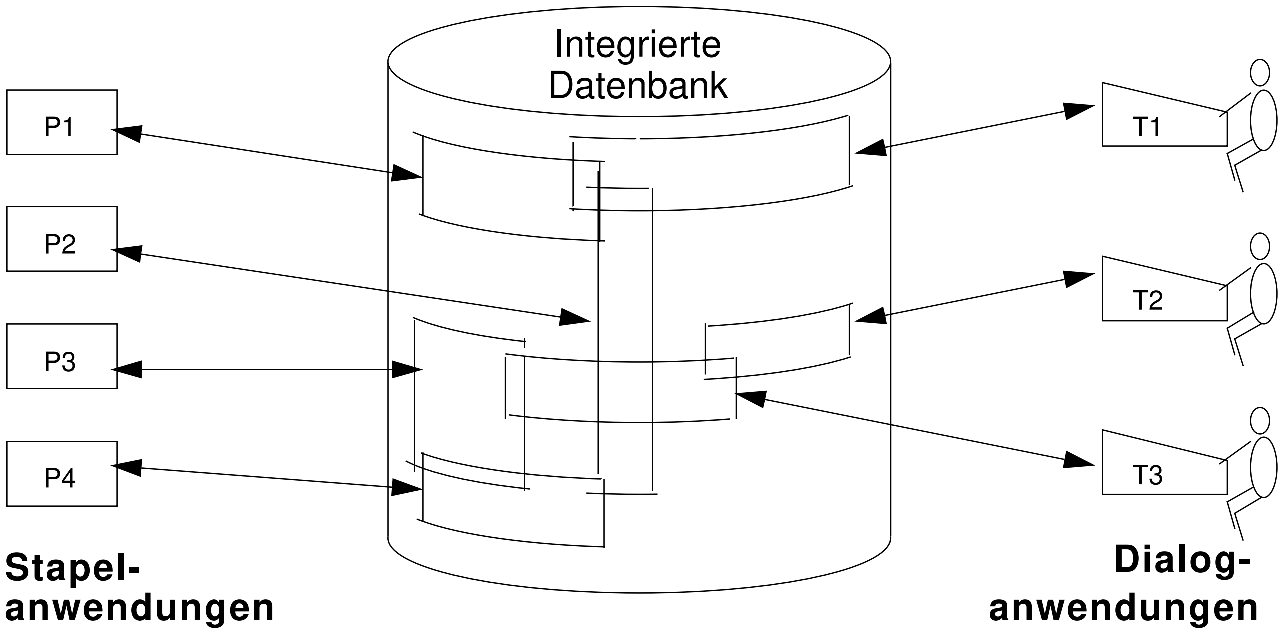
Datenbanksysteme (DBS): (zentrale) Hilfsmittel für KIS



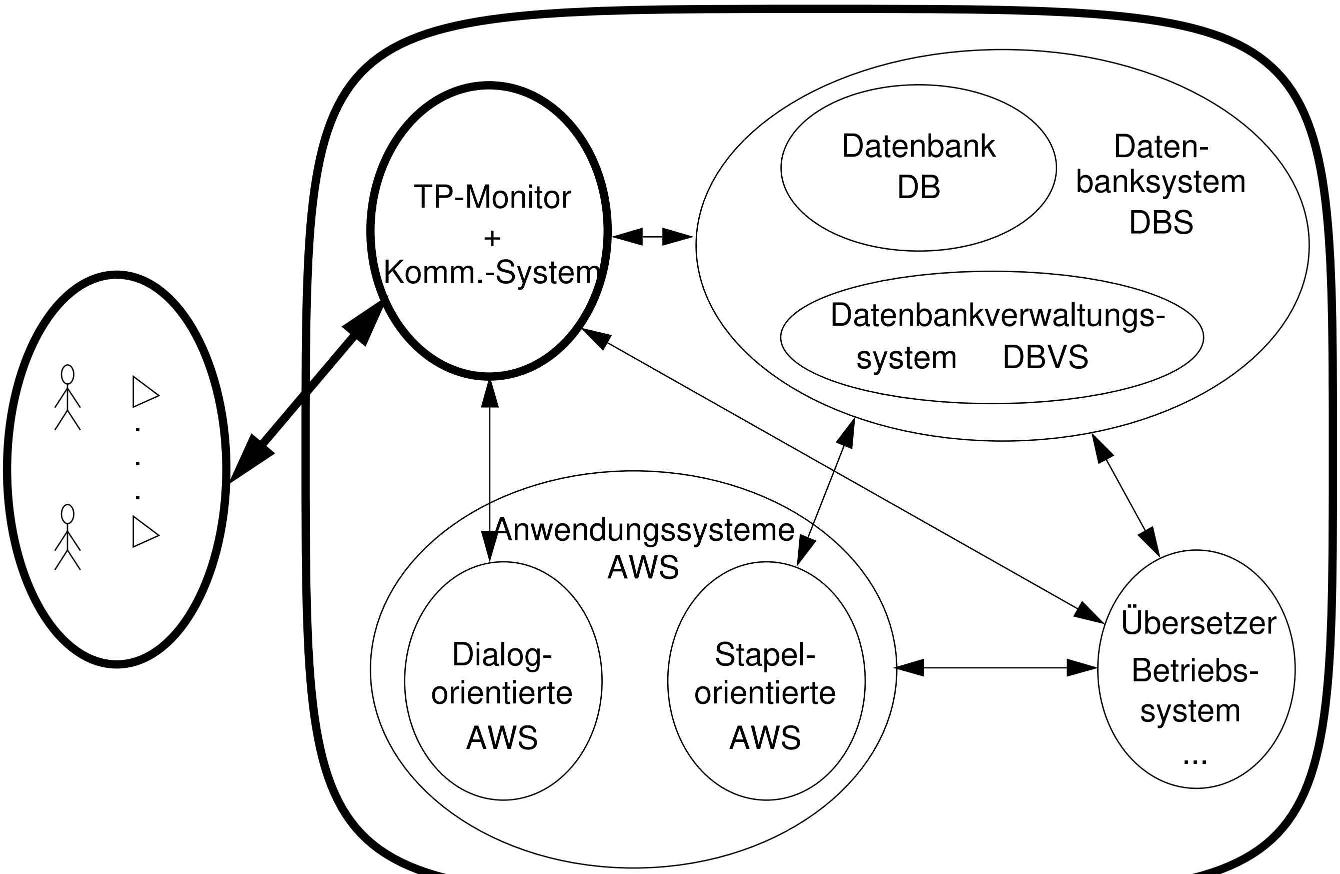
DBS = DB + Datenbankverwaltungssystem (DBVS, DBMS)

Eine ***Datenbank*** ist eine Sammlung gespeicherter operationaler Daten, die von den Anwendungssystemen eines Unternehmens benötigt werden.

Ein ***DBVS*** ist ein standardisiertes Softwaresystem zur Definition, Verwaltung, Verarbeitung und Auswertung der DB-Daten. Es kann mittels geeigneter Parametrisierung an die speziellen Anwendungsbedürfnisse angepaßt werden (hochgradig generisches System).



Erweiterung der Sicht eines KIS

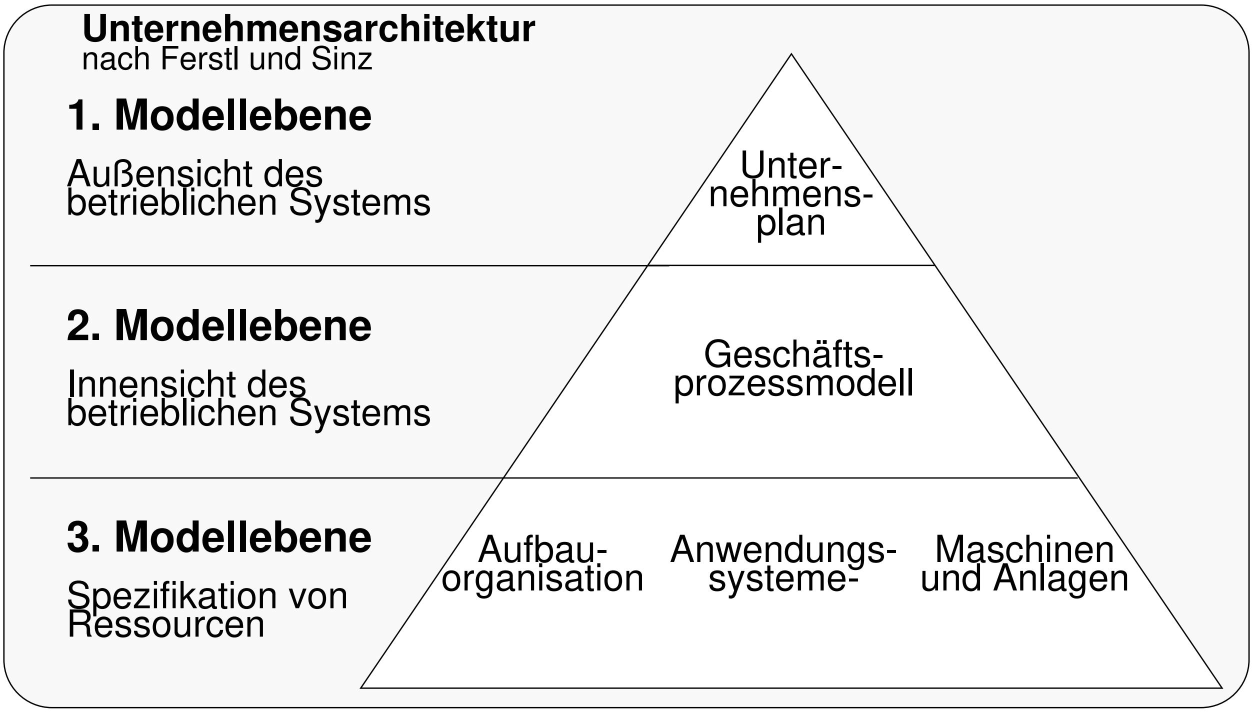


Dialogorientierte AWS: Transaktionsanwendungen mit Interaktion des Endbenutzers, harte Zeitrestriktionen, Terminal-E/A für Anforderungen und Ergebnisse

Stapelorientierte AWS: (Transaktions-) Anwendungen ohne Interaktion des Endbenutzers, keine Zeitrestriktionen, Datei-E/A für Anforderungen und Ergebnisse

Aufgaben betrieblicher Informationssysteme

- **Betriebliche Informationssysteme**
  - zentraler und strategisch bedeutsamer Bestandteil von Unternehmen
  - systematischer Aufbau und Ausrichtung an den Unternehmenszielen
  - Orientierung an Unternehmensarchitektur als dreischichtiger Pyramide



- **Administrative und operative Ebene als Fundament**
  - wird gebildet von Mitarbeitern (eingegliedert in Organisationseinheiten und Stellen einer Aufbauorganisation), Anwendungssystemen sowie Maschinen und Anlagen
  - Interaktion zwischen ihnen dient der Verfolgung definierter Ziele
- **Planungs- und Kontrollebene**
  - Interaktionen der 3. Ebene werden in Form von Geschäftsprozeßmodellen formuliert
  - Erreichen der Geschäftsziele wird überwacht
- **Strategische Ebene**

  Unternehmensplan formuliert Ziele und weitere Randbedingungen, die durch Ausführung von Geschäftsprozessen erreicht bzw. eingehalten werden sollen

Aufgaben betrieblicher Informationssysteme (2)

- **Unterscheidung nach Aufgaben des**
  - betrieblichen Lenkungssystems (Planung, Steuerung und Kontrolle)
  - betrieblichen Leistungssystems (Administration, Disposition und Durchführung)

- **Aufgaben und Typen (Beispiel)**
  - Administrationssysteme dienen der Rationalisierung und (Teil-)Automatisierung vorhandener Abläufe
  - Dispositionssysteme sollen die kurzfristige betriebliche Entscheidungsfindung vereinfachen/übernehmen
  - Planungssysteme unterstützen die mittel- bis langfristige Entscheidungsfindung (Erzeugung alternativer Pläne, weitreichndere Auswirkung)
  - Kontrollsysteme dienen dem Erkennen von außergewöhnlichen und daher bemerkenswerten Situationen (Datenkonstellationen). Sie erhalten von Administrationssystemen Ist-Daten, um Ist-Soll-Abweichungen erkennen zu können
    - ➥ hier: Erarbeitung der technischen Grundlagen

- **Was sind typische Aufgabenbereiche?**

Typ/Aufgabe	Produktion	Beschaffung Lagerhaltung	Vertrieb	Personal
Administration	Betriebsdaten-erfassung	Lagerver-waltung	Kundenver-waltung	Personal-verwaltung
Disposition		Bestell-disposition		
Kontrolle	Fertigungs-leitstand		Touren-planung	
Planung	Absatz-/Kapa-zitätsplanung		Marketing-planung	Personalein-satzplanung

Zielsetzungen für betriebliche IS

- **Anforderungen an ein Informationssystem in einem Unternehmen sind unterschiedlich, je nachdem, ob Aufgaben**
  - der operierenden Ebene (Sachbearbeitung)
  - der planenden Ebene (mittleres Management)
  - der strategischen Ebene (Unternehmensleitung)

  zu lösen sind

- **Verbesserung aller Prozesse und Aufgaben der operierenden Ebenen** des Unternehmens durch Auskunfts-, Berichts-, Buchungs-, Produktions-, Steuerungs-, Vertriebs- und Anwendungssysteme

  Kennzeichen: *Verarbeitung großer Datenmengen und große Änderungshäufigkeit der Daten*

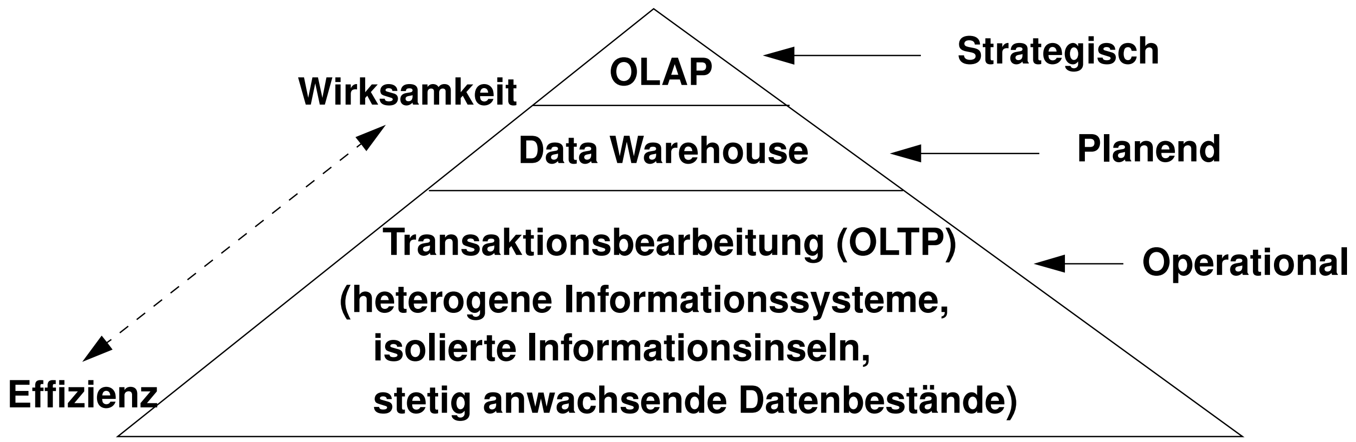
- **Unterstützung und evtl. Teilautomatisierung aller Prozesse und Aufgaben der planenden Ebene durch**:
  - benutzerorientierte Bereitstellung von Informationen
  - Suche und Auswertung von Daten im Dialog
  - Automatisierung von Routine-Entscheidungen
  - Einsatz von mathematisch-statistischen Methoden

  Kennzeichen: teilweise unvorhersehbarer Informationsbedarf, verdichtete Daten, kein Änderungsdienst

- **Unterstützung der strategischen Ebene** durch Bereitstellung von Daten für einen **überwiegend nicht vorhersehbaren Informationsbedarf**

Wichtige Anwendungsklassen betrieblicher IS

- **Anwendungsklassen – Terminologie**
  - OLTP (*On-Line Transaction Processing*)
    Abwicklung von Transaktionen auf den operationalen Daten
  - DW (*Data Warehouse*)
    als themenorientierte, zeitlich veränderliche, nicht-flüchtige Datensammlung (vorwiegend geschichtliche Daten)
  - OLAP (*On-Line Analytical Processing*):
    Analyse betrieblicher Datenbestände
    - viele Anwendungsfelder: Banken, Telecom, Handel, Versicherungen, ...
    - Einsatz von „Grabungstechniken für Wettbewerbsvorteile“
  - DSS (*Decision Support System*)
- **Schlagwörter**
  - Business Intelligence: Intelligente Nutzung großer Datenbanken
  - Data Mining: Aufspüren von inhärenten Daten-/Informationsmustern aus großen dynamischen Datenbeständen
    “In Data Mining applications, not only does the system define the semantics, it actually defines the queries. The user simply says ‘Go’, and the system produces what it believes to be useful answers.”
  - KDD (*Knowledge Discovery in Databases*), oft synonym zu Data Mining
- **Informationspyramide**



Daten in Informationssystemen
– strukturierte Daten

- **Strukturierte oder formatierte Daten:**
  (NAME = „Müller“, TAETIGKEIT = „Kalligraph“, GEBDAT = „780623“, . . .)
  - maximale Länge (= endlicher Wertevorrat)
  - Werte von Variablen, Feldern, Attributen; durch Namen beschrieben
  - Bedeutung weitgehend vorgegeben, relativ geringer Informationsgehalt

- **Klassische Datenbanktechnik**
  - <u>formatierte</u> Datenstrukturen, feste Satzstruktur
  - Beschreibung der Objekte durch Satztyp, Attribute und Attributwerte ($S_i/A_j/AW_k$)
  - jeder Attributwert AW_k wird durch Beschreibungsinformation (Metadaten) A_j und S_i in seiner Bedeutung festgelegt
  - **Beispiel:** Tabelle (Relation in Tabellendarstellung)

Schema

ANGESTELLTER

Satztyp (Relation)

Ausprägungen

PNR	NAME	TAETIGKEIT	GEHALT	ALTER
496	PEINL	PFOERTNER	2100	63
497	KINZINGER	KOPIST	2800	25
498	MEYWEG	KALLIGRAPH	4500	56

➥ **DB-Schema**: vollständige Strukturbeschreibung (Metadaten) ist vor der Speicherung von Objekten zu spezifizieren und dem DBS bekannt zu machen

- **Art der Anfrage und Aktualisierung**
  - deskriptiv (nicht-prozedural)
  - mengenorientiert

Daten in Informationssystemen
– unstrukturierte Daten

- **Unstrukturierte oder unformatierte Daten:**
  „Er heißt Müller. Er wird in seiner Arbeitszeit überwiegend als Kalligraph eingesetzt. Geboren ist er am 23. Juni des Jahres 1978 . . .“
  - beliebige Länge
  - teilweise selbstbeschreibend
  - Bedeutung nur schwach vorgegeben
  - hoher Informationsgehalt

- **Dokumente in Information-Retrieval-Systemen**
  - unformatierte Daten, keine dem IRS bekannte Dokumentstruktur
  - Beschreibung der Objekte durch Dokumenttyp und Wert (D_i/W_k)
  - Es gibt keine nähere Beschreibung oder Spezifikation von Struktur und Semantik, die W_k in seiner Bedeutung festlegt
    - ➥ IRS verwaltet „lange“ Werte (z. B. Texte eines Abschnitts, Kapitels oder Buches) und stellt dafür Container (verschiedenen Typs) zur Verfügung

- **Aufgaben/Eigenschaften von IRS**
  - Verwaltung von Dokumenten, Büchern, Abstracts usw.
  - effiziente Suche in großen Datenmengen
  - typischerweise nur Retrieval im Mehrbenutzerbetrieb
  - Anfragesprache für Retrieval
    (➥ Annäherung an natürliche Sprache erwünscht)

Daten in Informationssystemen
– semi-strukturierte Daten

- **HTML-Dokumente im WWW**
  - Aufbau des Dokumentes (syntaktische Struktur) ist festgelegt
  - Formatierungsanweisungen (Tags) lassen grobe Rückschlüsse auf den Inhalt des Dokumentes zu: TITLE, HREF, ADDRESS, ...
  - Inhalt des Dokumentes ist jedoch nicht weiter beschrieben; es gibt keine Metadaten, die die Bedeutung genauer festlegen

➥ WWW-Browser kann HTML-Dokumente aufbereiten und graphisch darstellen, ohne den Inhalt zu kennen

- **Beispiel: HTML-Dokument, semi-strukturiert**

```
<!DOCTYPE HTML PUBLIC „-//W3C//DTD HTML 3.2//EN“>
<HTML>
<HEAD>
  <TITLE>Publications 1998</TITLE>
  <META NAME=“GENERATOR“ CONTENT=“Mozilla/3.01Gold (X11; I; SunOS 4.1.3 sun4m) [Netscape]“>
</HEAD>
<BODY BACKGROUND=“http://wwwdbis.informatik.uni-kl.de/pics/paper.jpg“>
<P><BASE HREF=“http://wwwdbis.informatik.uni-kl.de/publications/“><A HREF=
“http://wwwdbis.informatik.uni-kl.de/“ target=“_top“>[Top]</A>
<A HREF=“publications.html“>[Up]</A> <A HREF=“p1997.html“>[Next]</A> [Deutsche Version]<BR>
<H1 ALIGN=CENTER>Publications 1998</H1>
<HR NOSHADE></P>
<H3>Last update: 11/30/98 </H3>
<P>De&szlig;loch, S., H&auml;rder, T., Mattos, N., Mitschang, B., Thomas, J.:<BR>
<A HREF=“http://wwwdbis.informatik.uni-kl.de:18070/publications/DHMMT98.VLDB.html“>Advanced Data
Processing in KRISYS: Modeling Concepts, Implementation Techniques, and Client/Server Issues</A>,
in: VLDB Journal 7:2, 1998, pp. 79-95.</P>
<ADDRESS><A HREF=“mailto:wwwhaerder@informatik.uni-kl.de“>wwwhaerder@informatik.uni-kl.de</A>
</ADDRESS>
</BODY>
</HTML>
```

➥ **Präsentation** an der Benutzerschnittstelle**:** für Menschen, nicht für Programme

Daten in Informationssystemen
– semi-strukturierte Daten

- **Formatierungssprachen dienen dem Austausch von Dokumenten**
  - Es gibt eine Vielzahl von Formatierungsprachen, die alle den internationalen Standard zur Textverarbeitung SGML als Meta-Sprache benutzen, um ihre Formate und Grammatik zu definieren

- **HTML**
  - ist eine Sprache zur Formatierung (Strukturierung) von Dokumenten (Texten) (HyperText Markup Language, Tag Language)
  - bietet eine vorgegebene Menge von Begrenzungs- und Formatierungs-anweisungen (>200) mit standardisierter Bedeutung
  - **Beispiel**

    <H2>Second-Level heading </H2>
    <P>This is a passage of text that probably belongs to the heading immediately above </P>
  - vermischt **Strukturierungs- und Darstellungsaufgaben**
  - kann die Suche von Dokumenten kaum unterstützen

- **DocBook**
  - weitere Sprache zur Textformatierung (Software-Dokumentation)
  - **Beispiel**

    <SECT2>
    <TITLE>Second-level heading </TITLE>
    <PARA> This is a passage of text that certainly belongs to heading above. We know this because both are contained in the same SECT2 element.
    </PARA>
    </SECT2>

  ➥ Jede Sprache ist auf eine bestimmte Kategorie von Dokumenten zuge-schnitten

Exkurs – XML

- **Neue Anforderungen**
  - **Erweiterbarkeit**, um nach Bedarf neue „Tags“ zu definieren
  - **Struktur**, um komplexe Daten zu modellieren und abzubilden
  - **Validierung**, um die strukturelle Korrektheit der Daten zu überprüfen
  - **Medienunabhängigkeit**, um Inhalte in verschiedenen Formaten zu publizieren
  - **Hersteller- und Plattformunabhängigkeit**, um entsprechende Dokumente mit standardisierter SW zu verarbeiten

- **Neuer Ansatz: XML (Extensible Markup Language)**
  - XHTML1.0 $\in$ XML $\subset$ SGML
  - XML ist eine Metasprache, die Definition von Tags erlaubt
  - XML-Tags haben keine vordefinierte Semantik (wie bei HTML)

- **Beispiel: wohl-geformtes XML-Dokument**

```
<book ISBN=“1575213346“>
        <title>Presenting XML</title>
        <author><firstname>Richard</firstname><lastname>Light</lastname></author>
        <author><firstname>Tim</firstname><lastname>Bray</lastname></author>
        <date>Sept.1997</date>
        <price currency=“USD“>19.99</price>
        <comment rating=“4“>
                <writtenby>
                        <firstname>Harald</firstname><lastname>Schöning</lastname>
                </writtenby>
                <text>A quite useful book for <userclass>beginners</userclass>.</text>
        </comment>
        <comment>
                <writtenby>
                        <firstname>A</firstname><lastname>Reader</lastname>
                </writtenby>
                <text> I did not like the cover</text>
        </comment>
</book>
```

➥ Die **Semantik** der Tags muß in Anwendungskontexten speziell festgelegt werden, z. B. durch Programme, Skripts oder deklarative Anweisungen für Formatvorlagen (*style sheets*)

Exkurs – XML (2)

- **XML**[1]
  - ist eine vereinfachte Form von SGML (und nicht eine erweiterte Form von HTML). Es wurden einige (für Web-Browser schwierige) Konstrukte von SGML weggelassen
  - erlaubt die Definition einer beliebigen Anzahl von Formatierungssprachen für verschiedene Zwecke (Kategorien von Dokumenten)
    (z. B. Molekülstrukturen, japanische Texte, 3D-Objekte usw.)
  - dient zur Beschreibung von Struktur und Inhalt von Dokumenten (selbstbeschreibend)
  - kann als Sprache (Modell) zur Darstellung und zum Austausch von Dokumenten aufgefaßt werden

➥ Damit wird der **Austausch von Dokumenten** zwischen Programmen möglich und ihre **Interoperabilität** gefördert

➥ **Organisation, Suche** und **Datenintegration** werden drastisch verbessert, da XML Inhalte genauer spezifizieren kann

- **Ziele**
  - XML kann für Daten erreichen, was Java für Programme bietet:
    Unabhängigkeit von Plattform (und Hersteller)

    ➥ Austausch von **Daten <u>und</u> Metadaten**

  - XML soll universelles, medienunabhängiges Publikationsformat bieten
    (für alle Klassen von Benutzern und alle Sprachen)

    ➥ „a single, completely internationalized format of almost unlimited power for both print and online publishing that is fully interoperable across all products and platforms“

1. http://www.w3.org/

Daten in Informationssystemen – Multimedia-Daten

- **Bedarf**

  Verwaltung großer Mengen von Multimedia-Datenobjekten auf verschiedenartigen Datenträgern (optischen Speichern, Videobändern, ...), so daß sie für möglichst viele Anwendungen auffindbar und zugreifbar (nutzbar) sind

- **Was sind Multimedia-Datenobjekte?**

  Digitalisierte, im Rechner abgelegte Bilder und Tonaufnahmen sowie Texte und Graphiken; Videoaufzeichnungen, Hologramme, Radarsignale, ...

- **Ziel**

  - einerseits Erhöhung der „Informations-Bandbreite“,
    bessere Benutzerschnittstellen
  - aber auch mehr Information im System:

    ➥ Inhalt eines Bildes (einer Tonaufnahme, ...) ist niemals vollständig in Medien wie Text oder Graphik wiederzugeben

- **Bezeichnungen**

  - **Medienobjekt** (oder Medien-Datenobjekt)
    Ein Datenobjekt, das einem einzigen Medium angehört, also ein einzelnes Bild ein Textstück
  - **Multimedia-Objekt**
    (Multimedia-Datenobjekt, auch „mixed-mode object“)
    Aggregation (Komposition) von Medienobjekten unterschiedlichen Typs, z. B. Video (Bild + Ton)
  - **Multimedia-Daten**
    Sammelbegriff für Medienobjekte und Multimedia-Objekte

- **Medienobjekte**
  sind aus formatierten und unformatierten Daten zusammengesetzt

Exkurs - Medienobjekte

- **Was ist eigentlich alles zu speichern?**
  - **Rohdaten**

    unformatiert: lange Folge (Menge, ...) von kleinen Elementen
    (Bits, Buchstaben, Pixel, Linien, Energieniveaus, ...)

  - **Registrierungsdaten** (Steuerungsdaten)

    obligatorisch
    erforderlich für korrekte **Interpretation** und **Identifikation** der Rohdaten

  - **Beschreibungsdaten**: optional

  - oft redundant:
    Darstellung der **Struktur** und/oder des **Inhalts** in einem anderen Medium, formatiert oder unformatiert

- **Operationen auf Medienobjekten**
  - **Erzeugung** (Eingeben, Erfassen, capture)
    mit Gerät – aus Programm – aus Datei
  - **Ausgeben** (Zeigen, Präsentieren, present)
    auf Gerät – an Programm – in Datei
  - **Bearbeiten** (Modifizieren, Editieren)
  - **Zusammensetzen**
    Erzeugen multi-medialer Objekte
  - **Weitergeben** (Versenden)
  - **Archivieren**
  - **Auswerten** (Aggregieren, Ableiten)
    Filtern, Analysieren, Erzeugen von Beschreibungsdaten
  - **Suchen** (Vergleichen)
    Mustererkennung / Ähnlichkeitssuche auf Rohdaten;
    oder inhaltsorientierte Suche auf Beschreibungsdaten

Beispiel: Medienobjekt „Rasterbild“

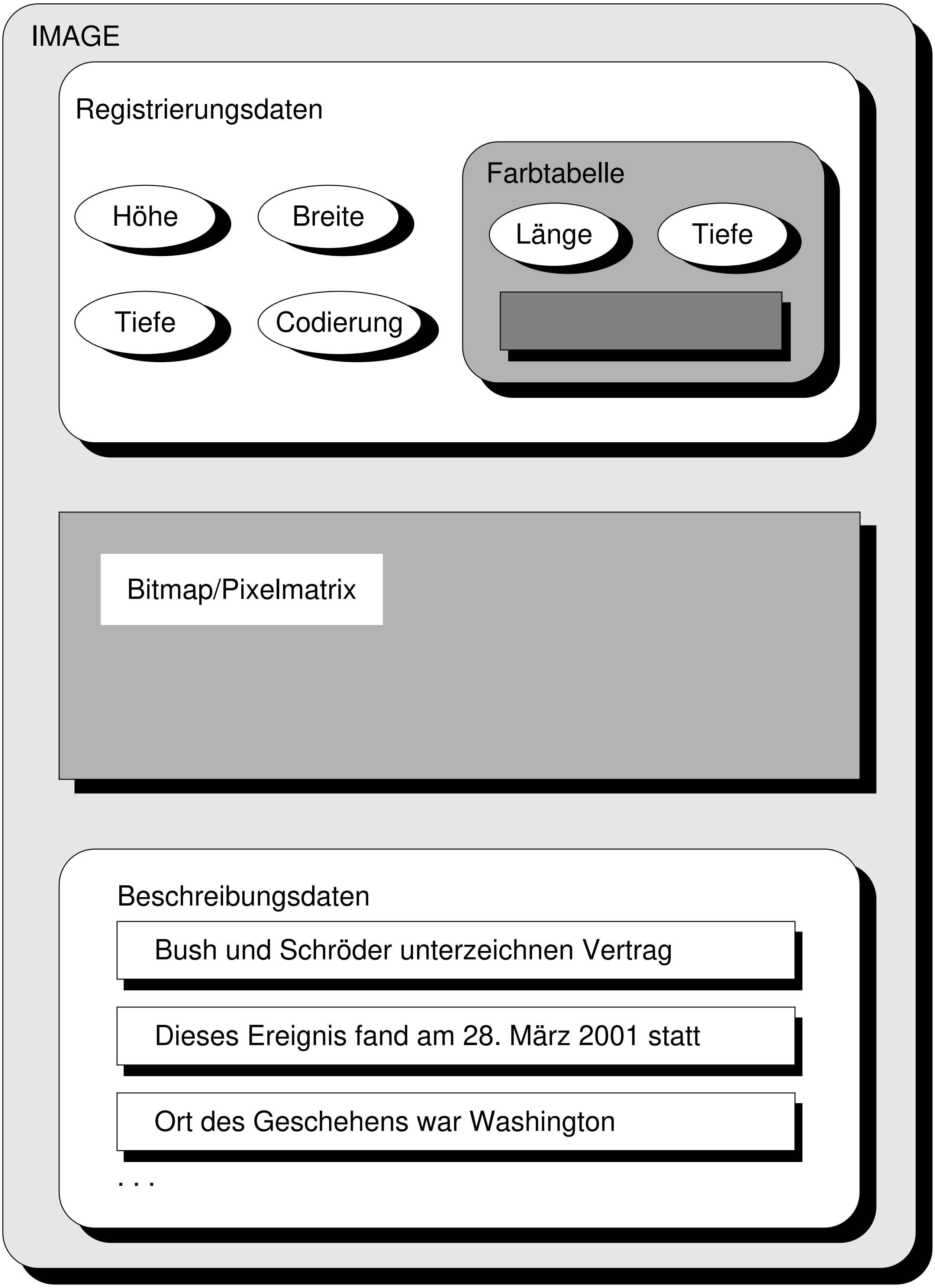
- **Rohdaten**
  - Matrix von Bildpunkten
    (Pixel = Picture Element, auch „Pel“ genannt)

- **Registrierungsdaten**
  - Anzahl Bits pro Pixel („Farbtiefe“, typisch: 1, 8, 24);
  - Anzahl Pixel pro Zeile (Breite des Bildes);
  - Anzahl der Zeilen (Höhe des Bildes);
  - Art der linearen Abspeicherung
    - zeilenweise oder spaltenweise;
  - Bedeutung eines Pixels:
    - Grauwert, Farbdefinition, Index einer Farbtabelle
      (Palette, Colormap);
  - ggf. Farbtabelle mit bestimmter Anzahl von Einträgen und Länge dieser Einträge (meist 24 Bit)
  - ggf. Art der Farbdefinition: RGB, IHS, YIQ, ...
    und anderes mehr

- **Beschreibungsdaten**
  - Text, Schlüsselworte, Wissensrepräsentation
  - Graphik: erkannte Linien und Flächen
  - darauf aufbauend zweidimensionale Objekte wie Kreise, Rechtecke usw.

Beispiel: Medienobjekt „Rasterbild“

Abstrakte Sicht auf eine Instanz vom Datentyp IMAGE:



Operationen auf Rasterbildern

- **Erzeugen** (Eingeben, Erfassen, capture)
  - mit Gerät: Scanner, Kamera
  - aus Programm: Hauptspeicherstrukturen für Rasterbilder
    (Ximage, Pixrect, ...)
  - aus Bild-Datei
    (GIF, TIFF, JPEG, PNG, GEM-Image, PBM, ...)

- **Ausgeben** (Zeigen, Präsentieren, present)
  - auf Gerät: Bildschirm, Drucker
  - an Programm: in Hauptspeicherstruktur
  - in Datei: Wahl eines bestimmten Formats

- **Bearbeiten** (Modifizieren, Editieren)
  - pixel-orientierte Zeichensysteme („Paintbrush-Systeme“)
    MacPaint, PhotoShop, ...
  - Setzen einzelner Pixel;
    Ändern der Farbtabelle;
    Bitmap-Operationen: Überlagern anderer Bilder
    (oft unerwünscht aus dokumentarischen Gründen)

Operationen auf Rasterbildern (2)

- **Weitergeben** (Versenden)
  - Nachrichtenformate (Austauschformate) für Bilder
    MIME, IIF

- **Archivieren**
  - Bild-Datenbanken (Pictorial Databases)

- **Auswerten** (Aggregieren, Ableiten)
  - Kontrastverstärkung, Vergrößern oder Verkleinern (Zoom)
  - Umsetzung Farbe nach Grauwert und Ausschnittbildung (window) – erzeugen wieder Bilder
  - Linienerkennung – erzeugt Graphik

- **Suchen** (Vergleichen)
  - Mustererkennung (Plattern Matching):
    wird unterstützt durch Zugriffspfade (*Iconic Indexing*), wichtig für einige Bildanalysetechniken, allgemein aber nicht sehr aussagekräftig
  - Bild-Bild-Vergleich:
    z. B. Phantombild oder Photo mit Archivbild zur Identifikation von Personen
    Ähnlichkeitsmaße? – schwierig

Zusammenfassung

- **Transakionsparadigma**
  - macht weitreichende Zusicherungen für die Verarbeitung von DB-Daten
  - ACID-Eigenschaften müssen in einer Rechnerumgebung (aufwendig) nachgebildet werden
  - erlaubt die Implementierung von „Vertragsrecht"

- **Information und Informationssysteme**
  - Daten: objektive Welt der nicht-interpretierten Daten
  - Information: subjektive Welt der bewerteten Daten
  - Heterogenität, Wachstum, Anforderungsvielfalt u. a. führen oft auf unabhängige IS, die zusammen als kooperatives IS die angestrebte Leistung erbringen müssen

    ➥ „grob": DBS + AWS = KIS
  - wichtige Anwendungsklassen für
    - operierende Ebene: OLTP
    - planende/kontrollierende Ebene: DW, OLAP
    - strategische Ebene: OLAP, DSS

- **Daten in Informationssystemen**
  - **strukturierte** Daten
    - fest vorgegebene Satz-/Tabellenstruktur
    - Bedeutung durch Metadaten weitgehend vorgegeben
  - **unstrukturierte** Daten
    - „lange" Werte bilden Dokumente (typischerweise in IRS)
    - Bedeutung nur schwach vorgegeben: Interpretation durch benutzerdefinierte Funkionen
  - **semi-strukturierte** Daten
    - HTML: Vermischung von Strukturierung und Präsentation
    - Verbesserung durch XML: Datenaustausch, Interoperabilität, Datenintegration
  - **Multimedia**-Daten
    - Beschreibung durch Rohdaten, Registrierungsdaten, Beschreibungsdaten
    - Verarbeitung durch medienspezifische Operationen

Beispiele für Informationssysteme

- **Informationssystem einer Universität**

Die Universitätsdatenbank ist die Sammlung aller für die Abwicklung der an einer Universität anfallenden Verwaltungsaufgaben benötigten Daten.

Eine Universität gliedert sich i. allg. in mehrere Fachbereiche, denen sowohl die Studenten als auch die Professoren zugeordnet sind.

Die Studenten belegen verschiedene Vorlesungen von Professoren und legen bei ihnen Prüfungen ab.

Typische Anwendungen sind z. B.:
Immatrikulation der Studienanfänger, Rückmeldung der Studenten, Ausfertigen von Studentenausweisen und Studienbescheinigungen, Stundenplanerstellung und Planung der Raumbelegung, Ausstellen von (Vor)diplomzeugnissen, Exmatrikulationen, Statistiken über Hörerzahlen, Raumauslastung, Prüfungsergebnisse, etc.

- **Informationssystem eines Produktionsbetriebes**

In einem Produktionsbetrieb werden Daten über die verschiedenen Abteilungen und deren Beschäftigte mit ihren Familienangehörigen gespeichert.

Die Angestellten arbeiten an verschiedenen Projekten mit. Jedes Projekt benötigt für seine Durchführung bestimmte Teile. Jedes Teil kann von Lieferanten bezogen werden. Die Projekte werden jeweils von einem Projektmanager geleitet.

Die in einem Betrieb hergestellten Endprodukte setzen sich i. allg. aus mehreren Baugruppen und Einzelteilen zusammen.

Typische Anwendungen sind z. B.:
Einstellung und Entlassung von Personal, Lohn- und Gehaltsabrechnung, Bestellung und Lieferung von Einzelteilen, Verkauf von Fertigprodukten, Lagerhaltung, Bedarfsplanung, Stücklistenauflösung, Projektplanung.

Beispiele für Informationssysteme (2)

- **Informationssystem einer Fluggesellschaft**

Eine Fluggesellschaft fliegt verschiedene Flughäfen an. Auf diesen Flugstrecken werden Flugzeuge bestimmter Typen mit dafür ausgebildetem Personal eingesetzt. Die Piloten haben Flugscheine jeweils nur für einige wenige Flugzeugtypen. Außer den Piloten gibt es noch anderes Bord- sowie Bodenpersonal.

Die Flugbuchungen der Passagiere sowie das Anfertigen der Passagierlisten werden ebenfalls automatisiert durchgeführt.

Typische Anwendungen sind z. B.:
Flugbuchungen von Passagieren, Personaleinsatzplanung, Materialeinsatzplanung, Flugplanerstellung, Überwachung der Wartelisten, Gehaltsabrechnung.

- **Informationssystem einer Bank**

Eine Bank gliedert sich gewöhnlich in mehrere Zweigstellen auf. Die Angestellten der Bank gehören jeweils fest zu einer bestimmten Zweigstelle. Auch die Bankkunden sind immer einer Zweigstelle zugeordnet. Es sind Daten über die verschiedenartigen Konten der Bankkunden bereitzustellen, wie z. B. Girokonten, Sparkonten, Hypothekenkonten, Kleinkreditkonten, Wertpapierkonten, etc.

Typische Anwendungen sind z. B.:
Buchung von Zahlungsvorgängen auf den verschiedensten Konten, Einrichten und Auflösen von Konten, Kreditgewährung bzw. Bereitstellen von Daten über die Kreditwürdigkeit eines Kunden, Zinsberechnung und -verbuchung, sowie alle Vorgänge der Personalverwaltung wie z. B. Gehaltsabrechnung.

Zur Rolle rechnergestützter Informationssysteme im Bankenbereich:

“In banking, by contrast, the data actually is the inventory – the two are synonymous. In increasingly many cases, the DB transaction is the financial transaction. There are no real, tangible tokens (greenbacks) moved as a result of the monetary transfer transaction. If the data is bad, money is lost or created. There is no possibility of counting the money (bits) in order to verify the status. Fiscal responsability dictates that creating or destroying money – even temporarily – is unacceptable.”
(Mike Burman, Bank of America)

Beispiele für Informationssysteme (3)

- **Straßeninformationssystem**

Beschreibung

Die Straßendatenbank ist Bestandsnachweis für das Straßennetz eines Bundeslandes.

Klassifikation

Es gibt verschiedene Straßentypen (Autobahnen, Bundesstraßen, Kreisstraßen, Gemeindestraßen etc.) innerhalb eines Bundeslandes.

Die Straßen sind aus Planungs- und Verwaltungsgründen in Abschnitte eingeteilt, die durch jeweils zwei Netzknoten, welche Abschnittsanfang und Abschnittsende markieren, definiert sind.

Organisation

Die Straßenabschnitte sind jeweils einem Bauamt zugeordnet, das für Planungsarbeiten und die geometrische Festlegung des Straßenverlaufs entlang des Abschnitts zuständig ist. Die Kosten der anfallenden Arbeiten am Straßenabschnitt trägt der Baulastträger (Gemeinde, Kreis etc.).

Die Straßenabschnitte gehen durch Gemeinden. Die Gemeinden gehören zu Kreisen.

Geometrische Darstellung

Für die geometrische Festlegung der Netzknoten ist jeweils ein Bauamt zuständig. Verschiedene Bauämter können innerhalb eines Kreises für Abschnitte oder Netzknoten zuständig sein.

Ein Straßenabschnitt kann mehrere Äste aufweisen (z. B. Aufteilung in 2 Einbahnstrecken). Ein Ast kann sich auch aus mehreren Abschnitten zusammensetzen. Eine Straße kann an einem Netzknoten unterbrochen sein und an einem anderen Netzknoten weiterführen. Ein Straßenabschnitt kann auf mehreren Straßen (z. B. Bundesstraße und Kreisstraße) gleichzeitig verlaufen.

Beispiele für Informationssysteme (4)

- **Straßeninformationssystem** (Fortsetzung)

Topologische Information

Zusätzlich sind jedem Straßenabschnitt Daten zugeordnet, welche den geometrischen Verlauf zwischen den begrenzenden Netzknoten festlegen (Trassierungselemente: Kreise, Geraden, Klothoiden).

Die Bauwerke (Brücken, Durchlässe, Signalanlagen etc.) sind dem geometrischen Verlauf des Straßenabschnitts ebenso zugeordnet wie Fußgängerüberwege, Radwege, Gehsteige, Daten des Fahrbahnaufbaus, Höheninformation, Entwässerungsschächte etc.

Besonderheiten

Unfalldaten, Verkehrsmengen, Frostsicherheit etc. sind weitere Attribute zum Straßenabschnitt.

Zeit

Die Straßendatenbank ermöglicht die Entnahme von Spezialplanungsunterlagen (z. B. Radwege, Gehwege) aber auch regionale Vergleiche des Straßennetzes und die Entnahme statistischer Daten.

Typische Fragen:

Auswahl aller Kreisstraßen im Kreis .... mit Breite < 5m und NN-Höhe > 500m.

Zusammenstellung aller Strecken mit Radwegen getrennt für Ortsdurchfahrt und freie Strecke.

Auswahl aller Bundesstraßenstrecken im Bauamt .... mit Neigungen größer als 7%.

Berechnung der befestigten Straßenfläche für alle im Jahr 1980 gebauten Bundesstraßenstrecken im Bundesland.

Beispiele für Informationssysteme (5)

- **Straßeninformationssystem** - Datenhaltung