



Universität Hamburg

50 Jahre Business Intelligence-Systeme

PROF. DR. WILH. HUMMELTENBERG
INSTITUT FÜR
WIRTSCHAFTSINFORMATIK (II)
UNIVERSITÄT HAMBURG
MAX-BRAUER-ALLE 60
22765 HAMBURG

Hamburg, Dez. 2008

Inhalt

Kap.	Überschrift	Seite
1.	Einleitung	3
2.	Methoden und Konzepte	4
2.1	Menschliche und organisationale Intelligenz	4
2.2	Entscheidungsprozesse	4
2.3	Design von Decision Support-Systemen	7
2.4	Spezifische Merkmale von Business Intelligence-Systemen	9
3.	Business Intelligence- und Decision Support-Systeme	10
3.1	Meilensteine	10
	1958: Content Management mit Text Mining und Profilvergleich	
	1978: Finanzplanungssoftware / Planungssprachen	
	1979: Spreadsheet-System VisiCalc	
	1988: Business Data Warehouse and Executive Information System	
	1989: Popularisierung von Business Intelligence	
	1992: Balanced Scorecard	
	1993: Cube Analysis / Online Analytical Processing (OLAP)	
	1996: From Data Mining to Knowledge Discovery in Databases (KDD)	
3.2	State of the Art	14
4.	Schlußbemerkungen	16
5.	Literatur	17

Bilder und Tabellen

Bild 1	Phasen und Informationsverarbeitung in Entscheidungsprozessen nach Simon	5
Bild 2	Intelligence Cycle	6
Bild 3	OODA Loop	6
Bild 4	PDCA Cycle	7
Tab. 1	Wissensarten und Erkenntnisziele im PDCA Cycle	7
Bild 5	Generische Aufgaben bei der Systemanalyse	8
Bild 6	Der Business Intelligence-Prozeß (BIP)	9
Tab. 2	Typisierungen von Business Intelligence- und Decision Support-Systemen	10
Bild 7	Anforderungen der Modelltypen an das Information Processing	13
Bild 8	Benutzergruppen und BI-Analysetechniken (ohne Content Management)	15

1. Einleitung

Vor 50 Jahren prägte der IBM-er Hans Peter Luhn 1958 bei seinen Forschungen zur selektiven Verbreitung von Informationen den Begriff „Business Intelligence System“ (BIS) mit folgenden Inhalten [Luhn58, S. 314]:

- Business: „Business is a collection of activities carried on for whatever purpose, be it in science, technology, commerce, industry, law, government, defense, et cetera.”
- Intelligence: “Intelligence is ... the ability to apprehend interrelationships of presented facts in such a way as to guide towards a desired goal.”
- Business Intelligence System: “An automatic system is being developed to disseminate information to the various sections of any industrial, scientific, or government organization. This intelligence system will utilize data processing-machines for auto-abstracting and auto-encoding of documents and for creating interest profiles for each of the ‘action points’ in an organization. Both incoming and internally generated documents are automatically abstracted, characterized by a word pattern, and sent automatically to appropriate action points.”

Nach heutiger Terminologie leistete das BI-System Text Mining mittels statistischer Verfahren zur Klassifikation von Dokumenten, um auf Basis eines regelbasierten Vergleichs des Inhaltsprofils mit den Wissensanforderungen an den Action Points zu entscheiden, wer über welche Dokumente informiert wird. Seitdem ist die BIS-Geschichte von zahlreichen technologisch bedingten Hypes geprägt. Dies mag daran liegen, dass häufig unter vereinfachenden Annahmen rasche Antworten und Problemlösungen für komplexe Aufgaben und Managementprozesse versprochen und durch euphorische Begriffsbildungen übertriebene Erwartungen geweckt werden. Hier ist die Wissenschaft aufgerufen, klärend und ordnend einzugreifen. Es verbietet sich deshalb eine einfache Historiographie, die nur aufzeigt „what happened“; vielmehr ist zu klären „why and how it happened“. Dies schließt eine Betrachtung der methodischen und konzeptionellen Grundlagen ein.

Bereits 1967 wies Ackhoff in seinem Beitrag „Management Misinformation Systems“ [Acko67] darauf hin, dass die Implementierung von Informationssystemen häufig von folgenden irrigen Annahmen ausgeht:

- (1) the critical deficiency under which most managers operate is the lack of relevant information
- (2) the manager needs the information he wants
- (3) if a manager has the information he needs his decision making will improve
- (4) better communication between managers improves organizational performance
- (5) a manager does not have to understand how his information systems works, only how to use it.

Um Fehleinschätzungen zu vermeiden, schlägt Ackoff vor, Managementinformationssysteme (MIS) in Managementkontrollsysteme (MKS) einzubinden. Entsprechend sind Business Intelligence-Systeme (BIS) in Business Intelligence Cycles (BIC) und diese wiederum in Business Decision and Management-Prozesse (BDMP) einzubinden. Hier soll nicht die BIS-Historie im Detail verfolgt werden; statt dessen sei auf [ChGl06; GlKe06; KeBa06; Powe07] verwiesen. Vielmehr wird der aktuelle Entwicklungsstand im Kontext von Decision Support-

Systemen (DSS) sowie der Managementaufgaben und –konzepte, zu deren Unterstützung sie entwickelt werden, anhand von Meilensteinen für Konzepte und Design von BIS beleuchtet.

2. Methoden und Konzepte

2.1 Menschliche und organisationale Intelligenz

Menschliche Intelligenz erkennt man daran, wie ein Mensch spricht, denkt und sich artikuliert. Er wird allgemein als intelligent bezeichnet, wenn er

1. komplexe Informationen rasch erfassen (Wahrnehmung und Speicherung),
2. auf Informationen adäquat reagieren und Erkenntnisse umsetzen (Flexibilität und Kreativität),
3. rasch lernen und weitere Konsequenzen ziehen kann (Lernfähigkeit).

Entsprechend läßt sich eine Organisation als intelligent bezeichnen, wenn sie

1. eine hohe Wissbegierde und Wissensaufnahmefähigkeit besitzt,
2. auf Veränderungen konsistent und flexibel zu reagieren vermag und
3. rasch lernfähig ist.

Im Gegensatz zum Menschen sind Organisationen soziotechnische Systeme mit je nach Größe und Zielbestimmung unterschiedlicher Komplexität. Bei ihnen besteht kein zwangsläufiger Zusammenhang zwischen der Intelligenz bzw. Lernfähigkeit ihrer Mitglieder sowie der kollektiven Intelligenz und dem kollektiven Lernen der Organisation. Organisationale Intelligenz (OI) resultiert nicht aus der algebraischen Summe der Intelligenz ihrer Mitglieder und sonstigen Artefakte (Maschinen, Prozesse, Programme etc.), sondern entsteht erst durch deren zielorientiertes, organisches Zusammenwirken im Sinne der OI-Elemente [s. Humm08a, b].

2.2 Entscheidungsprozesse

Simon schuf 1942 mit seinen Untersuchungen über Entscheidungsprozesse in kommunalen Verwaltungen die konzeptionellen Grundlagen für die Entwicklung von Decision Support-Systemen (DSS). Er konsolidierte seine Ergebnisse in [Simo47] „Administrative Behavior“ mit dem Modell für Entscheidungsprozesse aus den Phasen (1) Intelligence, (2) Design, (3) Choice, (4) Review. Im Idealfall impliziert das Modell:

- In Phase (1) identifizieren Entscheidungsträger alle möglichen Alternativen.
- In Phase (2) ermitteln sie mögliche Konsequenzen in Form von Ereignissen und Ergebnissen mit zugehörigen Eintrittswahrscheinlichkeiten.
- Die prognostizierten Ergebnisse werden in Phase (3) vergleichend bewertet. Phase (3) schließt die anschließende „Action“ ein.
- Phase (4) dient der Steuerung und Kontrolle sowie dem Sammeln von Erfahrungen (Lernen).

Wird die Notwendigkeit von Systemveränderungen offenbar, so wird das Phasenmodell zu einem Zyklus.

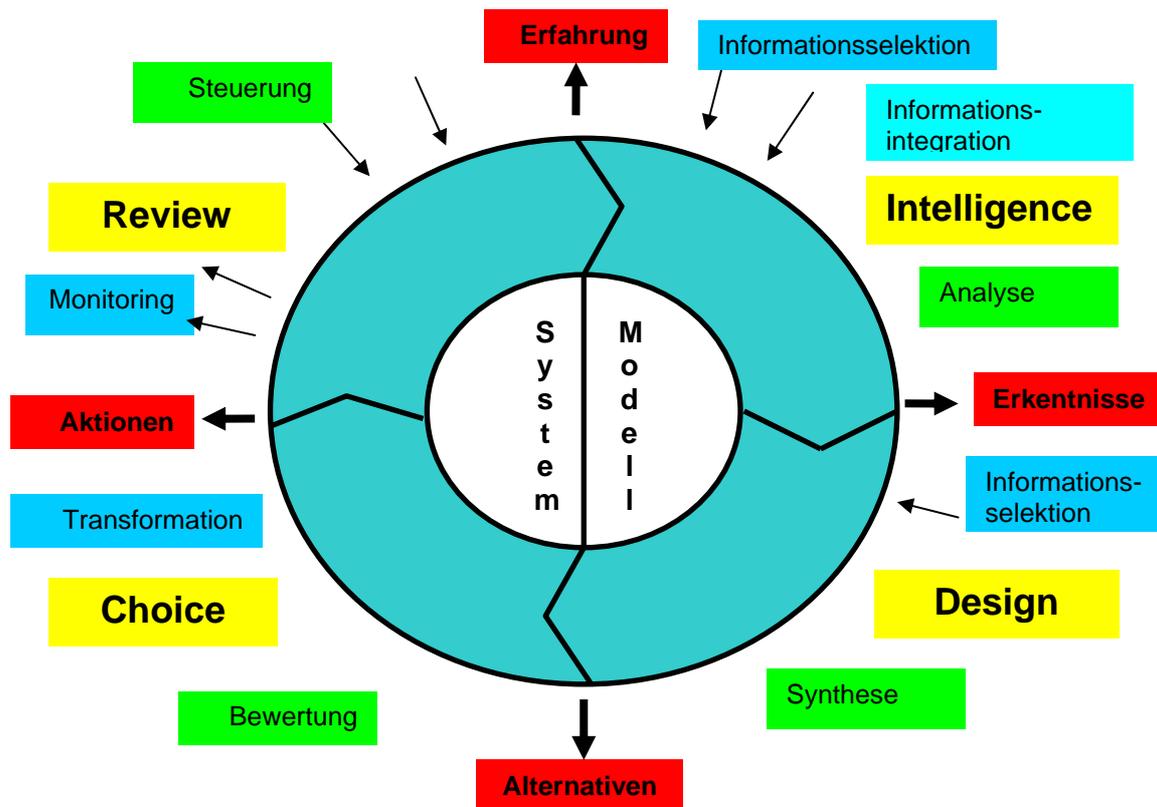


Bild 1: Phasen und Informationsverarbeitung in Entscheidungsprozessen nach Simon

Simon war sich bewusst, dass der Grad der Aufmerksamkeit, die Informationsqualität und Informationsquantität, der zeitliche Rahmen und Streß sowie die kognitiven Fähigkeiten der Entscheidungsträger wie auch deren Zielkonflikte und Handlungsbestreben (Satisfizierung, kontinuierliche Verbesserung, Optimierung) den Entscheidungsprozeß und den angestrebten Zielerreichungsgrad beeinflussen. Simon abstrahiert davon, wie Entscheidungsprozesse durch Zielsetzungen beeinflusst werden und welche Informationen zu beschaffen sind, um rationale, zieladäquate Entscheidungen treffen zu können. Für die Gestaltung von Informationssystemen (IS) und Decision Support-Systemen (DSS) ist ferner entscheidend, in welcher Beziehung der Benutzer zu dem System steht, das in ihm modelliert wird:

- Ist der Benutzer für das System verantwortlich, das im IS / DSS modelliert wird (Unternehmer, Manager, Mitarbeiter)?
- Hat der Benutzer lediglich ein formal begründbares, berechtigtes Informationsinteresse (Stakeholder, Aufsicht, potentieller Kunde/Käufer)?

Das von Nonaka / Takeuchi entwickelte SECI-Modell zur Wissensgenerierung mit den Phasen **S**ocialisation, **E**xternalisation, **C**ombination und **I**nternalisation (SECI) [NoTa95] begründet den Intelligence Cycle:

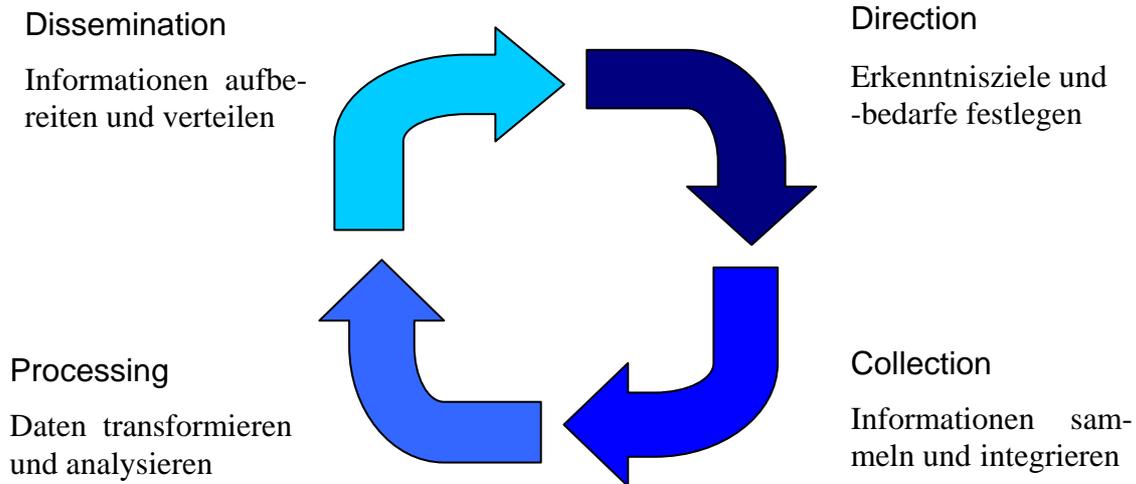


Bild 2: Intelligence Cycle

Erkenntnisfeld und Charakter einer Entscheidungssituation (strategisch / taktisch / operativ) bestimmen die Ausprägung der Phasen und die einzusetzenden Techniken. Strategische Entscheidungssituationen haben häufig einen spieltheoretischen, zuweilen militärischen Charakter und erfordern ein rasches Agieren. Der Militärstrategie John Boyd konzipierte den „Observe – Orient – Decide – Act (OODA) Loop“; er zeigt, wie Individuen oder Organisationen auf Umwelteindrücke reagieren [vgl. Humm08b]:

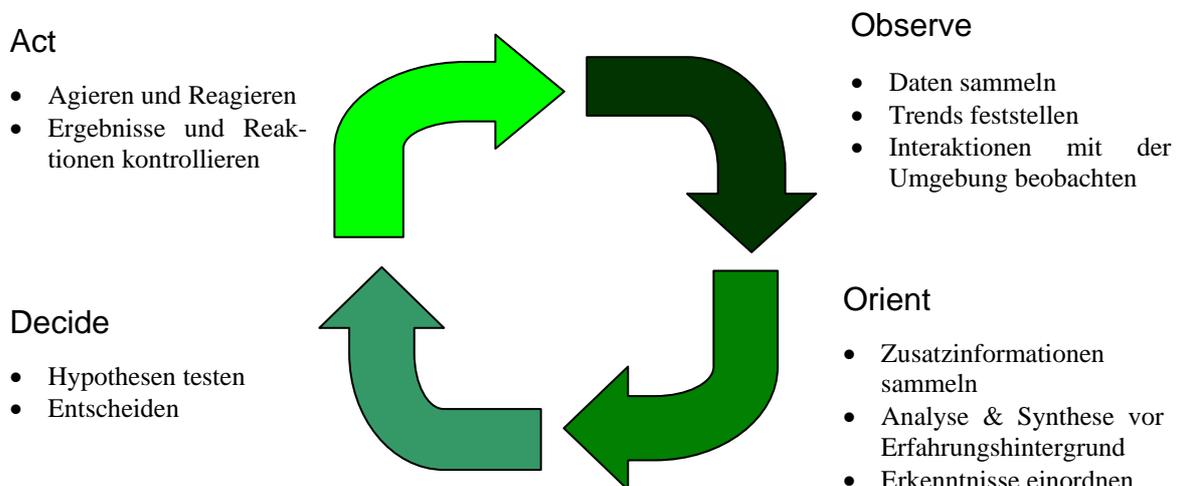


Bild 3: OODA Loop

Taktische Entscheidungssituationen betreffen vor allem Prozessgestaltung und Ressourceneinsatz. Operative Entscheidungen sollen eine optimale, störungsfreie und den Qualitätsanforderungen entsprechende Durchführung der Prozesse garantieren. Ein Konzept für die Umsetzung taktischer und operativer Entscheidungsprozesse liefert das Qualitätsmanagement mit dem „geschachtelten“ Plan – Do – Check – Act (PDCA) Cycle (Bild 4), wobei der innere Kreis die operative und der äußere die taktische Ebene widerspiegelt:

- (I) **Plan:** Erkenne eine Verbesserungsmöglichkeit, plane ihre Umsetzung.
- (II) **Do:** Führe Pilotstudie aus, teste die Auswirkungen der Maßnahme.
- (III) **Check:** Analysiere und bewerte die Testergebnisse.
- (IV) **Act:** Setze die Maßnahme um und bewerte sie.

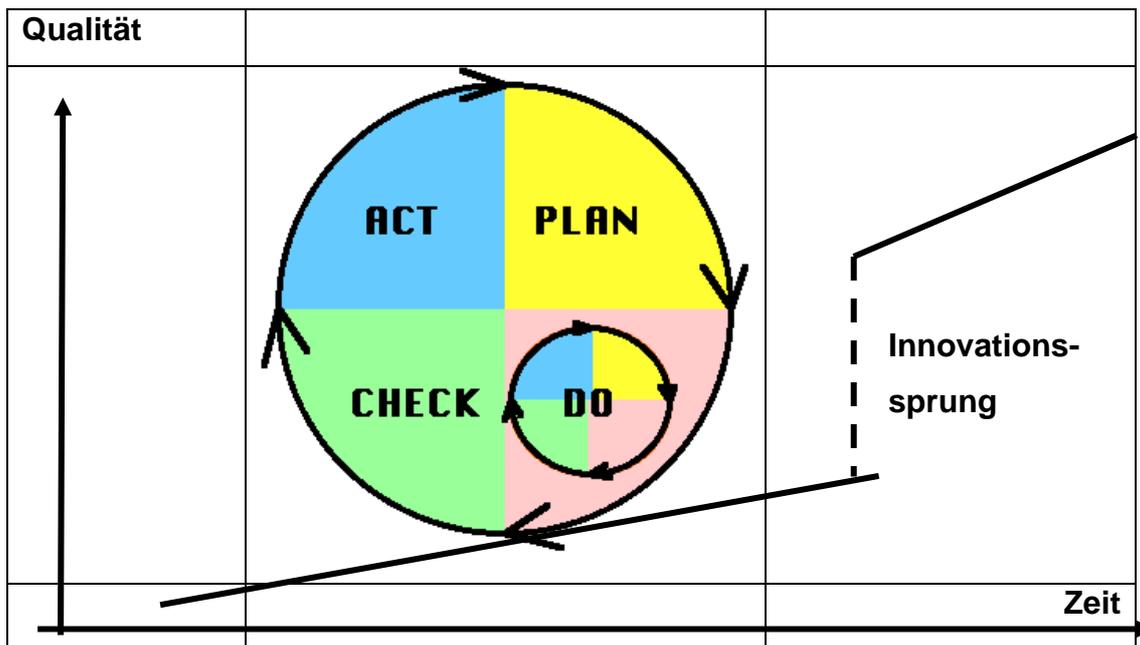


Bild 4: PDCA Cycle [vgl. Humm08b]

Für die BIS-Gestaltung ist entscheidend, welche Arten von Business Knowledge die Phasen des PDCA Cycle erfordern:

Phase	Arten an Business Knowledge	Nötige Business Intelligence
I	Problemlösungswissen	Alternativen, Systemanalyse und -design
II	Transformationswissen (Change Management)	Orientierung über Systemverhalten, Kontextanalyse
III	Systemleistungspotentiale, Bewertungsmethoden	Leistungsmessung, Risikoanalyse und Qualitätskontrolle
IV	Steuerungswissen	Monitoring, Leistungskennzahlen

Tab. 1: Wissensarten und Erkenntnisziele im PDCA Cycle

2.3 Design von Decision Support-Systemen

Decision Support-Systeme schwächen die Vision der (harten) Künstlichen Intelligenz ab, indem sie intelligentes Verhalten durch eine Mensch-Maschine-Interaktion unterstützen und sich dabei auf automatisierbare Prozesse bei Informationsselektion, Analyse und Synthese beschränken. Bei der Analyse wird der Untersuchungsgegenstand als (statisches oder dynamisches) System angesehen, aufgrund dieser Interpretation in seine Bestandteile zerlegt, um diese schließlich zu ordnen, deren Merkmale und Beziehungen zu untersuchen und zu bewerten. Das Gegenstück bildet die Synthese, indem sie versucht, aus den Elementen, die durch die Analyse gefunden wurden, ein neues Ganzes zusammensetzen. Bei komplexen Systemen sollte dies in 2 Schritten geschehen:

- 1) Design: Auf Basis bestimmter Symbole und einer Syntax zu ihrer Verwendung wird eine formale Synthese erstellt.
- 2) Implementierung / Konstruktion: In der materiellen Synthese erfolgt unter Verwendung bestimmter Elemente eine Umsetzung.

Ein Zusatz „OnLine“ kann anzeigen, daß Prozesse schrittweise und unter Benutzerinteraktionen erfolgen.

Die Informatik versucht mit Forschungen zur Künstlichen Intelligenz, intelligentes Verhalten zu automatisieren. Im Rahmen des ESPRIT-Projekts „Knowledge Acquisition and Documentation Structuring“ (KADS) wurde ein Modell für die Entwicklung Wissensbasierter Systeme entwickelt; es umfasst die Phasen: (1) Systems Analysis, (2) Design, (3) Implementation / Use, (4) Maintenance / Knowledge Refinement. Die KADS-Methodologie [vgl. Hick+89] unterscheidet beim Wissensmanagement zwischen den Phasen (1) Systemanalyse, (2) Systemmodifikation und (3) Systemkonstruktion. Die Systemanalyse kann sich auf Kennzahlen oder ein Verhalten (Zeitreihe, „What If“, Impact etc.) beziehen. Anschließend stellen sich folgende Fragen:

- a) Kann das System unverändert weiter betrieben werden?
- b) Sollen Systemeigenschaften (Prozesse, Ressourcen) verändert bzw. verbessert werden (→ Phase 2: Systemadaption, Process Improvement, Kaizen)?
- c) Müssen im System Innovationen eingeführt werden (→ Phase 3: Process Reengineering, Systems Redesign)?

Wird (a) verneint, ist eine neue Systemsynthese erforderlich. Im Fall (b) spricht man von einer Systemmodifikation, im Fall (c) von einer nötigen System(re)konstruktion.

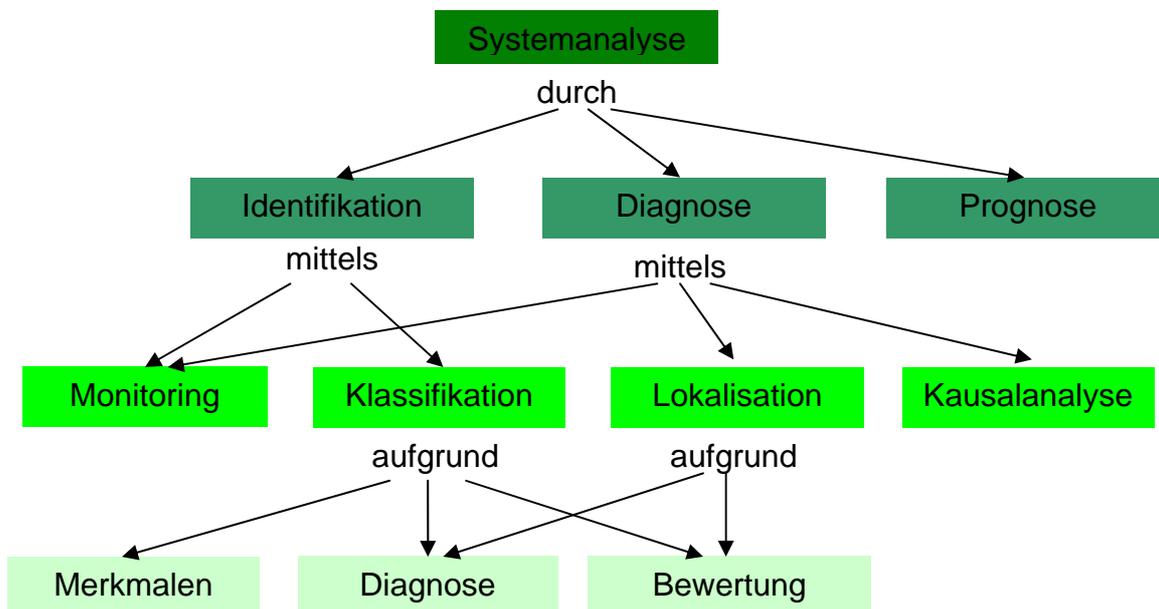


Bild 5: Generische Aufgaben bei der Systemanalyse [Hick+89, 53 u. 92]

In KADS werden Lösungswege für definierte Generic Tasks anhand sog. Interpretationsmodelle in Form von Inferenzstrukturen beschrieben [vgl. Hick+89].

2.4 Spezifische Merkmale von Business Intelligence-Systemen

Business Intelligence als Prozeß umfaßt sämtliche Formen der systematischen Sammlung, Auswertung, Analyse und Darstellung von Informationen. Das Design des Prozesses hängt von den Annahmen über den Erkenntnisgegenstand ab. Der Analyseprozeß endet in der Erkenntnis über das Wesen eines Phänomens oder Gegenstands und dessen innere Zusammenhänge (Struktur, Art der Kausalität etc.). In der Statistik unterscheidet man einerseits 3 Arten der Analyse, und zwar deskriptive, induktive und explorative Statistik, andererseits zwischen univariater und multivariater Analyse. Bei der multivariaten Analyse wiederum lässt sich zwischen strukturprüfenden Verfahren (Regressionsanalyse, Diskriminanzanalyse,...) und strukturentdeckenden Verfahren (Cluster- und Faktorenanalyse, Assoziationsanalyse, ...) unterscheiden. Schließlich sind Annahmen über die Abhängigkeiten zwischen Variablen (linear, nichtlinear, ...) zu treffen.

Bild 6 beschreibt die Phasen von Business Intelligence-Prozessen (BIP), deren Inputs und Outputs und läßt mögliche Iterationen erkennen. OLAP wird in Bild 6, wie in [Codd+93] geprägt, verstanden; über die Sinnhaftigkeit dieses Begriffs wird unten zu diskutieren sein.

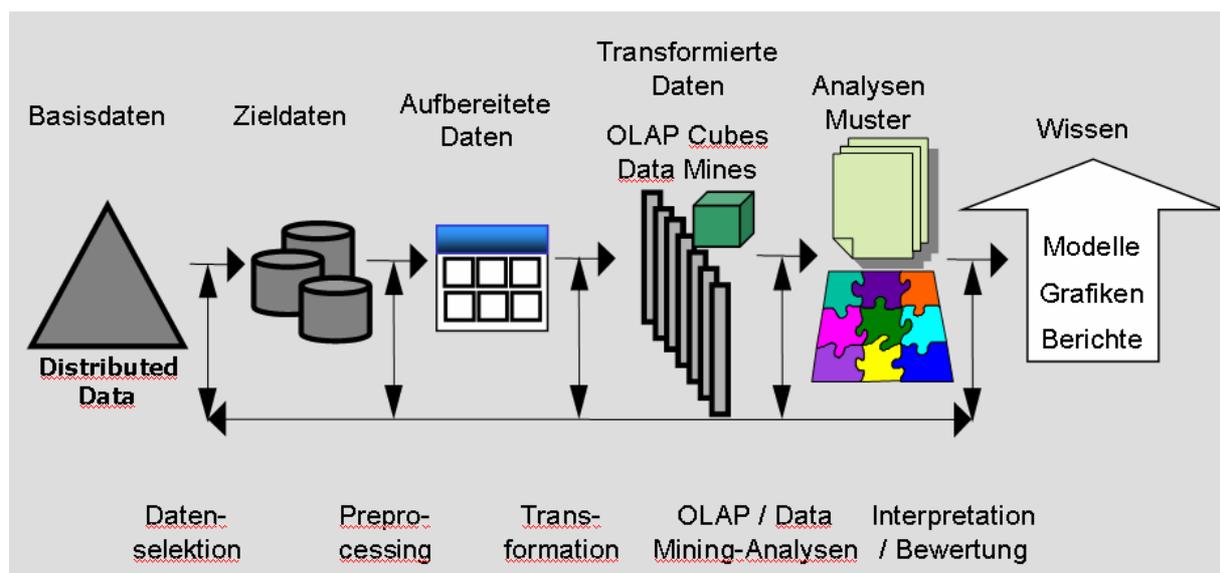


Bild 6: Der Business Intelligence-Prozeß (BIP) [vgl. Humm08a,b]

Business Intelligence als Ergebnis umfaßt in Form von Berichten, Modellen, Kommentaren oder sonstigen, den kognitiven Stilen und Fähigkeiten ihrer Benutzer entsprechenden und ihre Wahrnehmung unterstützenden Darstellungen jene Informationen, die ihnen helfen, das erforderliche Entscheidungswissen zu generieren.

3. Business Intelligence- und Decision Support-Systeme

3.1 Meilensteine

Die Geschichte von Business Intelligence-Systemen (BIS) und Decision Support-Systemen (DSS) ist vielschichtig, technikgetrieben und durch parallele Entwicklungspfade und Spezialisierungen geprägt. Dies verdeutlichen die unterschiedlichen Klassifikationen aus Sicht der „Driver“, „Orientation“ und „Kind of Support“ in Tabelle 2 [vgl. Powe07]. Die Typisierungen basieren vor allem auf Technologiesprüngen. Die durchgängige Unterscheidungen nach Art der Speicherform und Art der Wissensrepräsentation lassen eine Technologieorientierung erkennen.

Klassifikation nach		
Powell	Holsappel/Whinton	Arnott/Pervan
(1) model-driven	(1) text-oriented	(1) personal DSS
(2) data-driven	(2) database-oriented	(2) group DSS / organisational DSS
(3) communications-driven	(3) spreadsheet-oriented	(3) negotiation support system
(4) document-driven	(4) solver-oriented	(4) Intelligent DSS
(5) knowledge-driven	(5) rule-oriented	(5) Knowledge management-based DSS
(6) Web-based		(6) Executive Information System / Business Intelligence
		(7) Data Warehousing

Tab. 2: Typisierungen von Business Intelligence- und Decision Support-Systemen

1958: Content Management mit Text Mining und Profilvergleich

In dem von Luhn entwickelten BI-System in werden Dokumente (Publikationen o.ä.) aufgrund von Arbeitsplatzprofilen selektiert, analysiert und im Hinblick auf ihre Relevanz für bestimmte Mitarbeiter bewertet. Je nach Relevanz wird das gesamte Dokument, ein Abstract oder lediglich ein Hinweis darüber per Monitoring am Arbeitsplatz angezeigt. Klassifikation und Abstract-Erstellung gründen auf statistische Methoden; sie entsprechen einem Text Mining auf Basis von Häufigkeiten und Korrelationen unter Verwendung eines Thesaurus und Taxonomien. Die Stellenbeschreibungen, aufgrund derer Arbeitsplatzprofile erstellt werden, sind semi-strukturiert. Dem Umstand, dass Dokumentenmanagementsysteme in den 80-er Jahren als Groupware vor allem für Office-Anwendungen auf Microrechnern entwickelt wurden, dürfte zuzuschreiben sein, dass der BI-Ansatz von Luhn nicht in das Content Management (CM) Eingang fand.

1978: Finanzplanungssoftware / Planungssprachen

Das Interactive Financial Planning System (IFPS / Execucom) bietet als erstes eine Sprachumgebung, welche die Modellierung und Analyse einer mehrperiodischen Finanzplanung unterstützt. Die Sprache ist den Begrifflichkeiten und Rechenschritten in der Finanzplanung

angepasst, enthält Makros für finanzmathematische und statistische Methoden und unterstützt die Berichtserstellung und die Strukturierung einer hierarchischen Konsolidierung. Einfache Input-Output-Berechnungen können um Sensitivitätsanalysen („What If“, Impact-Analyse, parametrische Analysen) sowie Zielerreichungsberechnungen (Target Analyse, Goal Seeking) ergänzt werden. Der Vorteil gegenüber den damals vorherrschenden hierarchischen Datenbankmanagementsystemen lag in der tabellenorientierten 2- bis 3 dimensionalen Sicht sowie der Verwaltung und Manipulation von Zeitreihendaten inkl. Sensitivitätsanalysen.

1979: Spreadsheet-System VisiCalc

Mit VisiCalc (Visible Calculator) wird von der Firma Personal Software (später VisiCorp) das erste Spreadsheet-System für Apple II Personal Computer vorgestellt. Es folgen die Spreadsheetsysteme SuperCalc, Multiplan, Lotus 1-2-3. Mit Microsoft (Office) Excel trat 1993 das erste System auf den Markt, das dem Benutzer erlaubte, das Erscheinungsbild des Spreadsheets zu beeinflussen, und das analysierte, welche Zelleninhalte sich bei einer Neuberechnung ändern können, um dadurch die Performance zu steigern.

1988: Business Data Warehouse and Executive Information System

Nachdem 1986 Stecher und Hellemaa „An 'intelligent' extraction and aggregation tool for company data bases“ vorgestellt hatten [StHe86], erlaubten die Fortschritte in der relationalen Datenbanktechnologie 1988 Devlin/Murphy [DeMu88], ein integriertes System für Informationssuche und Berichtserstellung per Ad hoc-Anfragen vorzustellen. Da es nur einen beschränkten Analyse- und keinen Synthesupport liefert, sprechen sie von einem „Office Support System“. Die Informationen werden in einem Business Data Warehouse (BDW), welches Basis für ein „Business and information system“ ist, entsprechend den Geschäftserfordernissen gespeichert. Der Begriff „Data Warehouse“ wird ab 1990 von Inmon mit seinem Buch „Building the Data Warehouse“ [Inmo90] popularisiert.

Ebenfalls im Jahr 1988, jedoch unabhängig von den Entwicklungen im Data Warehousing, entscheidet sich die Autovermietung Hertz, ein Executive Information System (EIS) zu entwickeln. Es ist PC-basiert und wird mir dem Commander EIS von Comshare erstellt. Besonderes Merkmal ist ein Exception Reporting mittels Dashboard. Die Verwendung des Begriffs „Ampelfunktion“ wird hier vermieden, da bei alleiniger Verwendung der Ampelfarben „grün“, „gelb“, „rot“ keine Einbindung in einen Managementzyklus mit der begleitenden Phasenfolge „grün“ → „gelb“ → „rot“ → „rot-gelb“ → „grün“ markiert werden kann.

1989: Popularisierung von Business Intelligence

Es diente der Popularisierung von Business Intelligence-Systemen (BIS), daß Howard Dresner im Jahre 1989 diesen Begriff verwendete, um auf die Möglichkeiten der Beschleunigung und Verbesserung der Informationsversorgung durch „fact-based support systems“ auf Basis von Datenzentren (Data Warehouse / Data Marts) hinzuweisen. Die Tatsache, daß er in diesem Zusammenhang nicht auf die 31 Jahre vorher von Luhn publizierte Arbeit hinwies, mag den Eindruck erweckt haben, es handele sich um eine neue Klasse von Informationssystemen. Dies hat zu Parallelentwicklungen im Data Warehousing (DW), bei Business Intelligence-Systemen (BIS) und Content Management-Systemen (CMS) geführt.

Mit der Popularisierung von BIS war eine Abkehr von der umfassenden Definition von Luhn verbunden. Die Gartner Group fokussierte BIS auf Analysewerkzeuge und Berichtserstellung. Diese Einschränkung und Abgrenzung von BIS gegenüber einer zentralen Datenhaltung im Data Warehousing mit vorangehendem ETL-Prozeß verkennen, daß die Güte von Informatio-

nen von der Datenqualität und der Informationsnutzen maßgeblich vom Umfang der Datenintegration abhängen. Die Abgrenzung schlägt sich in den unterschiedlichen Magic Quadrants von Gartner für Data Warehouse DBMS und BI Tools nieder.

1992: Balanced Scorecard

Ein Armaturenbrett (Cockpit, engl. Dashboard) zeigt Meßgrößen über unterschiedliche Instrumente an, um Zustand und Verhalten einer Maschine zu beschreiben. Folgerichtig schlagen 1992 Kaplan und Norton [KaNo92] die Verwendung einer Scorecard zur Bewertung der Unternehmensleistung vor. Sie dient der Identifikation der Ausprägung von Kennzahlen mit anschließender Kausalanalyse. Als Orientierungsrahmen, um Ursache-Wirkungsanalysen zwischen zentralen Leistungsindikatoren (Key Performance Indicator) durchzuführen und Maßnahmen zu empfehlen, schlagen Kaplan und Norton die Verwendung einer Strategy Map vor. Sie geht auf die bereits in den 50er Jahren von Jay W. Forrester im Rahmen von System Dynamics entwickelten Methodik zur ganzheitlichen Analyse und Synthese von Systemen zurück [Forr77]. Balanced Scorecards und Strategy Maps werden heute irrtümlich auf allen Organisationsebenen für verschiedene Arten von Kennzahlensystemen und Entscheidungsprozessen in Form von Kennzahlen-Cockpits eingesetzt. Dabei wird häufig versäumt, die Balanced Scorecard durch eine dem betrachteten System adäquate Modellbeschreibung (Business Process Notation) und Modellanalyse (Business Process Animation / Simulation) zu ersetzen.

1993: Cube Analysis OnLine Analytical Processing (OLAP)

Im Jahre 1993 proklamieren E.F.Codd et al. mit „Providing OLAP to User-Analysts: An IT Mandate“ [Codd+93] die Notwendigkeit,

- (1) Daten aus unterschiedlichen Quellsystemen zu selektieren und zu integrieren,
- (2) sie in multidimensionale Datenmodelle zu überführen und
- (3) unterschiedliche Formen der statischen und dynamischen Datenanalyse auszuführen.

Die folgende Würdigung ihres Ansatzes konzentriert sich auf seine konzeptionellen Aussagen und Designvorschläge.

Codd et al. teilen Modelle zur Datenanalyse in 4 Typen ein (Numerierung hinzugefügt):

„(1) the categorical model; (2) the exegetical model; (3) the contemplative model; and (4) the formulaic model. The construction and manipulation of each of these four model types is different. These data models and their use follows a continuum from actual historic data to future anticipated data values and behaviours; the requirement for analysis of the data across multiple dimensions, the degree of analyst interaction and involvement required to construct and manipulate each of these models increases along this same axis.”

Hierunter fallen multidimensionale identifizierende Datenanalysen, aber auch statistische und prediktive Analysen wie Data Mining zur Mustererkennung (Pattern Recognition). Modelltyp (1) lässt sich häufig offline analysieren; die Modelltypen (2) – (4) erfordern eine zunehmende Benutzerinteraktion (vgl. Bild 7). Codd et al. definieren deshalb “OnLine Analytical Processing” (OLAP) als „category of database processing“ mit folgenden Funktionalitäten:

„OLAP is the name given to the dynamic enterprise analysis required to create, manipulate, animate, and synthesize information from exegetical, contemplative, and formulaic data analysis models. This includes the ability to discern new or unanticipated relation-

ships between variables, the ability to identify the parameters necessary to handle large amounts of data, to create an unlimited number of decisions (consolidation paths), and to specify cross-dimensional conditions and expressions.”

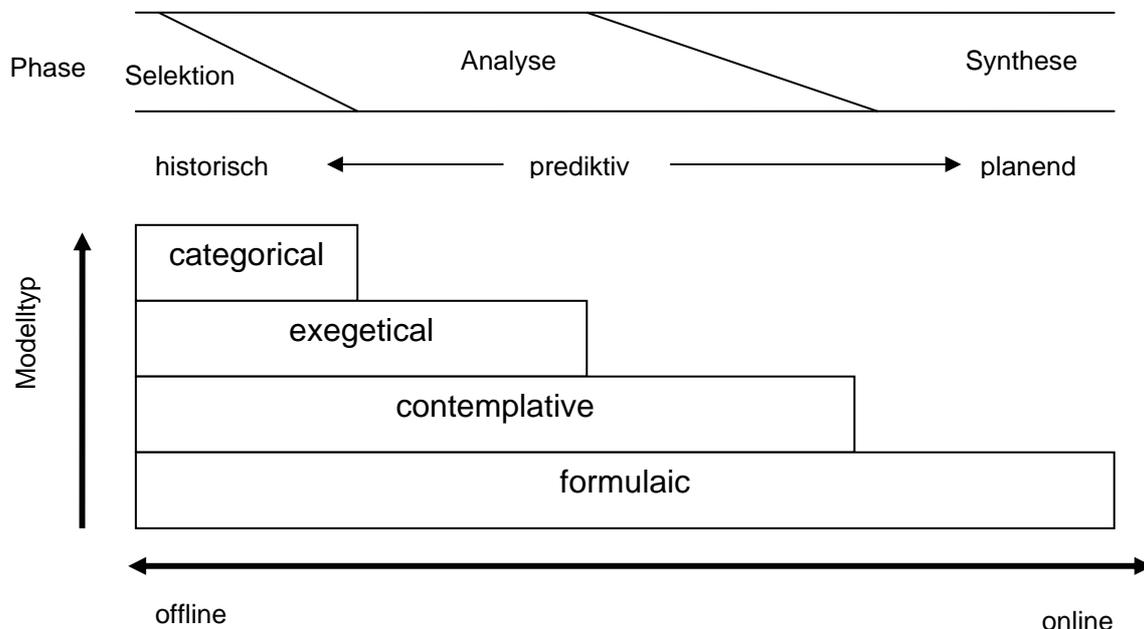


Bild 7: Anforderungen der Modelltypen an das Information Processing

Die Gründe für den Vorschlag eines eigenständigen OLAP-Servers bilden vor allem Performanceüberlegungen und die Notwendigkeit von Benutzerinteraktionen bei der Analyse umfangreicher Datenmengen und komplexer Phänomene. Bei relationalen Datenbanken bedeutet eine stringente Befolgung der 3. Normalform (3NF) die Modellierung im Snowflake Scheme. Aggregationen über Konsolidierungspfade erfordern dann rekursive Joins der Lookup-Tabellen. Schließt man „Materialized Views“ aus, so wird verständlich, dass aus Performancegründen ein separater Server vorgeschlagen wird. Dieser dient der Datenintegration aus unterschiedlichen Quellsystemen und ist als offenes System konzipiert, um

- a) auf ihm hierarchisch strukturierte Analysen (Drill Down / Roll Up, Slice & Dice, Rotate etc.) durchzuführen und
- b) Datenschnittstellen für Tabellenkalkulation, grafische Benutzerschnittstellen und statistische Analysesysteme bereit zu stellen.

Der von Codd et al. geprägte Begriff „Online Analytical Processing“ ist historisch zu begründen. Er beschreibt spezifische Formen der Mensch-Maschine-Interaktion in Informationssystemen. Er ist jedoch inhaltsleer in Bezug auf das zur Identifikation von Wissen verwendete Modell „Multidimensional Analysis and Reporting“.

1996: From Data Mining to Knowledge Discovery in Databases (KDD)

Die Konzentration auf die Multidimensionalität von OLAP-Systemen hat dazu geführt, dass OLAP häufig als multidimensionale Datenanalyse („Cube Analysis“) angesehen, also auf den Generic Task Identifikation reduziert wird. Dies erklärt die begriffliche Unterscheidung zwischen OLAP und Data Mining und, warum 1996 Fayyad et al. in [FaPS96] unabhängig vom

ETL-Prozeß für OLAP-Server und die darauf aufbauenden Cube-Analysen einen KDD-Prozeß für die Generierung von Wissen aus umfangreichen Datenbeständen mittels Data Mining kreiert haben. Sie näherten sich Data Mining aus Sicht der Künstlichen Intelligenz, dem maschinellen Lernen, und begründeten ihren KDD-Ansatz wie folgt:

„The traditional method of turning data into knowledge relies on manual analysis and interpretation. ... Be it in science, marketing, finance, health care, retail or any other field, the classical approach to data analysis relies fundamentally on one or more analysts becoming intimately familiar with the data and serving as an interface between the data and the users and products.”

Ebenso wie Codd et al. weisen Fayyad et al. auf die Notwendigkeit eines OnLine Processing hin, ohne es jedoch begrifflich - etwa im Sinne „OnLine Data Mining“ - herauszustreichen.

„Blind application of data-mining methods (rightly criticized as data dredging in the statistical literature) can be a dangerous activity, easily leading to the discovery of meaningless and invalid patterns.”

Um aus Data Mining-Analysen synthetische Urteile abzuleiten, wählte bereits Luhn einen regelbasierten Ansatz [s.a. Till03].

3.2 State of the Art

Business Intelligence-Systeme (BIS) gelangen heute in allen Funktionsbereichen und auf sämtlichen Ebenen von Organisationen sowie in deren Beziehungen zu Kunden, Lieferanten, Stakeholdern usw. zum Einsatz. Ihre Erscheinungsformen sind individuell und derart komplex, dass zu ihrer Differenzierung zwischen verschiedenen Disziplinen von Business Intelligence unterschieden werden sollte [vgl. Humm08a]. Entsprechend vielfältig und vielschichtig ist der Kreis der BIS-Anwender. Grob gesehen, lassen sich 5 Benutzergruppen unterscheiden. Je anspruchsvoller die Analyse ist, um so stärker sind die Benutzerinteraktionen. In Bild 8 (nächste Seite) wird deshalb die Bezeichnung OLAP für gewisse Formen der Analyse vermieden und „OnLine Analytical Processing“ durch „MultiDimensional Analytical Processing“ (MDAP) ersetzt. Unter Advanced Analysis werden neben statistischen Analysen ebenso Risikoanalysen subsumiert.

Die Vielfalt der in Bild 8 angeführten Benutzer legt nahe, nicht nur in einem Data Warehouse die von ihnen aus unterschiedlichen Quellsystemen bezogenen Daten zu integrieren, sondern auch die organisationale Intelligenz durch Unterstützung der kollaborativen Intelligenz [vgl. Humm08a] zu steigern. Bereits zu Beginn der 90er Jahre wurde der Begriff der „Organizational Decision Support Systems“ (ODSS) [vgl. TuAr98, 420ff.] geprägt, um die Kommunikation zwischen sowie die Koordination und Kooperation von Mitarbeitern getrennter Bereiche und auf unterschiedlichen Ebenen zu steigern. Die damaligen Versuche waren recht schwerfällig; kollaborative Intelligenz wird heute auf Basis der Web-Technologie mit Web-Servern und Portalen erfolgreich umgesetzt.

Bereits Devlin/Murphy sahen 1988 in ihrer Architektur für die „Distributed Data Storages“ in den Workstations neben dem aus dem Business Warehouse abgeleiteten „Public Storage“ einen „Personal Storage“ vor. Sie berücksichtigten so, dass zentralistische Ansprüche einer IT-Abteilung, das gesamte Wissen einer Organisation verwalten zu wollen, zum Scheitern von BI-Projekten führen.

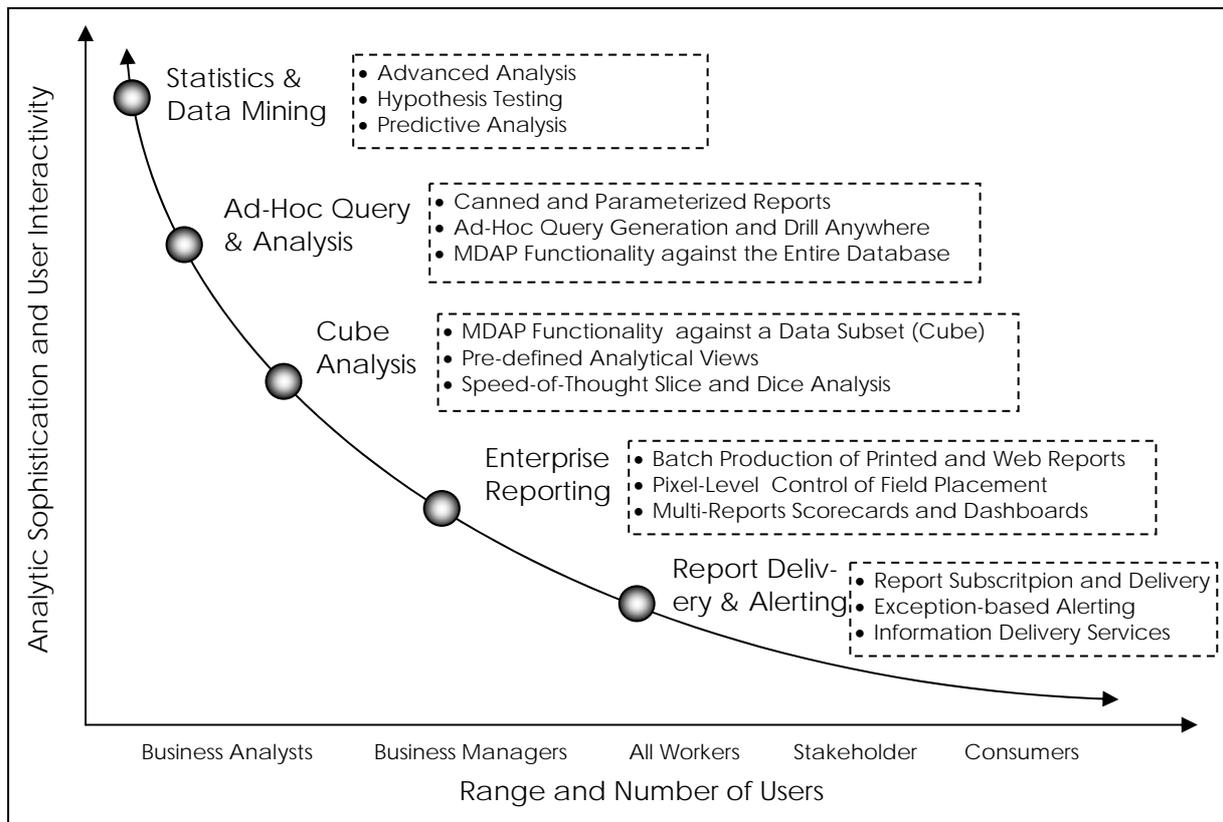


Bild 8: Benutzergruppen und BI-Analysetechniken (ohne Content Management) [vgl. Micr02]

Die im Markt der Office-Systeme führende Position von Microsoft erklärt, weshalb MS Access und MS Excel in Verbindung mit MS Word für Textverarbeitung und Dokumentenverwaltung/-austausch am Action Point die vorherrschenden Personal Storages bilden. MS Office liefert eine durchgängige Unterstützung sämtlicher Phasen des Intelligence Cycle. Ein Benutzer kann mit der MS Office Suite Daten unterschiedlicher Formate und Strukturen per ODBC, JDBC, OCR, XML/SQL-Converter oder sonstiger Import-Funktionen in lokale oder zentrale Datenbank-Server importieren oder sogleich in spezifische DSS (Excel etc.) importieren, um sie mittels data-driven Komponenten weiter zu verarbeiten. Die Analysefunktionen unter MS Excel ermöglichen Cube-Analysen unter Einsatz des Pivot-Tabellen-Berichtsgenerators, statistische Analysen und deren graphische Visualisierung etc. Cube-Import und weitere Access Services erlauben eine direkt Kopplung von Excel mit der BI-Plattform und deren Services. MS Excel unterstützt aber auch die sich an die Analyse anschließende Phasen von Synthese und Auswahl durch

- a) Planungsfunktionalitäten im Rahmen einer tabellenorientierte Modellbildung bis hin zur Mathematischen Optimierung sowie
- b) Instrumente für Risikoanalyse und Risikosimulation [vgl. Humm07].

Ein sog. Power User kann also seine Möglichkeiten der Wissensgenerierung und des Lernens, seine Effektivität und seine Produktivität in dem von ihm gewünschten und erreichbaren Maße steigern. Aufgabe zentraler IT-Services ist, derartige Funktionalitäten organisationsweit zur Verfügung zu stellen, zu integrieren und zu unterstützen, Qualitätsstandards zu setzen und die Performance zu erhöhen. Diese Koordinationsaufgabe stellt neue Anforderungen an die IT Governance und lässt sich z.B. mittels BI Competency Centern (BICC) lösen.

4. Schlußbemerkungen

Die Ansätze von Luhn, Devlin/Murphy und Codd et al. markieren die tragenden Säulen in der Entwicklungsgeschichte von Business Intelligence-Systemen:

- (1) Luhn: Business Intelligence für die betrieblichen Prozesse und die an ihnen beteiligten Action Points mittels Data Mining und wissensbasierten Systemen
- (2) Codd et al.: Business Intelligence über die betrieblichen Prozesse mittels multidimensionaler Datenanalyse, ergänzenden statistischen Analysen und unterstützenden Synthesewerkzeugen (Strategy Map, statische Planung mittels Spreadsheet-Systemen, dynamische Planung mittels Simulation und Risikoanalyse)
- (3) Devlin/Murphy: Data Warehouse / Data Marts als Data Center and Management Information Bases im Business Intelligence and Decision Support Process.

BIS bilden einen zentralen Gegenstand der Wirtschaftsinformatik. Ihre Darstellungen werden in ihr sowohl an akademischer Strenge (Rigorosität) als auch an der Praxisrelevanz gemessen. Wie die Erfahrung lehrt, ist es häufig schwierig, beide Erwartungen gleichermaßen zu erfüllen. Offensichtlich dominiert bei BIS / DSS die Praxis die Theorie. Unter den Gesichtspunkten der Rigorosität sind folgende Punkte zu kritisieren:

1. Mit der Trennung zwischen Data Warehousing (DW) und Business Intelligence-Systemen (BIS) werden ETL-Prozesse auch konzeptionell auf die Informationslogistik für eine zentrale Datenhaltung im DW und eine entscheidungsbezogene Informationsselektion in Office BIS am Action Point aufgeteilt.
2. Die Trennung zwischen OLAP und Data Mining erweckt den Eindruck, dass eine Mustererkennung keine Online-Analyseprozesse erfordert. Online als klassifizierendes Merkmal bei Cube Analyses, nicht jedoch bei Data Mining, ist daher fragwürdig.
3. Die Überbetonung der Notwendigkeit, eine wachsende Informationsflut zu bewältigen, hat in „fact-based“ DSS das „Systems Thinking“ verdrängt; es verharrt bei BIS häufig auf der Ebene einer „statischen“ strategy map.
4. Das Ziel der Wissensgenerierung, -konservierung und ständigen Aktualisierung (Lernen) wird in BIS vernachlässigt, indem die Erkenntnisse aus der Phase Review nicht ausgewertet, vor dem Hintergrund der bisher gesammelten Erfahrungen bewertet und archiviert werden. Ohne ein Organizational Memory wird organisationales Lernen durch BIS unvollkommen unterstützt.
5. Zwar wird für BIS ein „business driven approach“ propagiert, aber häufig eine „IT-driven solution“ implementiert. Entsprechend schwach ist dann das Business Modelling ausgeprägt.
6. Information Warehouse / Intelligence Marts müssen nicht „subject-oriented“, sondern „system and business process oriented“ entwickelt und angepaßt werden.
7. Bei der BIS-Entwicklung ist zwischen personal, group and organizational decision support zu differenzieren.

5. Literatur

[ACKO67] ACKOFF, R.L.

Management Misinformation Systems. *Management Science* 14(1967)4, D147-B156.

[CHGL06] CHAMONI, P.; GLUCHOWSKI, P. (HRSG.)

Analytische Informationssysteme: Business Intelligence-Technologien und – Anwendungen. 3. Aufl., Berlin/Heidelberg/New York 2006.

[CODD+93] CODD, E.F., ET AL.

Providing OLAP to User-Analysts: An IT Mandate. Codd & Associates, White Paper, o.O. 1993

[DEMU88] DEVLIN, B., MURPHY, P.T.

An Architecture for a Business and Information System. *IBM Systems Journal* 27(1988)1, S. 60-80.

[FAPS96] FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P.

From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, Fall 1996, S. 37-54

[FORR77] FORRESTER, J.W.

Industrial dynamics, 9. Auflage, Cambridge 1977

[GLKE06] GLUCHOWSKI, P.; KEMPER, H.-G.

Quo Vadis Business Intelligence? *BI-Spektrum* 01-2006, S. 12-19.

[HICK+89] HICKMAN, F.R.; KILLIN, J.L.; LAND, L.; MULHALL, T.; PORTER, D.; TAYLOR, R.M.

Analysis for Knowledge-Based Systems - A Practical Guide to the KADS Methodology. New York/... 1989.

[HUMM07] HUMMELTENBERG, W.

Risikosimulation. In: Freidank, C.-Chr.; Lachnit, L.; Tesch, J. (Hrsg.): *Vahlens Großes Auditing Lexikon*, München 2007, S. 1189-1194.

[HUMM08A] HUMMELTENBERG, W.

Disziplinen von Business Intelligence. In: von Kortzfleisch, H.F.O.; Bohl, O. (Hrsg.): *Wissen, Vernetzung, Virtualisierung*, Lohmar-Köln 2008, S. 41-56.

[HUMM08B] HUMMELTENBERG, W.

Business Intelligence. In Kurbel, K. et al. (Hrsg.): *Enzyklopädie der Wirtschaftsinformatik - Online Lexikon*, <http://www.oldenbourg.de:8080/wi-encyklopaedie/lexikon/daten-wissen/Business-Intelligence>.

[INMO90] INNMON, W.H.

Building the Data Warehouse. New York/... 1990.

[KAN092] KAPLAN, R.S.; NORTON, D.P.

The Balanced Scorecard - Measures That Drive Performance. Harvard Business Review 1992, 71-79

[KEBA06] KEMPER, H.-G.; BAARS, H.

Business Intelligence und Competitive Intelligence. HMD 247(2006), S. 7-20.

[LUHN58] LUHN, H.P.

A Business Intelligence System. IBM Journal, Oct. 1958, S. 314-319

[MICR02] VGL. MICROSTRATEGY (HRSG.):

The 5 Styles of Business Intelligence : Industrial-Strength Business Intelligence. McLean (VA) 2002. <http://www.microstrategy.com/Solutions/5StylesBook.asp>.

[NOTA95] NONAKA, I.; TAKEUCHI, H.

The Knowledge Creating Company. Oxford / New York 1995.

[POWE07] POWER, D.J.

A Brief History of Decision Support Systems. DSSResources.COM, <http://DSSResources.COM/history/dsshistory.html>, version 4.0, March 10, 2007.

[SIMO47] SIMON, H.A.

Administrative Behavior. New York, N.Y., 1947.

[STHE86] STECHER, P.; HELLEMAA, P.

An ,intelligent' extraction and aggregation tool for company databases. Decision Support Systems 2(1986)2, 145-158.

[TILL03] TILLMANN, CHR.

Data Mining zur Unterstützung betrieblicher Entscheidungsprozesse. Diss. Univ. Dortmund 2003.

[TUAR98] TURBAN, E.; ARONSON, J.E.

Decision support systems and intelligent systems. 5th Ed., Upper Saddle River, N.J., 1998.