Information System: (3-0-2)

Course Objectives

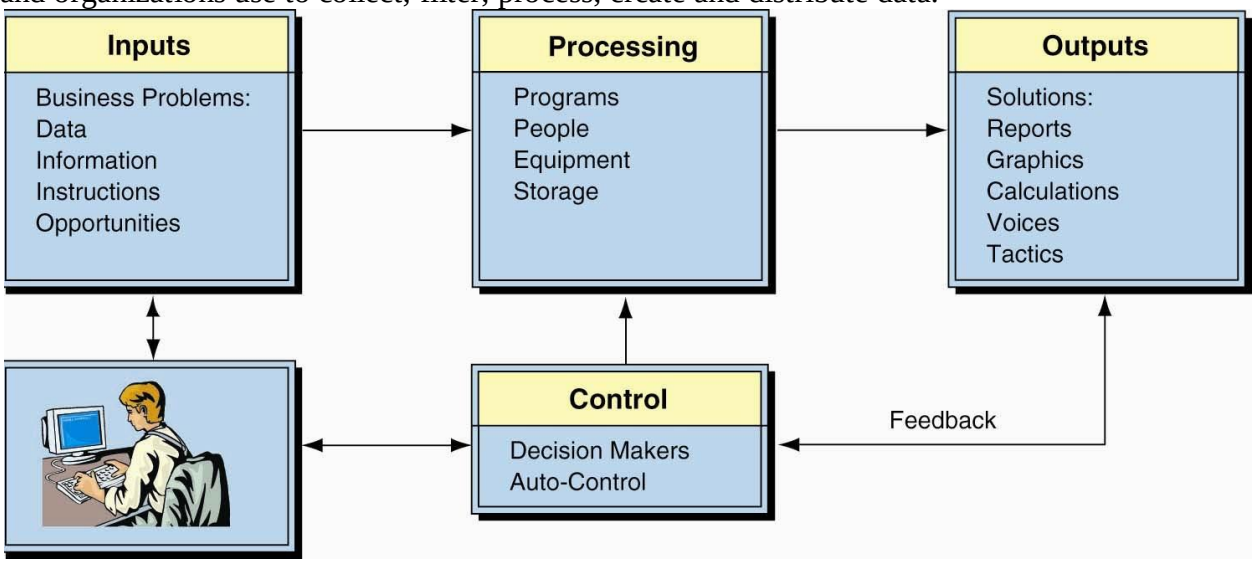
• To introduce & apply the knowledge of computer based IS
• To provide the concept in designing & setting up complex information system.

Chapter 1: Information System

– Classification & Evolution of IS
– IS in functional area
– Information System Architecture
– Qualities of Information systems
– Managing information system resources
– Balanced scorecard- case studies

1.1 Classification & Evolution of IS:

• **Information system (IS)** collects, processes, stores, analyzes, and disseminates information for a specific purpose.
• Includes ***inputs*** (data, instructions) and ***outputs*** (reports, calculations).
• ***Processes*** the inputs by using technology such as PCs and produces outputs that are sent to users or to other systems via electronic networks and a ***feedback*** mechanism that ***controls*** the operation.
Information system (**IS**) is the study of complementary networks of hardware and software that people and organizations use to collect, filter, process, create and distribute data.



• **Data Item.** Elementary description of things, events, activities and transactions that are recorded, classified and stored but are not organized to convey any specific meaning.
• **Information.** Data organized so that they have meaning and value to the recipient.
• **Knowledge.** Data and/or information organized and processed to convey understanding, experience, accumulated learning and expertise as they apply to a current problem or activity.
• **Information Technology Architecture.** A high-level map or plan of the information assets in an organization, which guides current operations and is a blueprint for future directions.
• **Information Technology Infrastructure.** The physical facilities, IT components, IT services and IT management that support an entire organization.
• **Computer-based Information System (CBIS)** An information system that uses computer technology to perform some or all of its intended tasks.

Basic Components of Information Systems

• Hardware
• Software
• Database
• Network
• Procedures
• People

Why Study Information Systems?

• You will be more effective in your chosen career if you understand how successful information systems are built, used, and managed.
• You also will be more effective if you know how to recognize and avoid unsuccessful systems and failures.
• According to the US Bureau of Labor Statistics, "Top seven fastest growing occupations fall within IT or computer related field"
• Developing "Computer" Literacy will only enhance your "Information" Literacy

Evolution of IS

• First business application of computers (in the mid- 1950s) performed repetitive, high-volume, transaction-computing tasks.
• The computers" crunched numbers" summarizing and organizing transactions and data in the accounting, finance, and human resources areas - called transaction processing systems (TPSs)
• Management Information Systems (MISs): these systems access, organize, summarize and display information for supporting routine decision making in the functional areas.
• Office Automation Systems(OASs): such as word processing systems were developed to support office and clerical workers.
• Decision Support Systems: were developed to provide computer based support for complex, Non routine decision.
• End- user computing: The use or development of information systems by the principal users of the systems' outputs, such as analysts, managers, and other professionals.
• Intelligent Support System (ISSs): Include expert systems which provide the stored knowledge of experts to non experts, and a new type of intelligent system with machine- learning capabilities that can learn from historical cases.
• Knowledge Management Systems: Support the creating, gathering, organizing, integrating and disseminating of organizational knowledge.
• Data Warehousing: A data warehouse is a database designed to support DSS, ESS and other analytical and end-user activities.
• Mobile Computing: Information systems that support employees who are working with customers or business partners outside the physical boundaries of their company; can be done over wire or wireless networks.
• eCommerce – Need to allow access to customers
• and many more……

Classification of Information Systems

These include classification by
• Organizational structure,
• Functional area within the organization,
• Use across multiple organizations,
• Mode of Data Processing
• Classification By System Objectives

Classification by Organization structure

• Organizations -made up of components such as divisions, departments, and work units,

• Organized in hierarchical levels
• For example,
– organizations have functional departments, such as production and accounting,
– which report to plant management,
– which report to a division head.
– The divisions report to the corporate Headquarters
– Although some organizations have restructured themselves in innovative ways, such as those based on cross-functional teams,
• today the vast majority of organizations still have a traditional hierarchical structure.
• Thus, we can find information systems built for headquarters, for divisions, for the functional departments, for operating units, and even for individual employees.
• Such systems can stand alone, but usually they are interconnected.
• Typical information systems that follow the organizational structure are ***functional (departmental), enterprise, and interorganizational.***
• These systems are organized in a hierarchy in which each higher-level system consists of several (even many) systems from the level below it.
• At a higher level, the enterprise system supports the entire company, and inter-organizational systems connect different companies.

An information system (IS) can span departments, business units and corporations .
• Departmental IS
• Enterprise-Wide IS
• Inter-Organizational IS

Classification by Function
• Functional organizations
– are hierarchical structures and
– center on a strong concept of supervisors and subordinates
• The controlling authority, often called top management, coordinates with each management level and functional department to keep the organization running smoothly
• A functional organization analyzes the strengths and weaknesses of each member,
• groups them into categories and assigns them to tasks that best utilize their skills
• Jobs that perform a similar function are grouped in functional areas
• Each functional area contains employees with varied skills grouped based on specialization and put in separate units or departments
• Information systems which served these functional departments are called functional information systems
An information system (IS) support each department in a corporation.
• Operations
• Accounting
• Finance
• Marketing
• Human resources

Evolution of Functional Organizations
• Functional organizations work best when a single product or service is involved
• The chain of command is linear, so everyone knows his position in the organization
• By clustering specialists with similar skills, leadership, tutoring and guidance concentrate on one area
• Employees have an obvious path for growth and promotion, either up or lateral
• As a company gets larger, some of the positives of functional organizations become negatives

• Since decisions travel through the chain of command, the process becomes bureaucratic, and information and decisions move slowly
• Functional grouping can result in a narrowed overall perspective
• Because of communication and decision making issues, the functional organization is slow to adapt to environmental changes

Enterprise Information Systems

• While a departmental information system is usually related to a functional area, other information systems serve several departments or the entire enterprise
• These information systems together with the departmental applications comprise the enterprise information system (EIS).
• most popular enterprise applications is enterprise resources planning (ERP),
– which enables companies to plan and manage the resources of an entire enterprise.

Inter-organizational Information Systems

• Some information systems connect two or more organizations-referred as inter-organizational information systems (IOS's).
• IOS's support many inter-organizational operations, of which supply chain management is the best known
• An organization's supply chain describes the flow of materials, information, money, and services from raw material suppliers through factories and warehouses to the end customers.
• supply chain includes both physical flows and information flows.
• Information flows and digitisable products (e.g., music and software) go through the Internet, whereas physical products are shipped.
• For example, when you order a computer from *www.dell.com, your information goes to Dell via the Internet.*
• When your transaction is complete (i.e., your credit card is approved and your order is processed), Dell ships your computer to you.

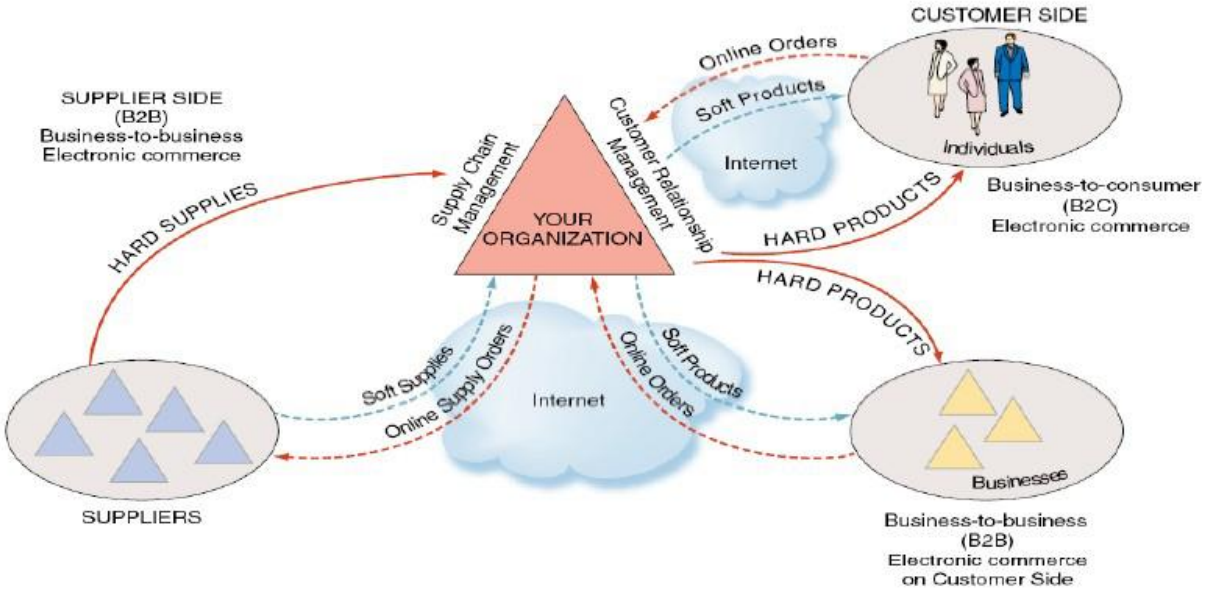
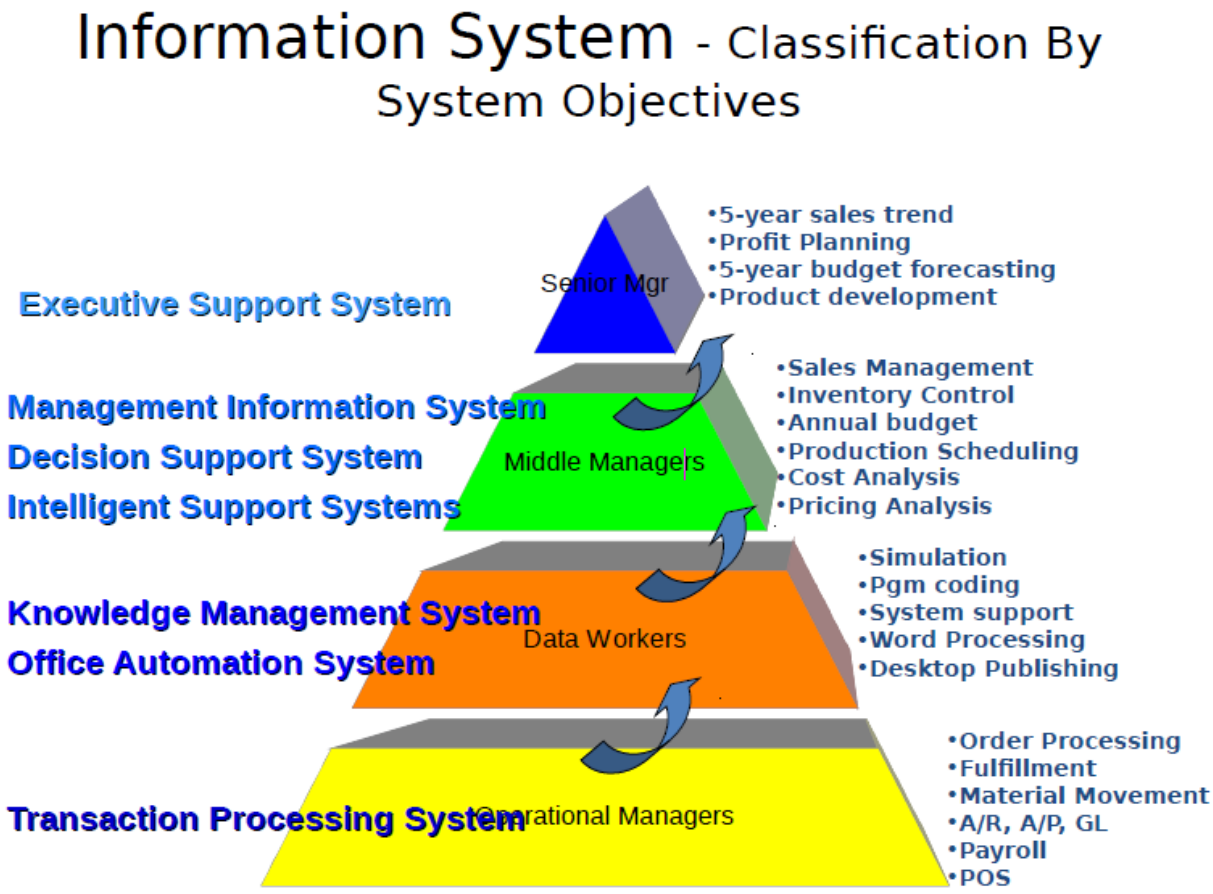


Figure: Inter-organizational IS

Classification by Mode of Data Processing

• Batch Processing Systems:
– Transactions are collected as they occur, but processed periodically, say, once a day or week.
• On-line Batch Systems:

– Transaction information is captured by on-line data entry devices and logged on the system, but it is processed periodically as in batch processing systems.
• On-line Real-time Systems:
– Transaction data capture as well as their processing in order update is carried out in real-time as the transaction is taking place



Transaction Processing System (TPS)
• TPS automates routine and repetitive tasks that are critical to the operation of the organization, such as preparing a payroll, billing customers, Point-of-Sale and Warehouse operations.
• Data collected from this operation supports the MIS and DSS systems employed by Middle Management
• Computerizes the primary and most of the secondary activities on the Value Chain.
• Primary purpose to perform transactions and collect data.

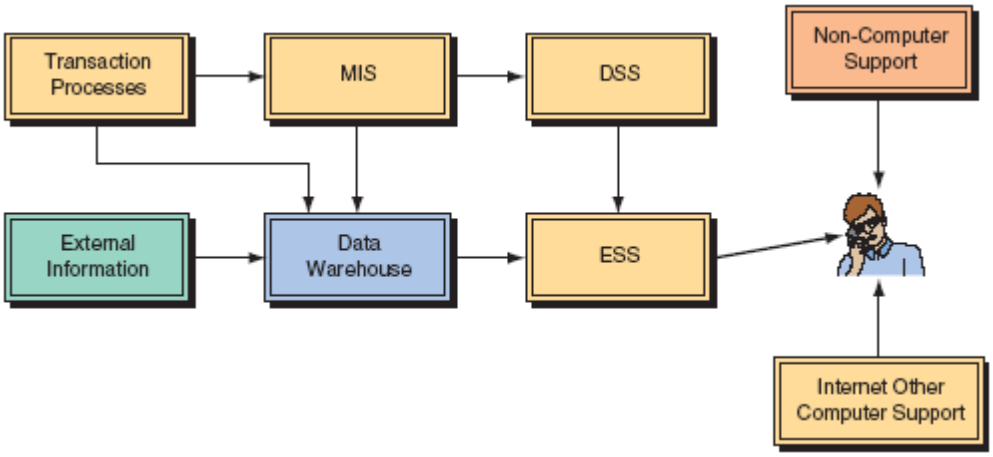
Management Information Systems (MIS)
• These systems access, organize, summarize, and displayed information for supporting *routine decision making* in the functional areas. Geared toward middle managers, MIS are characterized mainly by their ability to produce periodic reports such as a daily list of employees and the hours they work, or a monthly report of expenses as compared to a budget
• Typical uses would be in Replenishment, Pricing Analysis (Markdowns) and Sales Management
• Decisions supported are more structured.
• Primary purpose to process data into information

Decision Support Systems (DSS)
• These systems support complex non-routine decisions.

• Primary purpose to process data into information
• DSS systems are typically employed by tactical level management whose decisions and what-if analysis are less structured.
• This information system not only presents the results but also expands the information with alternatives.
• Some DSS methodologies
– Mathematical Modeling
– Simulation
– Queries
– What-If (OLAP-Cubes)
– Data mining

Intelligent Support Systems (ISS)
• Essentially, artificial intelligence (AI) these systems perform intelligent problem solving.
• One application of AI is expert systems. *Expert systems* (ESs) provide the stored knowledge of experts to non-experts, so the latter can solve difficult or time consuming problems. These advisory systems differ from TPS, which centered on data, and from MIS and DSS, which concentrated on processing information. With DSS, *users* make their decisions according to the information generated from the systems. With ES, the *system* makes recommended decisions for the users based on the built-in expertise and knowledge.

Executive Support Systems (ESS)
• ESS systems or Enterprise Information Systems (EIS) originally were implemented to support Senior management. These systems have been expanded to support other managers within the enterprise.
• At the senior management level they support *Strategic activities* which deal with situations that significantly may change the manner in which business is done.



Office Automation Systems (OAS)
• Electronic communication is only one aspect of what is now known as an *office automation system* (OAS). Other aspects include *word processing systems*, *document management systems* and *desktop publishing systems*.
• OAS systems are predominantly used by *clerical workers* who support managers at all levels. Among clerical workers, those who use, manipulate, or disseminate information are referred to as **data workers**.

Knowledge Management Systems (KMS)
• An additional level of *staff support* now exists between top and middle management. These are professional people, such as financial and marketing analysts that act as advisors and assistants to both

top and middle management. They are responsible for finding or developing new knowledge (External Content) for the organization and integrating it with existing knowledge (Internal Content).
• KMS that support these **knowledge workers** range from Internet search engines and expert systems, to Web-based computer-aided design and sophisticated data management systems

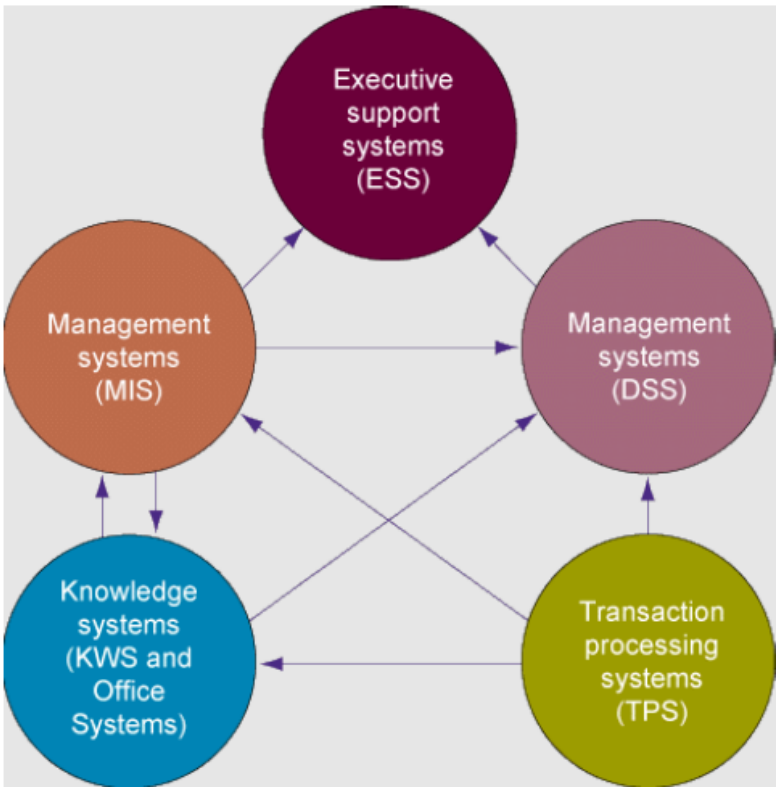
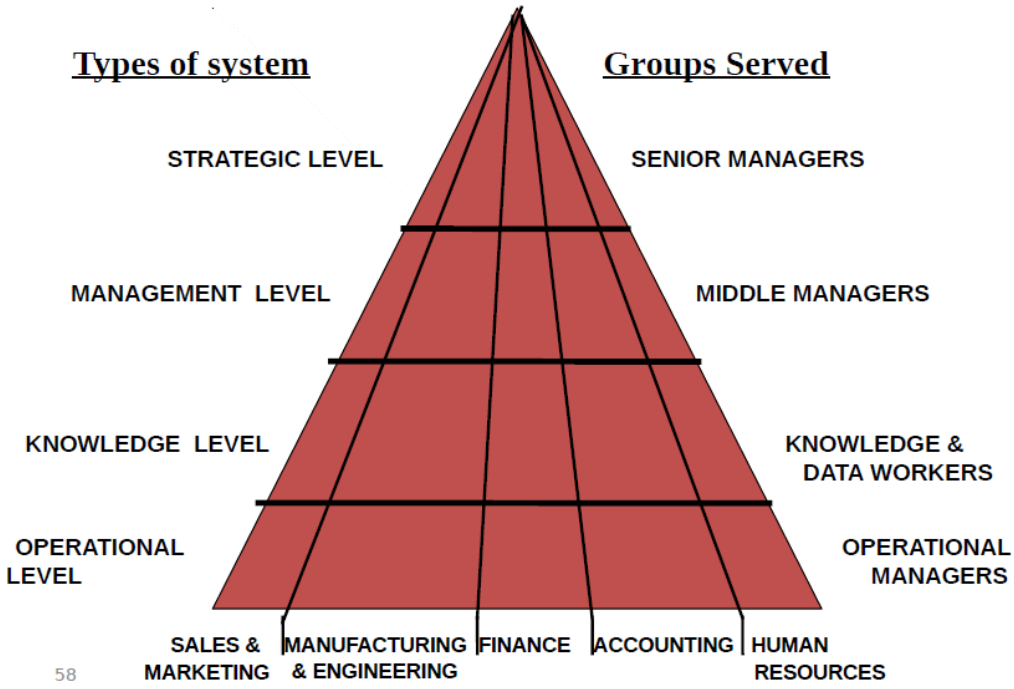


Fig. **INTERRELATIONSHIPS AMONG SYSTEMS**

1.2 IS in Functional Area



• **Functional Perspective *Marketing***
– Identify customers
– Determine what they want
– Planning products
– Advertising and promoting products
– Determine prices for products

• **Functional Perspective *Sales***
– Contact customers
– Sell the product
– Take the order
– Follow-up on the sale
– 5 year sales forecast

• **Functional Perspective *Manufacturing***
– Control Equipment and machinery
– Design new products
– When and quantity of products toproduce
– New production facilities
– Generate the work order

• **Functional Perspective *Purchasing***
– Which vendors
– Quantity to purchase
– Coop, rebate tracking
– Handle delivery discrepancies
– Generate the purchase order

• **Functional Perspective *Finance***
– Financial Assets
– Investment management
– Banking
– Long term budgets

• **Functional Perspective *Accounting***
– Accounts Receivable
– Disbursement (payment)
– Payroll
– Depreciation
– Earned Coop and Rebates

• **Functional Perspective *Human Resources***
– Employee wages, salaries & benefits
– Long term labor requirements
– Tracking vacation, sick,
– Track employee skills
– Interview and review employees

Major functions of systems: Sales management, market research, promotion, pricing, new products
Major application systems: Sales order info system, market research system, pricing system

1.3 Information System Architecture
• An information system architecture is a formal definition of

– the business processes and rules,
– systems structure,
– technical framework, and
– product technologies for a business or organizational information system.
• An information system architecture usually consists of four layers:
– business process architecture,
– systems architecture,
– technical architecture, and
– product delivery architecture.

• The architecture of an information system encompasses the hardware and software used to deliver the solution to the final consumer of services.
• The architecture is a description of the design and contents of a computerized system.
• If documented, the architecture may include information such as
– a detailed inventory of current hardware, software and networking capabilities;
– a description of long-range plans and priorities for future purchases, and
– a plan for upgrading and/or replacing outdated equipment and software.
• The architecture should document: What data is stored?, How does the system function?, Where are components located?, When do activities and events occur in the system?, and Why does the system exist?

Enterprise Architecture
• Enterprise architecture is an ongoing business function that helps an 'enterprise' figure out how to execute best the strategies that drive its development
• **Enterprise architecture and system development** defines
– the broad structure of a system, consisting of its parts,
– their interrelationships and
– other visible properties
Consists of constituent units such as
1. Business architecture -Describes the processes the business uses to meet its goals,
2. Technology architecture Describes the hardware and software infrastructure that supports applications and their interactions and
3. Information systems architecture -Describes how specific Applications are designed and how they interact with each other (application Architecture) & describes how the enterprise data stores are organized and accessed (Data Architecture).
Taken together, they ensure that an organization:
• Meets stakeholder needs
• Aligns its IT with business purposes
• Integrates all departments
• Promotes security
• Brings about data integrity and consistency and
• Reduces duplication and is cost-effective

• helps to establish the rules of governance and the ways in which the architecture process can be managed

• Enterprise architecture engagement ensures that EA standards and guidelines are put into effect.
• It sets out rules specifying ways in which

Enterprise Architecture and strategic planning can be carried out by projects.
• It is an intrinsic part of enterprise architecture governance process.

1.4 Qualities of Information System

• It is often pragmatically defined as: "The fitness for use of the information provided."

• "Information quality" is a measure of the value which the information provides to the user of that information

• **Authority/Verifiability:**

Authority refers to the expertise or recognized official status of a source.

Verifiability refers to the ability of a reader to verify the validity of the information irrespective of how authoritative the source is.

• **Scope of coverage:** Scope of coverage refers to the extent to which a source explores a topic.

• **Composition and Organization:**

Composition and Organization has to do with the ability of the information source to present it's particular message in a coherent, logically sequential manner.

• **Validity:** Validity of some information has to do with the degree of obvious truthfulness which the information caries

• **Uniqueness:** As much as 'uniqueness' of a given piece of information is intuitive in meaning, it also significantly implies not only the originating point of the information but also the manner in which it is presented and thus the perception which it conjures.

• **Timeliness:** Timeliness refers to information that is current at the time of publication. Consider publication, creation and revision dates. Beware of Web site scripting that automatically reflects the current day's date on a page.

• **Reproducibility:** Means that documented methods are capable of being used on the same data set to achieve a consistent result.

1.5 Managing Information System Resources

Resource-

• A **resource** is a source or supply from which benefit is produced

• Typically resources are materials, money, services, staff, or other assets that are transformed to produce benefit and in the process may be consumed or made unavailable.

• Benefits of resource utilization may include

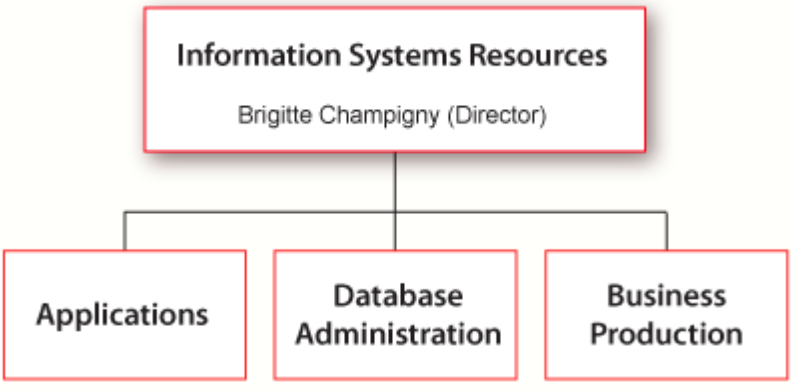
– increased wealth,

– meeting needs or wants,

– proper functioning of a system,

– or enhanced well being.

– From a human perspective a natural resource is anything obtained from the environment to satisfy human needs and wants.



• **Information Systems Resources (ISR)** enables the use of information systems in administrative functions so that those functions may be conducted in the most effective way possible.

Overall mission is to:
• Promote, enable, and participate in the incorporation of information systems into administrative functions;
• Facilitate access, use, exchange and management of administrative information resources;
• Provide and operate an information systems infrastructure and service;
• Supply systems development and systems management expertise;
• Facilitate and promote information systems development in departments through specification of architectures, standards and development of local expertise.
• Information Systems Resources are Networks, Hardware, Software, Data & People

1.6 Balanced Scorecard (BSC)

The Balanced Scorecard is a management tool that provides stakeholders with a comprehensive measure of how the organization is progressing towards the achievement of its strategic goals.
• Balances financial and non-financial measures
• Balances short and long-term measures
• Balances performance drivers (leading indicators) with outcome measures (lagging indicators)
• Should contain just enough data to give a complete picture of organizational performance… and no more!
• Leads to strategic focus and organizational alignment.

Historical Background-BSC
• It was originated by Drs. Robert Kaplan (Harvard Business School) and David Norton as a performance measurement framework that added strategic non-financial performance measures to traditional financial metrics to give managers and executives a more 'balanced' view of organizational performance.
• While the phrase balanced scorecard was coined in the early 1990s, the roots of the this type of approach are deep, and include the pioneering work of General Electric on performance measurement reporting in the 1950's and the work of French process engineers (who created the *Tableau de Bord* – literally, a "dashboard" of performance measures) in the early part of the 20th century.
• The balanced scorecard has evolved from its early use as a simple performance measurement framework to a full strategic planning and management system.
• The "new" balanced scorecard transforms an organization's strategic plan from an attractive but passive document into the "marching orders" for the organization on a daily basis.
• It provides a framework that not only provides performance measurements, but helps planners identify what should be done and measured.
• It enables executives to truly execute their strategies.

Why?

• To achieve strategic objectives.
• To provide quality with fewer resources.
• To eliminate non-value added efforts.
• To align customer priorities and expectations with the customer.
• To track progress.
• To evaluate process changes.
• To continually improve.
• To increase accountability

Advantages to this Approach
 Simple to Use and Understand

 Based on Vision and Strategy
 Multidimensional
•Quantitative and Qualitative Measures
•Current and Future
 Provides Measurement of and Method for Improving our Services
 Ties QI initiatives together
 Serves as a Communication Tool

• "The balanced scorecard retains traditional financial measures.
• But financial measures tell the story of past events, an adequate story for industrial age companies for which investments in long-term capabilities and customer relationships were not critical for success.
• These financial measures are inadequate, however, for guiding and evaluating the journey that information age companies must make to create future value through investment in customers, suppliers, employees, processes, technology, and innovation."

The Learning & Growth Perspective
• This perspective includes employee training and corporate cultural attitudes related to both individual and corporate self-improvement.
• In a knowledge-worker organization, people -- the only repository of knowledge -- are the main resource.
• In the current climate of rapid technological change, it is becoming necessary for knowledge workers to be in a continuous learning mode.
• Metrics can be put into place to guide managers in focusing training funds where they can help the most
• In any case, learning and growth constitute the essential foundation for success of any knowledge-worker organization.
• Kaplan and Norton emphasize that 'learning' is more than 'training';
• it also includes things like mentors and tutors within the organization, as well as that ease of communication among workers that allows them to readily get help on a problem when it is needed.
• It also includes technological tools; what the Baldrige criteria call "high performance work systems."

The Business Process Perspective
• This perspective refers to internal business processes.
• Metrics based on this perspective allow the managers to know how well their business is running, and whether its products and services conform to customer requirements (the mission).
• These metrics have to be carefully designed by those who know these processes most intimately;
• with the unique missions these are not something that can be developed by outside consultants.

The Customer Perspective
• Recent management philosophy has shown an increasing realization of the importance of customer focus and customer satisfaction in any business.
• These are leading indicators: if customers are not satisfied, they will eventually find other suppliers that will meet their needs.
• Poor performance from this perspective is thus a leading indicator of future decline, even though the current financial picture may look good.
• In developing metrics for satisfaction, customers should be analyzed in terms of kinds of customers and the kinds of processes for which we are providing a product or service to those customer groups.

The Financial Perspective
• Kaplan and Norton do not disregard the traditional need for financial data.

• Timely and accurate funding data will always be a priority, and managers will do whatever necessary to provide it.
• In fact, often there is more than enough handling and processing of financial data.
• With the implementation of a corporate database, it is hoped that more of the processing can be centralized and automated
• But the point is that the current emphasis on financials leads to the "unbalanced" situation with regard to other perspectives.
• There is perhaps a need to include additional financial-related data, such as risk assessment and cost benefit data, in this category.

Strategy Mapping
• Strategy maps are communication tools used to tell a story of how value is created for the organization.
• They show a logical, step-by-step connection between strategic objectives (shown as ovals on the map) in the form of a cause-and-effect chain.
• Generally speaking, improving performance in the objectives found in the Learning & Growth perspective (the bottom row) enables the organization to improve its Internal Process perspective Objectives (the next row up), which in turn enables the organization to create desirable results in the Customer and Financial perspectives (the top two rows).
• Traditional financial reports look backward
– Reflect only the past: spending incurred and revenues earned
– Do not measure creation or destruction of future economic value
• The Balanced Scorecard identifies the factors that create long-term economic value in an organization, for example:
– Customer Focus: satisfy, retain and acquire customers in targeted segments
– Business Processes: deliver the value proposition to targeted customers
• innovative products and services
• high-quality, flexible, and responsive operating processes
• excellent post-sales support
– Organizational Learning & Growth:
• develop skilled, motivated employees;
• provide access to strategic information
• align individuals and teams to business unit objectives

Not all Environments are Appropriate for a Balanced Scorecard
• Balanced Scorecard must be driven from the top:
– CEO/COO as sponsor
– Executive leadership team commitment
• A clear sense of purpose is required to:
– Drive change
– Clarify and gain consensus about strategy
– Build a senior executive team
– Focus the organization: align programs and investments
– Integrate cross-functionally
– Educate and empower the organization
• The dynamics of the senior executive team will determine whether the Balanced Scorecard becomes a strategic management system

Chapter 2: Control Audit & Security of IS

LEARNING GOALS

• Why controls are necessary in Information systems?
• Methods of controlling Information systems?
• How controls are introduced in Information systems?
• Why Information systems need auditing?
• How are systems audited?
• How the security of an Information system is ensured?

2.1 Control of Information systems

Control- Method to ensure that a system processes data as per design and that all data is included and are correct

Motivation for control:

• It is very important to ensure the reliability of reports produced by an information system
• If unreliability is seen by users the entire credibility of the system is lost
• Ensuring reliability is not difficult for small systems but when a system has to handle massive data it is a challenge
• Systematic controls are thus essential when a system is designed

Need of controls:

• Information systems handle massive amounts of data – accidents such as not including some data can cause serious damage
• Incorrect data entry can lead to high monetary losses
• Credibility in the information system may be lost if errors are found in operational systems

Objectives of control:

• To make sure data entering the computer are correct
• Check clerical handling of data before it is input to a computer
• Provide means of detecting and tracing errors which occur due to bad data or bad program
• Ensure legal requirements are met
• To guard against frauds

Types of controls: General controls and application controls

General controls

• Govern design, security, and use of computer programs and data throughout organization's IT infrastructure
• Combination of hardware, software, and manual procedures to create overall control environment

Types of general controls

• Software controls
• Hardware controls
• Computer operations controls
• Data security controls
• Implementation controls
• Administrative controls

Application controls

• Specific controls unique to each computerized application, such as payroll or order processing
• Include both automated and manual procedures
• Ensure that only authorized data are completely and accurately processed by that application

Types of application controls:

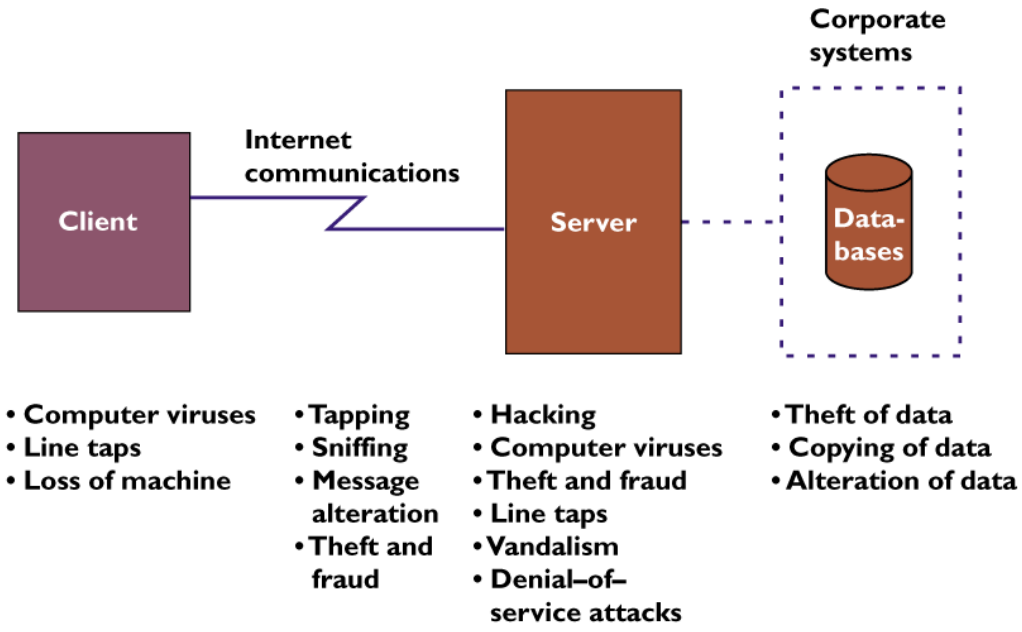
• Input controls
• Processing controls
• Output controls

Example :Application Control-Protecting the Digital Firm
• On-line transaction processing:
Transactions entered online are immediately processed by computer
• Fault-tolerant computer systems: Contain extra hardware, software, and power supply components
• High-availability computing: Tools and technologies enabling system to recover from a crash
• Disaster recovery plan: Runs business in event of computer outage
• Load balancing: Distributes large number of requests for access among multiple servers
• Mirroring: Duplicating all processes and transactions of server on backup server to prevent any interruption
• Clustering: Linking two computers together so that a second computer can act as a backup to the primary computer or speed up processing

Firewalls
• Prevent unauthorized users from accessing private networks
• Two types: proxies and stateful inspection

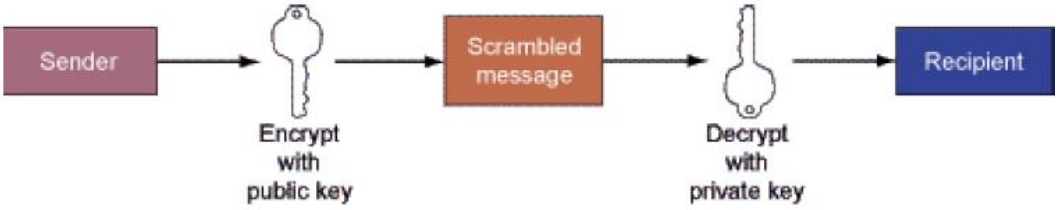
Intrusion Detection System
• Monitors vulnerable points in network to detect and deter unauthorized intruders

Internet Security Challenges:



Security and Electronic Commerce:
• Encryption: Coding and scrambling of messages to prevent their access without authorization
• Authentication: Ability of each party in a transaction to ascertain identity of other party
• Message integrity: Ability to ascertain that transmitted message has not been copied or altered
• Digital signature: Digital code attached to electronically transmitted message to uniquely identify contents and sender
• Digital certificate: Attachment to electronic message to verify the sender and to provide receiver with means to encode reply
• Secure Electronic Transaction(SET): Standard for securing credit card transactions over Internet and other networks

Public Key Encryption:



Developing a Control Structure: Costs and Benefits
Criteria for determining control structure
• Importance of data
• Efficiency, complexity, and expense of each control technique
• Level of risk if a specific activity or process is not properly controlled

The Role of Auditing in the Control Process: MIS audit
• Identifies all controls that govern individual information systems and assesses their effectiveness

2.2 Auditing of information systems:
Audit and testing: Ensure that the system is built as per specifications and that processed results are correct.
• Protect systems from frauds.
Motivation for audits:
• Many organizations are now entirely dependent on computer based information system
• These information systems contain financial data and other critical procedures
• It is essential to protect the systems against frauds and ensure that sound accounting practices are followed
• It is necessary to trace the origin and fix responsibilities when frauds occur
• Audit methods primary purpose is to ensure this.
Objectives:
• Ensure computer based financial and other information reliable
• Ensure all records included while processing
• Ensure protection from frauds

IS Audit: Any audit that wholly or partially evaluates automated information processing system, related non-automated processes, & their interfaces

Audit Planning:
• **Short-Term**: What do we need to audit this year?
• **Long-Term**: What should we plan to audit in the future?
• What should we test first?
– What parts of our business are the most susceptible to risk?
– What business/IS systems are changing?
– Are new evaluation tools available?
– What regulations must we test for?
– Are there new regulations to test for?

Steps in audit:

Step 1: Obtain Understanding of Audit Subject Area: May include:
• Tour facilities related to audit
• Read background material
• Review business and IT strategic plans
• Interview key managers to understand business
• Review prior audit reports

• Identify applicable regulations
• Identify areas that have been outsourced

Step 2: Perform Risk Assessment: Risk-Based Auditing
Inherent Risk: Susceptibility to a problem – E.g., a bank's inherent risk is a robber
Control Risk: A problem exists that will not be detected by an internal control system
– For bank: A thief accesses another's account at Money Machine but is not detected
Detection Risk: An auditor does not detect a problem that does exist
– For bank: Fraud occurs but is not detected
Overall Audit Risk: Combination of audit risks

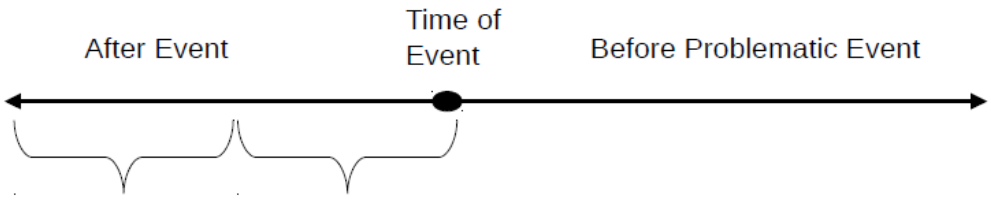
Step 2: Perform Risk Assessment: Prepare Audit Plan
• Develop risk-based approach
• Include audit objectives, scope, timing, required resources
• Comply with applicable law
• Develop audit program and procedures

Step 3: Add Detail to Plan
• Translate basic audit objective into specific IS audit objectives
• Identify and select the audit approach to verify and test controls
• Identify individuals to interview
• Obtain departmental policies, standards, procedures, guidelines to review
• Develop audit tools and methodology

Step 4: Evaluate Audit Area: Tools for the Auditor
ISACA has Standards and Guidelines related to Audit
• Section 2200 General Standards
• Section 2400 Performance Standards
• Section 2600 Reporting Standards
• Section 3000 IT Assurance Guidelines
• Section 3200 Enterprise Topics
• Section 3400 IT Mgmt Processes
• Section 3600 IT Audit and Assurance Processes
• Section 3800 IT Audit and Assurance Mgmt

Step 5: Evaluate Controls
Review IS Organization: Separation of duties
Review IS Policies, Standards, Procedures: Defined, periodically updated
Review IS Documentation: Policy, Procedures, Design, Test, Operations, Contract/SLAs, Security
Interview personnel: Segregation of duties, security awareness, competency
Observe personnel: Document everything in sufficient detail

Evaluate Controls: IT Control Classifications



Corrective Controls:
Fix problems and prevent future problems
Includes:
Contingency planning
Backup procedures
Reruns

Detective Controls:
Finding fraud when it occurs
Includes:
Hash totals
Check points
Duplicate checking
Error messages
Past-due account reports
Review of activity logs

Preventive Controls*:
Preventing fraud

Includes:
Programmed edit checks
Encryption software
Access control S/W
Well-designed procedures
Physical controls
Employ only qualified personnel

Compensating Control: A strong control supports a weak one.
Overlapping Control: Two strong controls

Step 6 & 7: Audit Test
Evidence: Audit findings must be based on sufficient and reliable evidence and appropriate interpretation of the evidence
Documentation: The audit work and audit evidence to support conclusions must be fully documented
Supervision: Audit staff is supervised to ensure that audit is professionally completed
Professional Skepticism: The auditor must keep an eye open for irregularities and/or illegal acts, unusual relationships, material misstatements
– when irregularities are encountered, the auditor should:
• Investigate fully
• document all communications, tests, evidence, findings
• report the irregularity to governance body in a timely manner

Step 6: Compliance Testing
• Are controls in place and consistently applied?
– Access control
– Program change control
– Procedure documentation
– Program documentation
– Software license audits
– System log reviews
– Exception follow-ups

• Control: Is production software controlled?
– Test: Are production executable files built from production source files?
– Test: Were proper procedures followed in their release?
• Control: Is Sales DB access constrained to Least Privilege?
– Test: Are permissions allocated according to documentation?
– Test: When sample persons access DB, can they access only what is allowed?

Step 7: Substantive Testing
• Are transactions processed accurately?
• Are data correct and accurate?
• Double check processing
– Calculation validation
– Error checking
– Operational documentation
• If Compliance results are poor, Substantive testing should increase in type and sample number

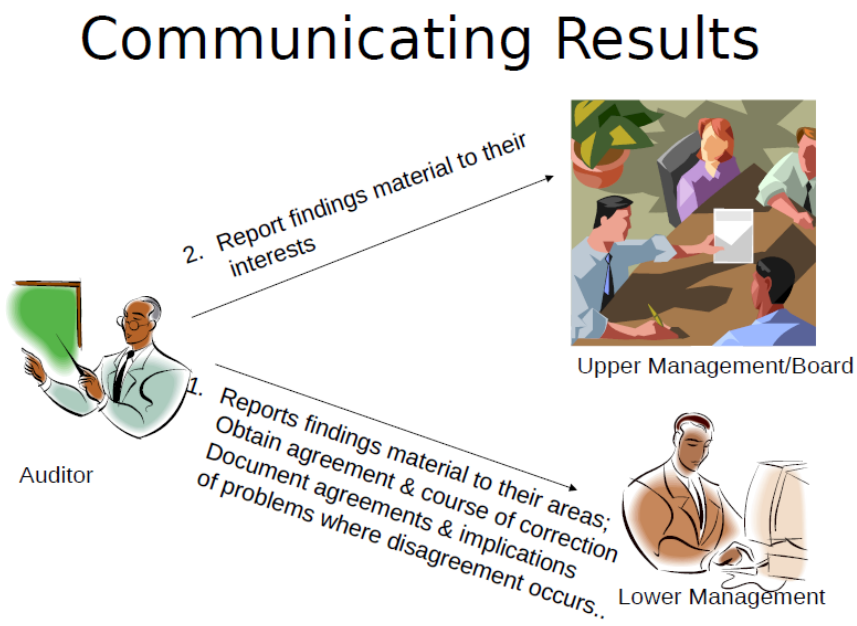
• Audit: Is financial statement section related to sales accurate?
– Test: Track processing of a sample transactions through the system, performing calculations manually
– Test: Test error conditions
• Audit: Is tape inventory correct?
– Test: Search for sample days and verify complete documentation and tape completeness
Sampling
Statistical Sampling:
• N% of all items randomly tested
• Should represent population distribution
Nonstatistical (or Judgment) Sampling:
• Auditor justifies another distribution for sample selection
• Which items are most risky?
Under what conditions do you think one is better?

Generalized Audit Software (GAS)
• File Access: Read records & file structures
• File reorganization: Allow sorting, indexing, merging/linking with other files
• Data Selection: Select a set of records
• Statistical functions: Perform sampling, stratification, frequency analysis
• Arithmetic Functions: Perform arithmetic operations on data sets

Step 8: Prepare Audit Report
Identify:
• Organization, recipients, restriction on circulation
• Scope, objectives, period of coverage, nature, timing and extent
• Findings, conclusions, recommendations/follow up, and reservations or qualifications
– Grouped by materiality or intended recipient
– Mention faults and constructive corrections
• Evidence to support results (may be separate)
• Overall findings, conclusion, & opinion
• Signed & dated

Evidence
Forms of Evidence:
• Notes from Interviews
• Test Results
• Email or mail correspondence
• Documentation
• Observations
Best Sources:
• External: Sources from outside organization

• Qualified: Most knowledgeable
• Objective: Evidence not prone to judgment
• Timing: Should match period under review



Step 9: Follow-up
• Has management taken appropriate action to fix problems in a timely manner?
• Request and evaluate information on follow-up
– Management should schedule implementation of correction
– May be scheduled for convenient time
– Next audit these follow-ups should be Checked

Final IMPORTANT Recommendation
IS Audits can result in system failures, problems, etc.
Protect Yourself:
• Get an approval signature for your audit plan before you begin.
• If you will be impacting the system at all, send an email to all affected and talk to the administrators before starting any tests
• When working with data or devices, be careful not to be the CAUSE of any problems; be careful not to change live data or configurations for test purposes: Work on a copy!
• Preferably have an escort for all that you do
There is one difference between a hacker and auditor: **Permission!!!**

Classifications of Audit:
Financial Audit: Assure integrity of financial statements
Operational Audit: Evaluate internal controls for a given process or area
Integrated Audit: Includes both Financial and Operational aspects
Forensic Audit: Follows up on fraud/crime
IS Audit: Does IS safeguard data, provide CIA in efficient way?
Administrative Audit: Assess efficiency of a process or organization

Computer-Assisted Audit Techniques (CAAT):
• Software tools enable auditor to
– Access and analyze data in database

– Perform compliance tests
– Perform penetration and vulnerability tests
– Test Application
• May include utility software, debug or scanning software, test data, application trace, expert systems, generalized audit software
• Special use:
– Referenced in audit plan & report
– Download sample data and use in read-only mode

Computer Assisted Auditing Techniques & Tools:
• Query systems, report writers, utilities, computer languages
• Complete files can be read speedily
• Can use parameters that may be altered each time program is run
• Once programs are set up, time savings are significant
• Allows auditor independence

CAAT: types of software
• Automated audit workpapers
• Data Analysis
• Risk assessment
• Scheduling
• Timekeeping
• Flowcharting
• Report generation

Use in fraud detection and investigation
• Terminated employees being paid
• Ghost employees
• Purchases to homes instead of business
• "On-call" pay abuse Unusually high salary increases
• Telephone use abuse
• Travel reimbursement abuse

Use in network security
• Port scanning tools
• Network intrusion detection
• SANS "Top 20 Network Vulnerabilities"
• Computer Intrusion Response Teams

Control Self-Assessment
• Internal audit system that enhances external audit
• Control monitoring occurs in functional areas
• Includes designing and assessing controls locally, often in workshops
• Benefit: Involves and trains employees, often reducing risk quicker

Emerging Audit Techniques:
Automated Work Papers: Automated tools for risk & audit reporting
Integrated Audit: Combines financial and IS audit via team effort
Continuous Audit: Provides audit reports on continuous basis (not just quarterly)

2.3 Security of Information System

Protection of data resources, programs, and equipment from illegal use , theft , vandalism, accidents, disasters etc. Security means protection of data from accidental or intentional modification, destruction or disclosure to unauthorized persons. IS security refers to precaution taken to keep all the aspects of information systems (e.g. H/w, S/w, N/w equip & data).

Motivation for security

• Systems contain sensitive data about the organization and also about persons working in the organization
• Sensitive data should be protected from spies, thieves or disgruntled employees.
• Thus access should be carefully controlled and provided only on a need to know basis
• When computers are networked corruption/erasure may take place due to viruses Services may be disrupted due to denial of service attacks
• Thus systems should be designed with appropriate Security

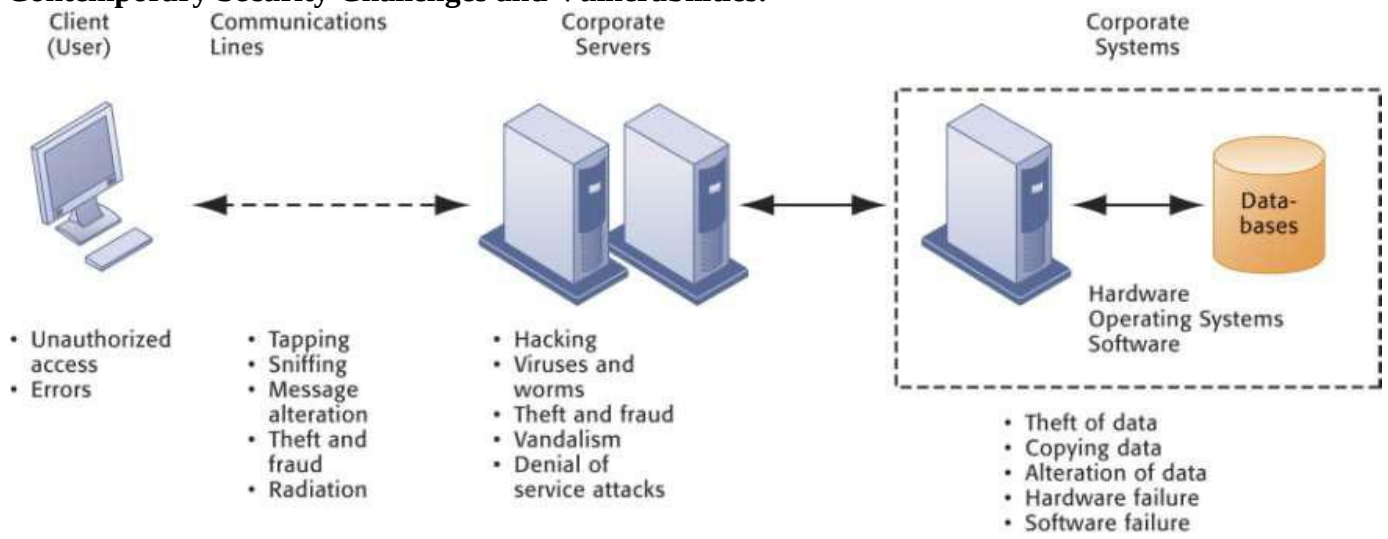
Potential threats to security

• Natural disasters such as fire, floods, earthquakes
• Accidents such as disk crashes, file erasure by inexperienced operators
• Theft/erasure of data by disgruntled employees & consultants
• Links to business associates- electronic information can be risky when it travels between or among business affiliates as a part of doing business
• Frauds by changing programs, data by employees
• With the advent internet & related telecom technologies systems have become more vulnerable
– Viruses/Worms as an email attachment
– Hackers who break into systems connected to the internet
– Denial of service attacks by flooding with mail

Most IS are compromised through

• Unauthorized access
• Information modification
• Denial of Service & viruses as well as
• Spam
• Spyware &
• Cookies

Contemporary Security Challenges and Vulnerabilities:



System vulnerabilities & abuse:

• **Viruses:** Rogue software program that attaches itself to other software programs or data files in order to be executed

• **Worms:** Independent computer programs that copy themselves from one computer to other computers over a network
• **Trojan horses:** Software program that appears to be benign but then does something other than expected
• **Spyware:** Small programs install themselves surreptitiously on computers to monitor user Web surfing activity and serve up advertising
• **Key loggers:** Record every keystroke on computer to steal serial numbers, passwords, launch Internet attacks
• **Hackers and crackers**
Activities include
– **System intrusion**
– **Theft of goods and information**
– **System damage**
– **Cyber vandalism**
• Intentional disruption, defacement, destruction of Web site or corporate information system
Spoofing:
• Misrepresenting oneself by using fake e-mail addresses or masquerading as someone else
• Redirecting Web link to address different from intended one, with site masquerading as intended destination
Sniffer: Eavesdropping (spying)program that monitors information traveling over network
Denial-of-service attacks (DoS): Flooding server with thousands of false requests to crash the network
Distributed denial-of-service attacks (DDoS): Use of numerous computers to launch a DoS
Botnets: Networks of "zombie" PCs infiltrated by bot malware
Cookies: A message passed to the web browser on a user's computer by web server. The browser then stores a message in a text file, and the message is sent back to the server each time the user's browser request a page from that server. Cookies are stored on a Hard disk of your computer without your Knowledge

Computer crime:
• Defined as "any violations of criminal law that involve a knowledge of computer technology for their perpetration, investigation, or prosecution"
Computer may be target of crime, e.g.:
• Breaching confidentiality of protected computerized data
• Accessing a computer system without authority
Computer may be instrument of crime, e.g.:
• Theft of trade secrets
• Using e-mail for threats or harassment

• **Identity theft:** Theft of personal Information (social security id, driver's license or credit card numbers) to impersonate someone else
• **Phishing:** Setting up fake Web sites or sending e-mail messages that look like legitimate businesses to ask users for confidential personal data.
• **Evil twins:** Wireless networks that pretend to offer trustworthy Wi-Fi connections to the Internet
• **Pharming:** Redirects users to a bogus Web page, even when individual types correct Web page address into his or her browser

Click fraud: Individual or computer program clicks online ad without any intention of learning more or making a purchase
Global threats - Cyberterrorism and cyberwarfare: Concern that Internet vulnerabilities and other networks make digital networks easy targets for digital attacks by terrorists, foreign intelligence services, or other groups

How to protect data/ programs

• Regular back up of data bases every day/or week depending on the time criticality and size
• Incremental back up at shorter intervals
• Backup copies kept in safe remote location particularly necessary for disaster recovery
• Duplicate systems run and all transactions mirrored if it is a very critical system and cannot tolerate any disruption before storing in disk.
• Physical locks
• Password system
• Biometric authentication (Eg: Finger print)

• Encrypting sensitive data/programs
• Identification of all persons who read or modify data and logging it in a file
• Training employees on data care/handling and security
• Antivirus software
• Firewall protection when connected to internet
    • Types(can both be hardware & software)
        • Packet filter
        • Application level control: measures security only for certain application such as file transferring
        • Circuit level control: measures security when certain kind of connection(circuit) is made
        • Proxy server
• Audit - control software

Data security, privacy and integrity
• Data security is concerned with protecting data from erasure, theft, unauthorized access and unauthorized modifications
• Data privacy is concerned with protecting data regarding individuals from being accessed and used without the permission/knowledge of concerned individuals
• Data integrity is concerned with the quality and reliability of raw as well as processed data

What is Layered security

• **Layered security**, also known as **layered defense**, describes the practice of combining multiple mitigating security controls to protect resources and data.
• The term bears some similarity to defense in depth,
• A term adopted from a military strategy that involves multiple layers of defense that resist rapid penetration by an attacker but yield rather than exhaust themselves by too-rigid tactics.
• Because potential Internet security risks can occur at a variety of levels, you need to set up security measures that provide multiple layers of defense against these risks.
• In general, when you connect to the Internet, you should not wonder if you will experience intrusion attempts or denial of service attacks.
• Instead, you should assume that you will experience a security problem.
• Consequently, your best defense is a thoughtful and proactive offense.
• Using a layered approach when you plan your Internet security strategy ensures that an attacker who penetrates one layer of defense will be stopped by a subsequent layer.
• Your security strategy must include measures that provide protection across the following layers of the traditional network computing model.
• Generally, you need to plan your security from the most basic (system level security) through the most complex (transaction level security).

2.4 Consumer Layered security (strategy)

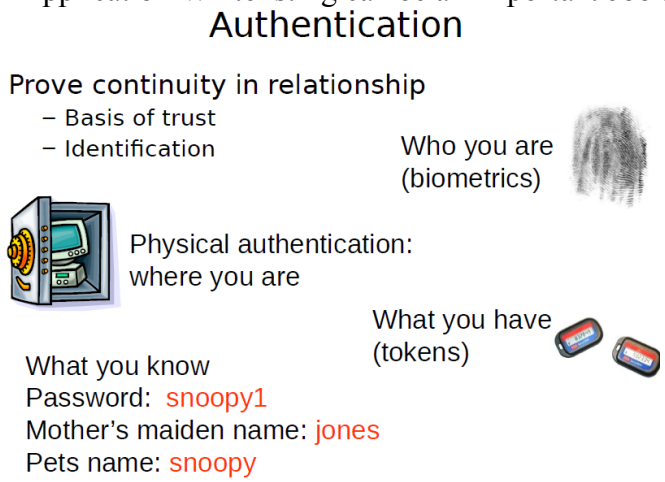
• Consumer (Application) level security measures control how users can interact with specific applications.
• In general, you must configure security settings for each application that you use.

• However, you need to take special care to set up security for those applications and services that you will use from or provide to the Internet.
• These applications and services are vulnerable to misuse by unauthorized users looking for a way to gain access to your network systems.

Various consumer layered security strategies
• Extended validation (EV) SSL certificates
• Multifactor authentication (also sometimes known as versatile or two-factor authentication)
• Single sign-on (SSO)
• Fraud detection and risk-based authentication
• Transaction signing and encryption
• Secure Web and e-mail
• Open fraud intelligence network

2.5 Enterprise Layered Security Strategy
• Workstation application whitelisting
• Workstation system restore solution
• Workstation and network authentication
• File, disk and removable media encryption
• Remote access authentication
• Network folder encryption
• Secure boundary and end-to-end messaging
• Content control and policy-based encryption

Workstation application whitelisting
• Whitelisting-listing the allowed users
• Application Whitelisting is one of the latest tools offered to enhance your "Defense in Depth" security strategy.
• With increasing numbers of attempted intrusions, cautionary tales of security breaches and the potential for resulting damages at your site,
• Application Whitelisting can be an important addition to your security arsenal.



Network Authentication
• Password
• One-time Passwords (ex. tokens)
• Network address
– Caller-id - credit card
– IP address

– MAC address – banks
• Cryptographic protocols

File, disk and removable media encryption:
• Develop and test appropriate a Data Recovery Plan
• Use compliant encryption algorithm and tools when creating a password, follow strong password requirements
• Do NOT use the same password from other systems.
• Use a secure password management tool to store sensitive information such as passwords and recovery keys
–Where passwords need to be shared with other users, ensure that passwords are sent separately from the encrypted file. E.g. call the person to verbally communicate the password.
– Do NOT write down the password and store it at the same location as the storage media (e.g. post-it note with the password next to the encrypted USB drive)
• After the covered data is copied to a removable media,
– Verify that the removable media works by following instructions to read the encrypted covered data
– If applicable, securely delete unencrypted covered data following secure deletion guidelines
• Removable media (e.g. CD, hard disks) should be labeled with the following information:
– Title. For example "Project XYZ Data"
– Data owner (researcher or research unit name)
– Encryption date

• When unattended, the removable media should be stored in a secured and locked location (e.g. cabinets, lock boxes, etc) where access is limited to users on a need-to-know basis.
• Document the physical location of removable media, along with the label information (specified above) for tracking and future reference.

2.6 Extended Validation Certificate (EV)
• It is an X.509 public key certificate issued according to a specific set of identity verification criteria.
• These criteria require extensive verification of the requesting entity's identity by the certificate authority (CA) before a certificate is issued.
• Certificates issued by a CA under the EV guidelines are not structurally different from other certificates (and hence provide no stronger cryptography than other, cheaper certificates), but are designated with a CA-specific policy identifier so that EV-aware software can recognize them.
• The criteria for issuing EV certificates are defined by the Guidelines for Extended Validation Certificates, currently (as of May 2012) at version 1.4.
• The guidelines are produced by the CA/Browser Forum, a voluntary organization whose members include leading CAs and vendors of Internet software, as well as representatives from the legal and audit professions.

Multifactor authentication (versatile or two-factor authentication):
• Multifactor authentication (MFA) is a security system in which more than one form of authentication is implemented to verify the legitimacy of a transaction.
• The goal of MFA is to create a layered defense and make it more difficult for an unauthorized person to access a computer system or network.
• It is an approach to authentication which requires the presentation of two or more of the three authentication factors:
– a *knowledge* factor ("something the user *knows*"),
– a *possession* factor ("something the user *has*"), and
– an *inherence* factor ("something the user *is*").

• Knowledge factors: "something the user knows"
– 1 Password
– 2 PIN
– 3 Pattern
• Possession factors: "something the user has"
– 1 Tokens with a display (disconnected tokens)
– 2 Connected tokens
• 2.1 USB tokens
• 2.2 Smartcards
• 2.3 Audio Port tokens
– 3 One-time pads
– 4 Mobile phones
• Inherence factors: "something the user is"
– Biometrics

Single sign-on (SSO)
• **Single sign-on** (**SSO**) is a property of access control of multiple related, but independent software systems.
• With this property a user logs in once and gains access to all systems without being prompted to log in again at each of them.
• Conversely, **Single sign-of** is the property whereby a single action of signing out terminates access to multiple software systems.
• As different applications and resources support different authentication mechanisms, single sign-on has to internally translate to and store different credentials compared to what is used for initial authentication.

Fraud detection and risk-based authentication
• Fraud detection would be a critical security layer
• Includes Risk-based Authentication as a mechanism for fraud detection
• Risk-based authentication is a technique that uses both contextual and historical user information, along with data supplied during Internet transaction, to assess the probability of whether a user interaction is authentic or not.
• Contextual information includes the username and password, who the user is, their IP addresses, location information, what kind of device they are using.
• Historical user data includes specific attributes provided from the session as well as user behavior and transaction patterns.
• This information represents an additional authentication factor that supplements the username and password, making this an enticing multifactor authentication technique.

Transaction signing and encryption
• Eliminating information piracy, data theft, etc. and ensuring security of information transmitted online is even more necessary as e-payments are fast becoming the norm than the exception.
• Example: A **digital signature** is a mathematical scheme for demonstrating the authenticity of a digital message or document.
• A valid digital signature gives a recipient reason to believe that the message was created by a known sender, such that the sender cannot deny having sent the message (authentication and non-repudiation) and that the message was not altered in transit (integrity).
• Digital signatures are commonly used for software distribution, financial transactions, and in other cases where it is important to detect forgery or tampering.

Secure email & Web
• **Email security** includes a combination of email encryption, digital signatures, content monitoring and policy compliance capabilities that are automatic and invisible to everyday users.

• Use of antivirus also helps in web & email security
• Advanced threat dashboards, forensic reporting and data capture, sandbox analysis of malware, and data-aware defenses that provide containment of sensitive information.

Open fraud intelligence network
• Open Fraud Intelligence Network represents a powerful new weapon on fraud.
• by consolidating information from a variety of sources, including fraud patterns that have been experienced by leading global financial services organizations.
• The combination of fraud data and the ability to share fraud experiences promises to be a significant advancement in the way fraud can be detected and addressed.

2.7 Remote Access Authentication:

• The ability to verify identity (authentication) is even more important for remote users than for those who are on-site
• Without a secure authentication scheme anyone could get into the network and view, copy, change or even destroy important data
• Remote access servers can be configured as dial-in servers or VPN servers. Dial-in servers use the Point-to-Point Protocol (PPP) or in the case of some older servers, the Serial Line Internet Protocol (SLIP) as the link layer protocol.
• VPN servers can use the Point-to-Point Tunneling Protocol (PPTP), Layer 2 Tunneling Protocol (L2TP), or IPSec tunnel mode to establish a secure "tunnel" over the Internet.
• Windows remote access servers support the following set of authentication methods:
– Password Authentication Protocol (PAP)
– Challenge Handshake Authentication Protocol (CHAP)
– Microsoft's implementation of CHAP (MS-CHAP)
– Updated version of MS-CHAP (MS-CHAP2)
– Extensible Authentication Protocol/Transport Layer Security (EAP/TLS)

Network folder encryption:

• With PGP NetShare, organizations can protect their intellectual property and other sensitive information as required by partner and regulatory mandates for information security and privacy.
• Persistent File Encryption on Network Servers PGP NetShare Protected Files & Directories with PGP NetShare, authorized users can save and share files on file servers and use all applications as they currently do while benefiting from the comprehensive protection offered by PGP encryption.
• Content such as documents, spreadsheets, presentations, video, audio, and Web materials is automatically encrypted when saved to a PGP NetShare–protected folder.
• The end-to-end principle is the core architectural guideline of the Internet

2.8 Content control and policy based encryption

• Policy-based Encryption service allow customers to set up filters based on the content of a message, if the message meets the set criteria it will be encrypted.
• Once signed up for this service, all messages sent from MS Exchange mailboxes to external recipients are processed according to configured policies and encrypted, if required.

Policy based encryption

• The Policy Based Encryption (PBE) service encrypts specific emails based on a policy

• i.e. a set of rules -
– analyzes all email, and
– encrypt any email that matches the predefined conditions
• Policy Based Encryption uses the Email Content Control rules to identify which email needs to be encrypted
• The Policy Based Encryption Service is managed through the same control panel that is used to manage Antivirus and Anti-spam settings

Policy Based Encryption and Email Content Control
• The Policy Based Encryption service is closely integrated with the Email Content Control service
• The rule that defines whether an email is to be encrypted is set up in the Email Content Control configuration screens in the Boundary Defense for Email Control Panel
• The encryption rule has an action to redirect any emails that meet the rules conditions to a specified encryption email address
• This email address will be sent to the administrator when the service is purchased
• This email address is used solely to process and encrypt the email

Creating an Encryption Group
• Prior to creating any encryption rules, an encryption group must be created
• This group needs to be added to each rule, as an exception in order for the mail to be forwarded to the Policy Based Encryption Gateway

Defining an Encryption Rule
• To trigger mail to be encrypted, an Email Content Control rule must be configured with an action to redirect the mail to the specified email address for the service you are using
• A rule is defined to include the specific conditions to cause email to be encrypted,
• for example, specific words contained in the header or body of the email
• The Email Content Control service scans email against the rules in the order they are listed in the Control Panel portal
• If an email triggers a rule with an exit action, it is subject to that action and does not pass on to be scanned for further rules
• The redirection action for rules is an exit action
• So it is important to put encryption rules towards the bottom of the rule set, so that other rules defined to comply with the organization's acceptable usage policy are acted on first

• **NOTE: If an email triggers a rule with an exit action, such as a block action higher in the rule set, the** email will not be encrypted, because the first rule, blocking the email, will take precedence.
• **NOTE: It is recommended that test groups be added to a rule initially for testing, to ensure the new** encryption performs as expected.
• This will prevent potential problems on mail flow for the entire organization.
• Test groups are created, only with valid corporate email addresses added to the rule instead non-valid email address

2.9 Example of security in Ecommerce Transaction

E-commerce
• Trade between two parties: where exchange is negotiated under the set of mutual acceptance conditions, so both parties emerge satisfied with result.
• Depends on whether two parties trust each other .

What is e-commerce

• E-commerce is a short version of the term Electronic Commerce
• transactions related to online buying and selling of products or services
• done using electronic systems such as the Internet and other computer networks
• penetration and spread of the internet has fuelled e-commerce
• Examples os e-commerce are electronic funds transfer, supply chain management, Internet marketing, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data collection systems
• definition of e-commerce in modern times implies that it typically uses the World Wide Web at least at any point in the transaction's lifecycle
• Online retailers are sometimes known as e-tailers
• online retail is sometimes known as e-tail

Types of e-commerce

• B2B: E-commerce that is conducted between businesses is referred to as Business-to-business
– (1) open to the entire public or (2) limited to a group of businesses who have been part of the specific group
– Transaction cost reduced through reduction in
• search costs
• costs of processing transactions (e.g. invoices, purchase orders and payment schemes)
• cost in trading processes
– eliminating intermediaries and distributors
– increase in price transparency
– creates supply-side cost-based economies of scale

• B2C Commerce

– commerce between companies and consumers
– involves customers gathering information; purchasing physical goods or information goods
– online retailing companies such as Amazon.com, Drugstore.com, Beyond.com, Flipkart.com, Lenskart.com
– reduces transactions costs
– increasing consumer access to information
– reduces market entry barriers

• B2G e-commerce

– commerce between companies and the government
– use of the Internet for procurement
– licensing procedures

• C2C e-commerce

– commerce between private individuals or consumers
– online auctions
– auctions facilitated at a portal, such as eBay, which allows online real-time bidding on items being sold in the Web;
– peer-to-peer systems, such as the Napster model (a protocol for sharing files between users used by chat forums similar to IRC) and other file exchange and later money exchange models; and
– classified ads at portal sites such as Excite Classifieds and eWanted (an interactive, online marketplace where buyers and sellers can negotiate and which features “Buyer Leads & Want Ads”).
– Consumer-to-business (C2B) transactions involve reverse auctions, which empower the consumer to drive transactions.

– There is little information on the relative size of global C2C e-commerce. However, C2C figures of popular C2C sites such as eBay and Napster indicate that this market is quite large. These sites produce millions of dollars in sales every day

M-Commerce
– buying and selling of goods and services through wireless technology
– handheld devices such as cellular telephones and personal digital assistants (PDAs) are used
– m-commerce will become the choice for digital commerce transactions
– bill payment and account reviews can all be conducted from the handheld devices
– consumers are given the ability to place and pay for orders on-the-fly
– delivery of entertainment, financial news, sports figures and traffic updates to a single mobile
– different server than that accessed by the regular online users
– allow users to book and cancel rail, flight, movie tickets through their mobile devices

Features provided by E-commerce
• Privacy : One's information not known to others.
• Confidentiality : Secure transmission is demanded.
• Repudiation: Only if customer truly initiated the transaction, must be allowed.
• Accountability: Is it's the user who is owner of credit card.
• Integrity : Services are available.
All these depend on the how secure are the endpoints.

Security issues of Ecommerce
• **Network security** This is probably the most obvious issue for ecommerce applications, since the amount and severity of hack attacks are increasing. Fortunately, significant progress has been made in this area through firewall security products that protect against basic network-level attacks. A proper security strategy should not end here, though.
• **Identity** Since e-commerce implies trading with potentially unknown and untrusted partners, identification of trading partners can be crucial. Once again, much has been done to provide standardized methods to identify users by using certificates based on the X.509 standard. Unfortunately the deployment of these certificates for general e-commerce applications has been slow.
• **Encryption** Encryption is a key technology in e-commerce.
• **Authorisation** In order to automate trading processes, it is often required to verify more than identity. Various emerging standards such as certificates, authorization servers and the use of a database with registered users and privileges inside the application, all contribute to address the authorisation issue of who may do what.
• **Host and application security** The protection offered by most operating systems falls short in a global networked environment. Efforts such as signed applets and signed executables are commendable, but will most probably not solve the problem of **virus and Trojan attacks**. In addition to these obvious flaws, more subtle problems such as buffer overflow attacks on certain networked applications can also lead to security compromises.
• **Transaction security** The protocols used for electronic transactions range from the primitive to the very sophisticated. Older secure protocols, for example those used in Point of Sale terminals, rely on the DES algorithm, and typically require some form of secure storage for the cryptographic keys. Many newer protocols (eg. SET) are based on public key mechanisms, but have not yet achieved widespread adoption.
• **Human error and malice** One of the most significant security problems faced by every system - no matter how secure technologically
- is that trust is eventually placed in an individual. Auditing, agreements and regulatory frameworks can help, but are often too slow to react to the current demand for e-commerce.

Steps involved

• ***Step One***. The consumer enters an order along with their credit card information and sends it to the business.
• ***Step Two***. The business sends the consumer an invoice, their certificate and their bank's certificate.
• ***Step Three***. The consumer acknowledges and approves this information and returns it to the business.
• ***Step Four***. The business then generates an authorization request for your credit card and sends it to their bank.
• ***Step Five***. The business's bank then sends the credit authorization request to the Acquirer.
• ***Step Six***. The Acquirer sends an acknowledgement back to the business's bank after receiving an acknowledgement from the consumer's bank.
• ***Step Seven*** Once the consumer's bank authorizes payment, the business's bank sends an acknowledgement back to the business with the authorization number.

Different attacks

• Snooping the shopper's computer
• Sniffing the network
• Using DDOS attacks
• Guessing passwords
• Using known OS-bugs

Simple Risks

• If URL with '?' separating various values, implies that the application is passing parameters to other scripts or programs. One can change these parameters and adjust the way program behaves.
• Price changed to negative number
• Buffer Overflow problems :
  - Do not make assumptions about user input size
  - Do not pass unchecked user input to shell commands
  - Cryptographic checksum is used to prevent unauthorized access Once the file is placed , checksum should be run on it and periodically it is run and checked with the original. To prevent false alarms, update of checksum should be a done.

Cookies:

• The primary use of cookies is to store authentication and session information, your information, and your preferences. A secondary and controversial usage of cookies is to track the activities of users.
• Temporary cookies: These cookies are valid only for the lifetime of your current session, and are deleted when you close your browser. These are usually the good type. They are mostly used to keep your session information.
• Permanent cookies: These are for a time period, specified by the site, on the shopper's computer. They recall your previous session information.
• Server-only cookies: These cookies are usually harmless, and are only used by the server that issued them.
• Third-party cookies: These are usually used for tracking purposes by a site other than the one you are visiting. Your browser or a P3P policy can filter these cookies.
• Encrypted cookies which are non-persistent are best as clear text cookie even though non-persistent , an attacker can design a proxy system between client and server to capture the required content

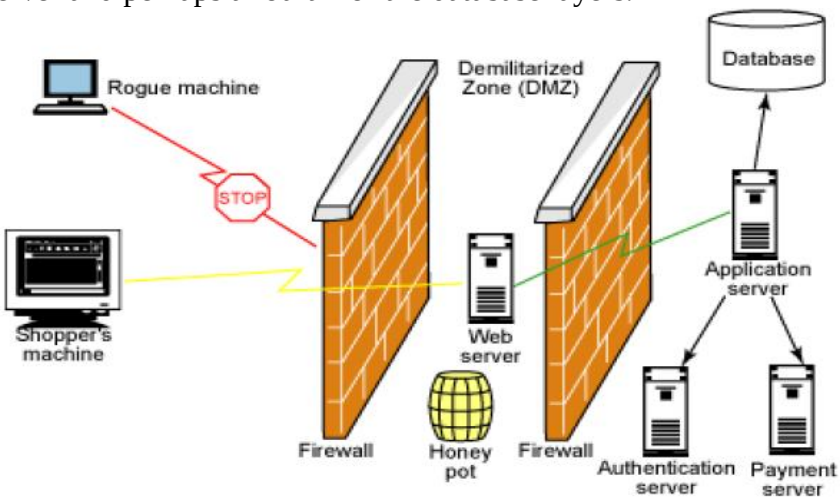
Different attacks

• Access attacks - That penetrate security perimeters to steal information, DDoS attacks paralyze Internet systems by overwhelming servers, network links, and network devices (routers, firewalls, etc.) with bogus traffic. Also called as Smurf attack

• Bandwidth attacks- These DDoS attacks consume resources such as network bandwidth or equipment by overwhelming one or the other (or both) with a high volume of packets. The most common form of bandwidth attack is a packet-flooding attack
• Application attacks—These DDoS attacks use the expected behavior of protocols such as TCP and HTTP HTTP half-open and HTTP error attacks
• Blackholing describes the process of a service provider blocking all traffic destined for a targeted enterprise as far upstream as possible, sending the diverted traffic to a "black hole" where it is discarded in an effort to save the provider's network and its other customers.

Multiple Layers of firewall

• Without robust physical and network security, sensitive corporate data is at risk of intrusion
• The use of firewalls, intrusion detection, client PC virus software, server-based virus checking and keeping all systems up to date with security patches will prevent most type of threats
• Firewalls should restrict access from the Internet and from one internal network (e.g. application servers) to another network (e.g. database).
• It's recommended to use multiple layers of firewalls for distinctly different functional portions of the network – one for the demilitarized zone (DMZ), a second for the web server, a third for the application server and perhaps a fourth for the database layers.



Chapter-3: Enterprise Management Systems(EMS):

3.1 EMS: Learning objectives

• Describe the purpose of an enterprise system.
• Explain how supply chain management is used.
• Summarize the challenges in supply chain management.
• Describe customer relationship management systems.
• Explain Information Management & Technology of Enterprise software
• Explain role of IS & IT enterprise management
• Discuss Enterprise Engineering

ES - Background

• Around the globe, companies increasingly being more connected
• Interested in reacting instantaneous when
– Customer places an order
– Shipment from the supplier delayed
• Managers
– Wants to know the impact of these events on every part of business &
– How business is performing at any point of time?

Enterprise systems provide integration to make this possible

What is Enterprise management system?

• It is an enterprise wide information system designed to coordinate all the resources, information and activities needed to complete business processes.

3.2 Enterprise Software: ERP/ SCM/ CRM

3.2.1 ERP (Enterprise Resource Planning)

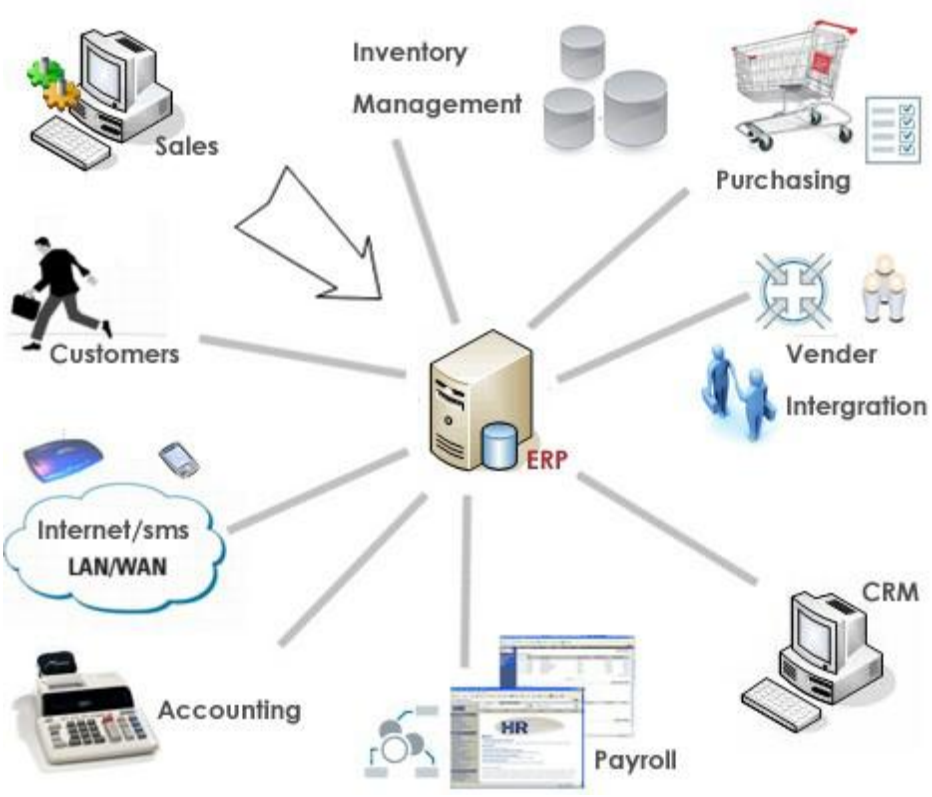


Figure: ERP software

• The focus of ERP is on resource management within constraints to maximize the return on investment.
• The ERP package design is built on the principle of 'Best Practices'.
• When a new information is entered, made immediately available to other business processes
    i. e. ERP - integrates all department functions onto a single system
• Serves the needs of the entire company
• Some of ERP's functions include:
– Bookkeeping & Accounting
– Human Resource Management
– Planning Production
– Supply-Chain Management

History of ERP
• ERP has its roots in the MRP and MRPII systems of the 70's and 80's.
• Came about as companies realized that the management and flow of information was just as important as materials and inventory management.
• ERP has also evolved considerably with computer and technology advances.

ERP models and modules
• Business forecasting, planning and control
• Sales, distribution, invoicing
• Production, planning and control
• Material management
• Finance and accounting
• Personnel management

Sub module:

Business	*Sales*	*Production*	*Materials*	*Finance*	*Personnel*
Forecasting	Forecasting	Planning	Purchase	Accounting	Human Resource
Planning	Planning	Order Control	Inventory	Funds Mgt.	Payroll
Goals	Sales Budget	WIP	Stores	Balance Sheet Processing	Accounting
Objectives	Order Processing	Quality	Valuation	Schedules	Skill Attendance Inventory
Targets Strategy Control	Order Execution Delivery Invoicing	Scheduling Despactch	Analysis Control	Analysis Control	Analysis Control
Fixed Assets	Maintenance	Quality Control	Cost Accounting	Management Accounting	Consolidation of business operations

How the Enterprise Systems work?
• ES consists of a set of interdependent s/w modules to support basic internal business processes mentioned above

• s/w enables data to be used by multiple functions & businesses for precise org's coordination & control
• s/w is built around thousands of business processes
• Org's implementing s/w would 1st select the functions of the system they wish to use &
• Then map their business processes to the predefined business processes available in the s/w
• Companies could use configuration tables provided by the s/w to tailor a particular aspect of the system to the way it does business
• If the ES does not support the way org does business, companies can rewrite some of the s/w
• However, if the company wants to reap the maximum benefit from ES, the company must change the way they work.
• Major ES S/W vendors- SAP, Oracle, PeopleSoft etc.

How should we implement ERP Systems?

• People
– Project Structure
– Should be aligned to processes
• Process
– Implementation Process (outlined in detail)
– Adapt your processes to those of the ERP.
• Technology
– Hardware
– Software
– Integrated Systems

Process

1. Definition and Analysis
• Hold discussions with various functional personnel to establish the actual number of systems operating at client site, what they are used for, why and how often
• Produce the Project Scoping Document outlining current situation, proposed solution and budgeted time
Challenge : **REQUISITE EXPERTISE** - No two clients are the same
2. Design
• Prepare various functional reports - specifies current scenario and wish list
• Prepare Design document which specifies how the system is going to work
• Prepare test scripts to be followed on system testing
• Map out the interface paths to various modules
Challenge : **INFORMATION SHARING** - Availability of staff
3. Build
• Configure system as per set up document specifications i.e. transfer conceptual model into reality
• Test system to verify accuracy (preliminary tests)
Challenge : **TECHNICAL ENVIRONMENT** - System functionality
4. Transition
• Train users on their specific areas
• Assist in test data compilation and system testing by users
• Finalise the Live system and captured opening balances
Challenge : **USER RESISTANCE**
Understanding and acceptance data Preparation
5. Production
• Official hand holding
• Efectiveness assessment
• Business and Technical Direction Recommendations

Technology
• Technology is an enabler, not the driver (it is there to assist the organisation to achieve business goals)
• It is a means to an end, not the end

ERP product characteristics:
• The process design of data, transaction, application or system processing is collaborative and parallel.
• The design reduces processing time of transaction enables faster decision-making and reduces cost of business processing.
• ERP applications in different functional areas are capable of accepting the required input through import/export facility, system interface or direct connectivity.

Features of ERP
• Business rules embedded in process.
• Extensive use of stored procedures, triggers and alerts.
• Stage-wise cost data capture for analysis and decision making.
• Resource planning, scheduling and management.
• Creates knowledge databases using data warehousing and data mining applications.
• Can run in different network environments.
• Use object and component technologies.

Benefits of ERP
• Better management of resources.
• Increasing productivity.
• Customer satisfaction due to shorter delivery.
• Simultaneous activation of the decision centers.
• Business operation transparency.
• Releasing burden on middle level management.
• Different view of business.
• Access of database distributed over the organization.
• Makes management alert at number of points for making decisions.
• Work flow automation.
• Faster communication and direct connection.
• Scalable architecture
• Higher level maintenance.
• Consistency in operations.
• Knowledge driven management.
• Scope can be enlarged through internet/intranet.
• Improvement in quality.
• Decision making tools are user friendly.

ERP product evaluation
• Business scope vs. application scope.
• The degree of deviation from the std. ERP products.
• Ease of use.
• Flexible design.
• Ability to migrate to the ERP environment from present
• Level of intelligent use of 'Help', error messages.
• Versatility of the solution.
• Rating on performance, response and integration.
• Product quality in terms of security, reliability, precision of results.

• Solution architecture and technology.
• Up-gradability.

EMS and MIS

• Today's enterprise has ERP SCM, CRM implementations up and running efectively, an integrated system of ERP, SCM, CRM is called EMS.
• The MIS is required to maximize the process productivity and performance.
• The conventional MIS design is more or less embedded in the ERP solution.
• The ERP through MIS design, improves the decision making skills of individuals very efectively.
• The qualitative change is MIS design due to paradigm shift of traditional business to e-business.

Super structure of EMS

• Data warehousing and data mining
• Executive information system
• OLAP and Query processing
• Decision Support Systems
• Knowledge Management System

3.2.2 Supply Chain Management (SCM)

• Background - today's competitive business environment companies need to focus on how they manage their supply chain
• SCM refers to the close linkage & coordination activities involve in buying, making & moving a product
• SCM integrates business processes for speedy information ,production, fund flow up/down the supply chain to reduce time, redundancy effort removal , reduce cost etc.
• **Supply chain**
– Integrated network
– Consisting of an organization, its suppliers, transportation companies, and brokers
– Used to deliver goods and services to customers
– Exist in both service and manufacturing organizations
SCM: Raw Materials → Supplier → Manufacturing → Distribution →Customer → Consumer

• **Supply chain management (SCM)**
– Process of working with suppliers and other partners in the supply chain to improve procedures for delivering products and services
– Coordinates:
• Procuring materials
• Transforming materials into intermediate and finished products or services
• Distributing finished products or services to customers
– Goods starts from raw materials, move through the logistics & production system until they reach customers
– Returned items flow in the reverse direction from buyer to the seller
• In manufacturing firm, information in an SCM system flows between the following areas:
– Product flow
– Information flow
– Finances flow
• Four key decisions in supply chain management:
– Location
– Inventory
– Production

– Transportation

Supply Chain Processes:
• Many processes & sub processes involve in managing supply chain
• Supply chain council developed the supply chain operations reference model(SCOR)
• SCOR identifies five major supply processes
Plan: Consists of processes that balance aggregate demand & supply to develop a course of action to meet sourcing, production & delivery requirements
Source: Consists of processes that produce goods & services needed to create a specific product or service
Make: Consists of processes that transform the product into finished state to meet planned or actual demand
Deliver: Consists of processes that provide finished goods & services to meet actual or planned demand
Return: Consists of processes associated with returning products including post delivery customer support
Logistics plays an important role i.e. dealing with planning& control of all factors which have high impact on transporting products/service at reduced cost & time

Dell Computer's Supply Chain
• Modified its supply chain from a "push" to a "pull" manufacturing process
– Also known as "built to order (BTO)"
• Main sales channel is direct sales to customers
• Dell has been able to reduce costs by eliminating intermediaries and shortening delivery time

SCM Technologies(Applications)
• Information technologies and the Internet play a major role in implementing an SCM system
• **Electronic Data Interchange (EDI)**
– Enables business partners to send and receive information on business transactions
– Expedites delivering accurate information
– Lowers the cost of transmitting documents
– Advantage of being platform independent and easy to use
• Using EDI has some drawbacks
– Uses the X.25 standard
– Beneficial when more companies are in the EDI network
– Often was not affordable for small suppliers and distributors
• Open EDI
– Based on XML
– Traditional EDI has declined in popularity

Internet-Enabled SCM:
• Improves information sharing throughout the supply chain
• Can improve the following SCM activities:
– Purchasing/procurement
– Inventory management
– Transportation
– Order processing
– Customer service
– Production scheduling

E-marketplaces:
• Third-party exchange

– Provides a platform for buyers and sellers to interact with each other and trade more efficiently online
• Benefits
– Increases efficiency and effectiveness in the supply chain
– Provides opportunities for sellers and buyers to establish new trading partnerships
– Provides a single platform for prices, availability, and stock levels that's accessible to all participants
– Solves time constraint problems for international trade and makes it possible to conduct business around the clock
– Makes it easy to compare prices and products from a single source instead of spending time contacting each seller
– Reduces marketing costs more than traditional sales channels can
• E-distributor
– Marketplace owned and operated by a third party that provides an electronic catalog of product
– Maintenance, repair, and operations (MROs) services
• Includes services from different vendors
• E-distributor coordinates them into one package for customers
– Example of a horizontal market

Online Auctions:
• Bring traditional auctions to customers around the globe
– Make it possible to sell far more goods and services than at a traditional auction
• Brokerage business model
• **Reverse auctions**
– Invite sellers to submit bids for products and services

Collaborative Planning, Forecasting, and Replenishment:
• Coordinate supply chain members through point-of-sale (POS) data sharing and joint planning
• Any data collected is shared with all members of the supply chain
• Coordinating the supply chain can be difficult
• CPFR ensures that inventory and sales data are shared across the supply chain
– So that everyone knows the exact sales and inventory levels
• Costs for each partner are shared or minimized
• Unforeseen problems can crop up

3.2.3 Customer Relationship Management (CRM):
– Consists of the processes a company uses to track and organize its contacts with customers
• Main goal of a CRM system
– Improve services offered to customers
– Use customer contact information for targeted marketing
• Marketing strategies in a CRM system
– Focus on long-term relationships with customers instead of transactions
• Identifies segments of customers
• Improves products and services to meet customers' needs
• Improves customer retention
• Identifies a company's most profitable (and loyal) customers
• Helps organizations make better use of data, information, and knowledge to understand customers
• Gives organizations more complete pictures of their customers
– Integrates demographic and other external data with customers' transaction data to better understand customer behavior
• Pays external agencies for additional data about you that might be public or semiprivate
• With a CRM system, organizations can:
– Provide services and products that meet customers' needs

– Offer better customer service through multiple channels
– Increase cross-selling and up-selling
– Help sales personnel close deals faster by offering data on customers' backgrounds
– Retain existing customers and attract new ones
• CRM systems include:
– Sales automation
– Order processing
– Marketing automation
– Customer support
– Knowledge management
– Personalization technology

CRM Applications
• Time Warner Cable Business Class
– CRM system from Salesforce.com
– Analyze business data, improve the accuracy of forecasts, improve problem solving, and monitor sales and business activities
• Important features of the system include:
– Dashboards, features for "drilling down," Web-based knowledge base for employees and customers, and Web log for sales personnel communication
• Has increased productivity by 10%

• On-premise CRM or Web-based CRM (SaaS)
• Several software packages are available for setting up a CRM system:
– Amdocs CRM, Optima Technologies ExSellence, Infor CRM, SAP mySAP, Oracle PeopleSoft CRM, and Oracle Siebel
• Features:
– Salesforce automation
– eCRM or Web-based CRM
– Survey management
– Automated customer service

• **Personalization**
– Satisfies customers' needs, builds customer relationships, and increases profits
– Designs goods and services that meet customers' preferences better
• **Customization**
– Allows customers to modify the standard offering
• Such as selecting a different home page to be displayed each time you open your Web browser
• Using personalization requires gathering a lot of information about customers' preferences and shopping patterns
• Amazon
– Known for using personalization to recommend products to customers
• Nordstrom.com
– Suggests shoes or a tie that go with the suit or a similar suit in the same category
• Apple iTunes
– Other songs that listeners like you purchased are suggested
• Google account holders
– Personalized search results that are reordered based on their searching histories
• Implement a personalization system
– Internet, databases, data warehouse/data marts, data-mining tools, mobile networks, and collaborative filtering

• **Collaborative filtering (CF)**
– Searches for specific information or patterns, using input from multiple business partners and data sources
• Improve CRM systems by identifying, storing, and disseminating "know-how" -facts about how to perform tasks
• Knowledge is an asset
– Should be shared throughout an organization to generate business intelligence and maintain a competitive advantage in the marketplace
• Knowledge is more than information and data
– Also contextual

3.3 Information Management & Technology of Enterprise Software

Background

• Modern information technologies focuses on integration of information from multiple functional areas
• Based on Enterprise wise service platform - provides great degree of cross functional integration
• Service platform integrates multiple applications from various business sectors
• Example : order - to - cash process
– Customer Relationship Management system handles lead generation, marketing campaign, order entry
– once order received Enterprise s/w prepares manufacturing schedule & verifies parts availability
– Order is then handled by processes for distribution planning , warehousing, order fulfillment & shipping which is the responsibility of SCM
– Enterprise financial application handles billing of order to customer
– If the purchase at some point required customer service CRM system handles
– To accomplish above firms need a business process management plan & application integration software that ties the various pieces together.
• Business process management (BPM) deals with the org's need for business process change to remain competitive
• BMP includes tools for creating models of improved processes that can be translated into s/w systems
• CRM, SCM, ES, home grown legacy system can be used as a building block for BMP
• Existing Enterprise system like SAP has tools for building cross application service named xApps.
• xApps uses web services standards to pull together data from
– the firm's SAP s/w suite,
– from internal legacy systems or
– from external system for use in new business processes that span multiple functions & application areas
• The s/w synchronizes the existing business processes embedded in these system
• SAP(system application & products) now has prepackaged xApps for merger & acquisitions, new product launches & resource & program management
• Besides SAP similar tools from other vendors are Seibel's Universal application network & PeopleSoft's AppConnect etc.
• These new services are then delivered through portals
• Portals provides framework for building new composite services
• Presenting it to users through a web interface which appears to be coming from a single source

3.4 Business and IT

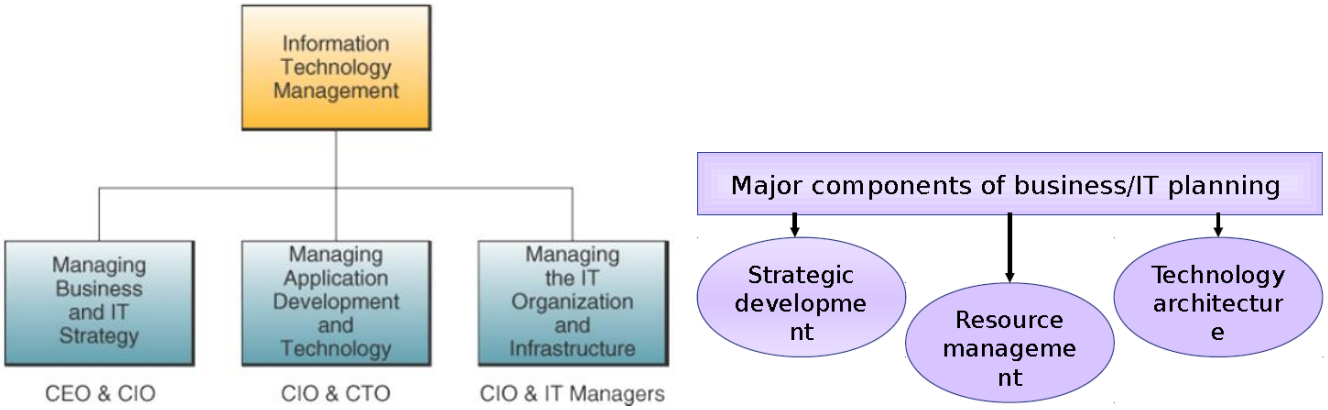
As the 21st century unfolds, many companies are transforming themselves into global powerhouses via major investments in:

- Global e-business

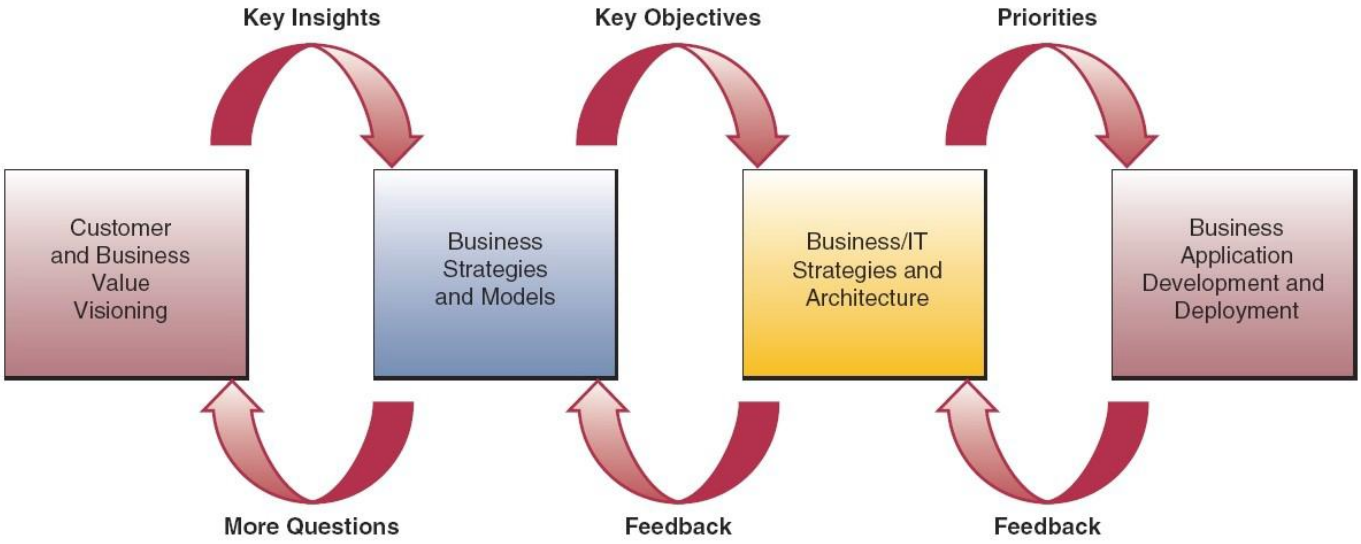
- E-commerce
- Other IT initiatives

Business managers and professionals must know how to manage this vital organizational function

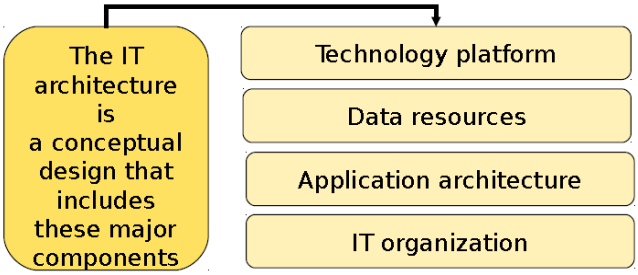
Components of IT Management



The Business/IT Planning Process:



Information Technology Architecture:



Managing the IT Function

Three things recently happened:

- The Internet boom inspired businesses to connect their networks
- Companies have essential applications on their intranets, without which they cannot function

- It became apparent that maintaining PCs on a network is very, very expensive

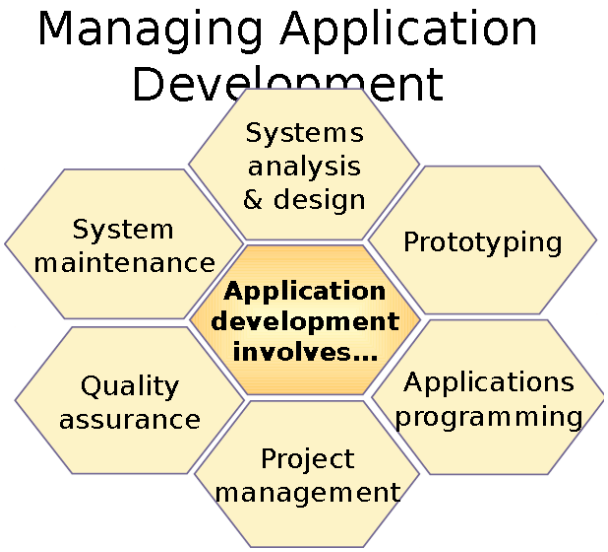
Created an urgent need for centralization

Organizing IT

Early Years: Centralization of computing with large mainframes

Next: Downsizing and moving back to decentralization

Current: Centralized control over the mgmt of IT Serving the strategic needs of business units Hybrid of centralized and decentralized components



Managing IS Operations

• IS operations management

– Concerned with the use of hardware, software, network, and personnel resources in data centers

• Operational activities that must be managed

– Computer system operations

– Network management

– Production control

– Production support

System Performance Monitors:

Software packages that:

- Monitor computer job processing
- Optimize computer system performance
- Facilitate capacity planning and control

Features of System Performance Monitors

Chargeback Systems: Allocates costs to users based on the information service rendered

Process Control Capabilities: Systems that monitor and automatically control computer operations at large data centers

IT Staff Planning:

- Evaluating employees and rewarding good job performance with salary increases, promotions
- Designing career paths
- Setting salary and wage levels
- Recruiting, training, and retaining qualified IS personnel

IT Executives:

Chief Information Officer (CIO):

- Oversees all uses of IT
- Aligns IT with strategic business goals

Chief Technology Officer (CTO):

- In charge of all IT planning/ development
- Manages the IT platform

Other IT Positions

- Systems analyst
- E-commerce architect
- Chief Security Officer
- Technical team leader

Technology Management:

• All IT technologies must be used as a technology platform for integrating business applications
– Both internally or externally focused
– Includes Internet, intranets, electronic commerce, collaboration technologies, CRM software, enterprise resource planning, and supply chain management
• Often the primary responsibility of a chief technology officer

Managing User Services:

• Business units that support and manage end-user and workgroup computing
– Can be done with information centers staffed with user liaison specialists or with web-enabled intranet help desks
• Key roles
– Troubleshooting problems
– Gathering and communicating information
– Coordinating educational efforts
– Helping with end-user application development

Outsourcing: The purchase of goods or services from third-party partners that were previously provided internally

Failures in IT Management:

IT not used effectively: Computerizing traditional business processes instead of developing innovative e-business process
IT not used efficiently: Poor response times, frequent downtime, poorly managed application development

Management Involvement & Governance:

• Managerial and end user involvement

– Key ingredient to high-quality information system performance
– Optimizes business value of IT
• Governance structures
– Steering committees, executive councils
– Encourages active participation in planning and controlling business uses of IT
– Helps avoid post-development problems

IT Governance Approaches:
• Control Objectives for Information and Technology (COBIT)
– Framework for IT management
– Set of generally accepted measures, indicators, processes, and best practices
• Covers four domains
– Planning and organization
– Acquisition and implementation
– Delivery and support
– Monitoring

The International Dimension:
Companies around the world are developing new models to operate competitively in a digital economy:

- These models are structured, yet agile, global, yet local
- Concentrates on maximizing risk adjusted return from both knowledge and technology assets

Global IT Management Dimensions:

- Global IT Management
    - Cultural, Political, and Geo-economic Challenges
        - Business/IT Strategies
        - Business Application Portfolios
        - Internet-Based Technology Platforms
        - Data Resource Management
        - System Development

Global IT Management Challenges:

a. Political challenges

- Many countries regulate or prohibit the transfer of data across their national boundaries
- Others severely restrict, tax, or prohibit imports of hardware and software
- Some have local content laws that specify the portion of the value of a product that must be added in that country if it is to be sold there
- Others require a business to spend part of the revenue they earn in a country in that nation's economy

b. Geo-Economic Challenges:

- Physical distances still a major problem
- Takes too long to fly in specialists
- Hard to communicate in real time across time zones
- Poor telephone and telecommunications services
- May be hard to find skilled local workers
- Differences in the cost of living and labor costs

c. Cultural Challenges:

- Language and cultural interests
- Religions and customs
- Political philosophies
- Cultural training needed before assignments
- Work styles and business relationships

Transnational Strategies: Companies are moving toward a transnational strategy

- Business depends heavily on information systems and Internet technologies to help integrate global business activities
- This requires an integrated and cooperative worldwide IT platform

Global Business Drivers:

• Business requirements caused by the nature of the industry and its competitive or environmental forces
• Examples of global drivers
– Customers
– Products
– Operations
– Resources
– Collaboration

Global IT Platforms:

a. Hardware Difficulties

- High prices
- High tariffs
- Import restrictions
- Long lead times for government approvals
- No local services are spare parts
- Lack of localized documentation

b. Software Difficulties

- Packages developed in Europe may be incompatible with American or Asian Versions
- Software publisher may refuse to supply markets that disregard software licensing and copyright agreements

International Data Communications Issues:

1. Network Management Issues
    a. Improving the operational efficiency of networks
    b. Dealing with different networks
    c. Controlling data communication security
2. Regulatory Issues
    a. Dealing with trans-border data flow restrictions
    b. Managing international telecommunication regulations
    c. Handling international politics
3. Technology Issues
    a. Managing Network Infrastructure across countries
    b. Managing international integration of technologies
4. Country-Oriented Issues

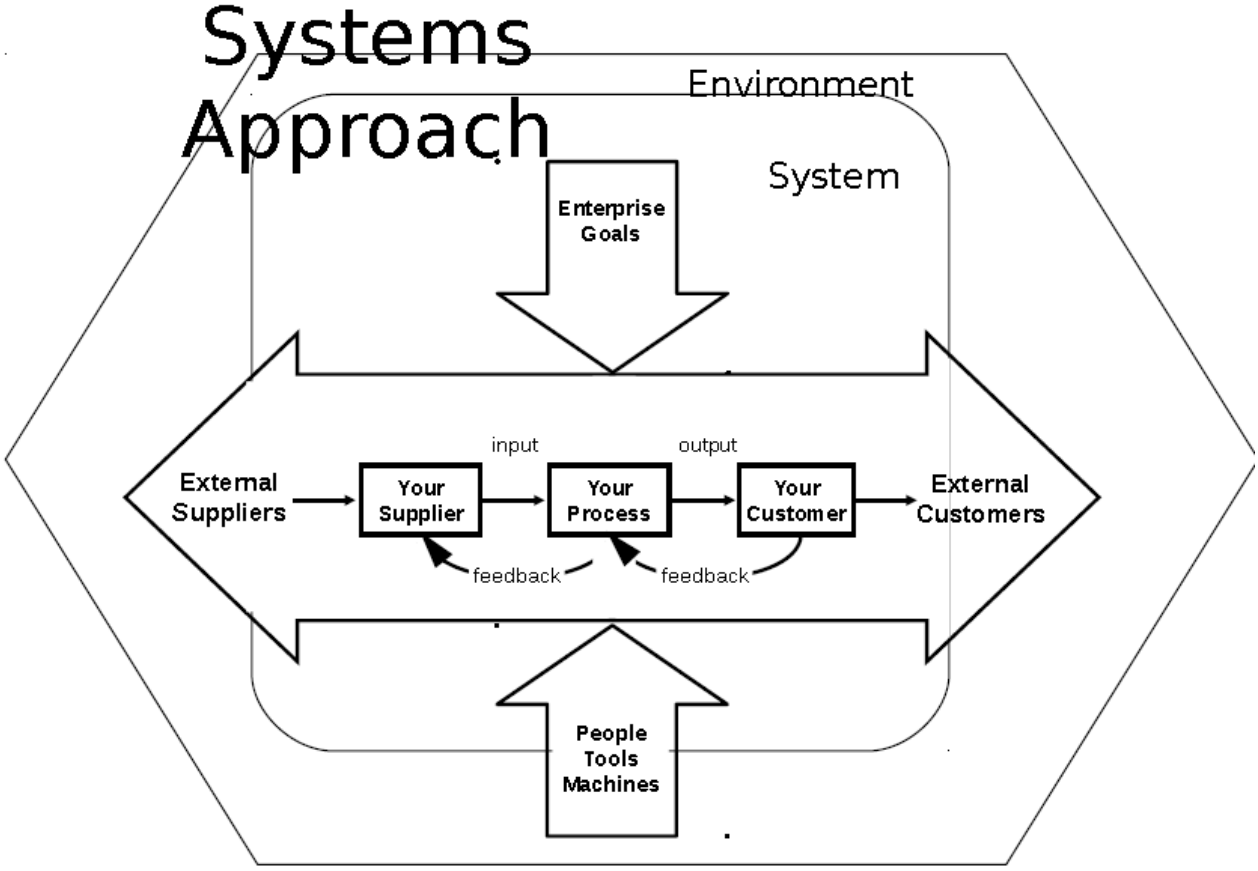
a. Reconciling national differences
b. Dealing with international tariff structures

The Internet as a Global IT Platform

- An interconnected matrix that reaches tens of millions of users in over 100 countries
- Business environment free of traditional boundaries and limits
- Expand markets
- Reduce communications and distribution costs
- Without incurring massive cost outlays for telecommunications, companies
- Internet can improve profit margins

3.5.1 Enterprise Engineering

What is an Enterprise?
An Enterprise is a complex system of cultural, process, and technology components. Enterprise is a system engineered to accomplish organizational goals.



Wrong use of Automation

- System must fit the users and not the reverse?
- Not always, frequently the users must change their ways in order to maximize profits from automation

Redesign, then automate!
• Little change, little payoff
• Big change, big payoff
• A small change with some payoff may mean it is much more difficult to make the right change later.

Key concept: JOINT creativity of business and computer people

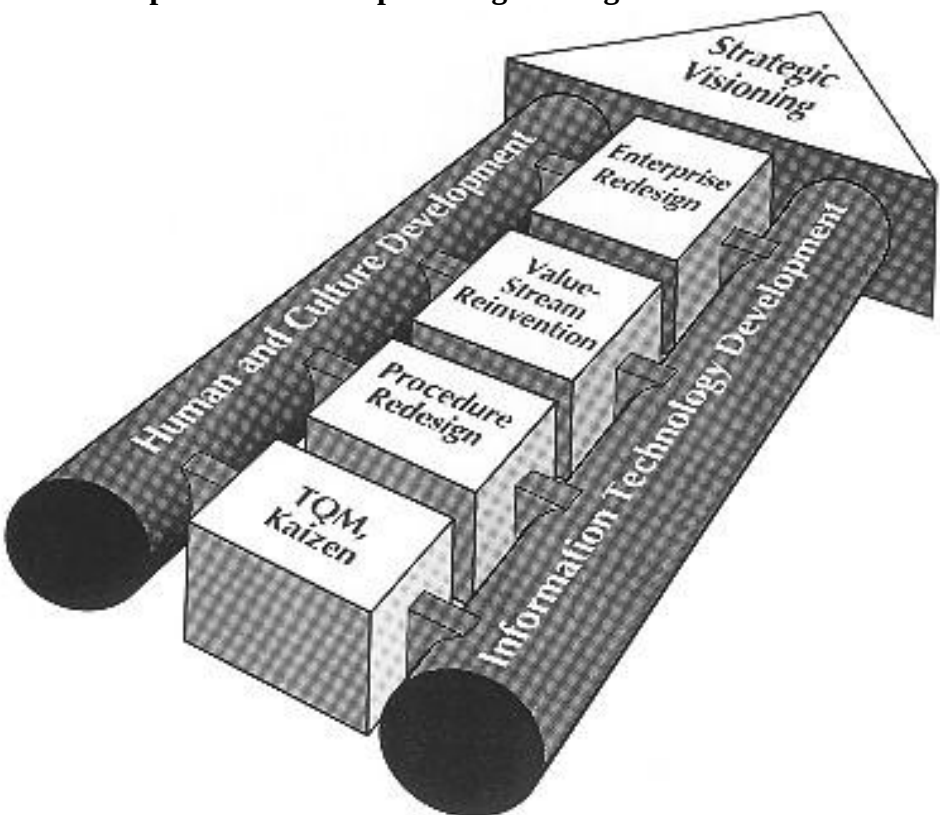
Architecture - Martin
• "The architecture of an enterprise is the basic overall organization within which work takes place.
• Note how this compares with later definitions

According to Martin, Enterprise Engineering is:

• An integrated set of disciplines for building or changing an enterprise, its processes, and systems.
• It integrates the most powerful change methods and makes them succeed.
• The goal is a human-technological partnership of maximum efficiency in which learning takes place at every level.

Goal of the Enterprise Engineer
• "Identify and integrate the most valuable and successful ways to change an enterprise, and to take them into a professional discipline with a teachable methodology and measures of effectiveness."
What do Enterprise Engineers do?
• Identify and Integrate best and most successful ways to change an enterprise
• Two aspects
• Understand new mechanisms
– New ways of organizing work
– New Corporate Architectures must be understood
• Understand methods that can change an enterprise
Two questions Enterprise Engineers always ask:
• What should the enterprise be?
• How do we get there from here?

Seven Components of Enterprise Engineering:



a. TQM, Kaizen

• Continuous change applied across an enterprise
• Kaizen - Japanese term for continuous improvement
• Everybody improves everything all the time

b. Procedure Redesign

• Discontinuous reinvention of existing processes
• Quick hit
• Low lying fruit

c. Value Stream Reinvention

• Discontinuous reinvention of "end to end" streams
• Breakthrough improvement for the CUSTOMER

d. Enterprise Redesign

• Discontinuous redesign
• Holistic change to a new world architecture, sometimes accomplished by building new business units of subsidiaries.

All for changing processes

• Simplifying work
• Improving results

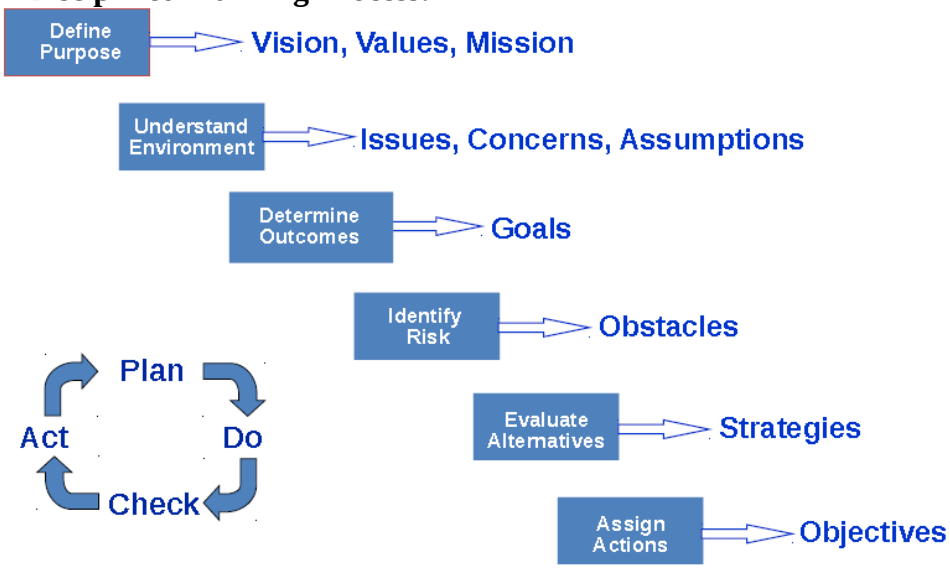
Simplification of Work (note order)

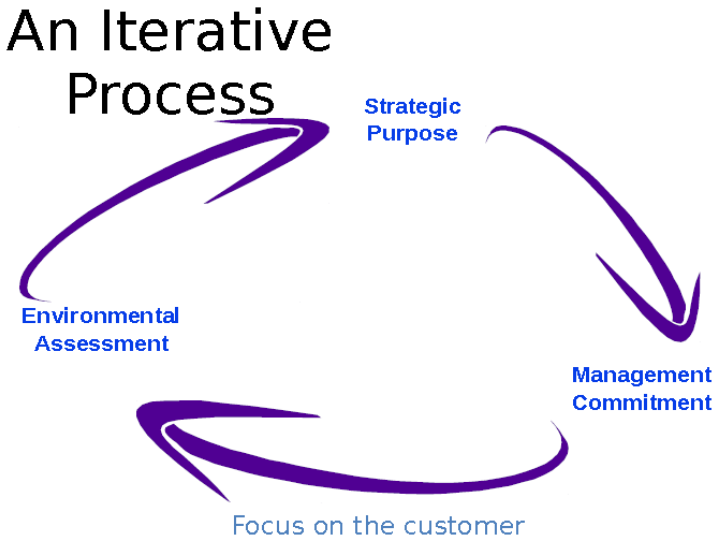
• Eliminate (bureaucracy and non-value added)
• Simplify (work flow, etc.)
• Work Smarter
• Reduce Middlemen (eliminate)
• Refine IIS
• Automate
• Automate Automation

Strategic Visioning: What is a vision?

Resisting the Tide of Change: "Doing your best is not enough."

A Disciplined Planning Process:





World View

• Enterprise Engineering
– Enterprise can be viewed as a complex system
– Enterprise is to be viewed as a system of processes that can be engineered both individually and holistically
– Engineering rigor is required in transforming an enterprise
• Enterprise CAN be engineered

Enterprise Engineering

• Enterprise engineering is a sub discipline of systems engineering,
• It applies the knowledge and methods of systems engineering to the design of businesses.
• Examines each aspect of the enterprise, including business processes, information flows, and organizational structure.
• Enterprise engineering may focus on the design of the enterprise as a whole, or on the design and integration of certain business components.
• Deals with the modeling and integration of various organizational and technical parts of business process and functions"

• Encompassing the application of knowledge, principles, and disciplines related to the analysis, design, implementation and operation of all elements associated with an enterprise
• This is an interdisciplinary field which combines systems engineering and strategic management
• It seeks to engineer the entire enterprise in terms of the products, processes and business operations
• This field is related to engineering management, operations management, service management and systems engineering
• In the context of software development,
– a specific field of enterprise engineering has also appeared that deals with the modeling and integration of various organizational and technical parts of business process and functions
• In the context of information systems development,
– this has become an area of activity for the organization of systems analysis, and an extension to the existing scope of Information Modeling
• It can also be viewed as
– An extension and generalization of the systems analysis and systems design phases of the software development process
• Enterprise modeling can form part of
– The early,
– middle and
– Late
• Enterprise engineering involves
– formal methodologies,
– methods and techniques are designed, tested and used extensively in order to offer organizations reusable business process solutions
• Uses
– Computer Integrated Manufacturing Open Systems Architecture (CIMOSA) methodology
– Integrated DEFinition(IDEF) methodology
– Petri Nets
– Unified Modeling Language(UML) or Unified Enterprise Modeling Language (UEML)
– Enterprise Function Diagrams (EFD)

• **Computer Integrated Manufacturing Open Systems Architecture**
– Provides templates and interconnected modeling constructs to encode business, people and information technology (IT) aspects of enterprise requirements.
– This is done from multiple perspectives: Information view, Function view, Resource view and Organization view

Integrated DEFinition(IDEF) methodology(IDEF)
• first developed as a modeling language to model manufacturing systems, has been used by the U.S. Airforce since 1981 and
• Originally ofered four diferent notations to model an enterprise from a certain viewpoint
• IDEF shows how a business process flows through a variety of decomposed business functions with corresponding information inputs, outputs and actors.
• Like CIMOSA, it also uses diferent enterprise views
• Moreover, IDEF can be easily transformed into UML-diagrams for the further development of IT systems
• These positive characteristics make it a powerful method for the development of Functional Software Architectures

Petri Nets
• Petri Nets are established tools used to model manufacturing systems
• The most advantageous properties are

– the ability to create simple representation of states,
– concurrent system transitions and capabilities
– thereby allowing modeling of the duration of transitions.
• As a result Petri Nets can be used to model certain business processes with corresponding state and transitions or activities therein as well as out
• **Unified Modeling Language (UML)**
• **Enterprise Function Diagrams**
• EFD is a used as a modeling technique for the representation of enterprise functions and corresponding interactions.

3.5.2 Electronic Organism:
• As systems become more complex, the design of these systems must be automated.
• Automation of Automation
• Reaction times shrink, complexity increases, decisions become less intuitive.
• Algorithmically controlled computers *are simple*, because they need not contain the creative infrastructure of the algorithmic division of labor
• but they also *have to be simple*, because otherwise they could not be algorithmically controlled of-line by humans.
• Electronic organisms, in fact all organisms , *have to be complex*, because they have to contain all the creative infrastructure necessary for their creation, reproduction, maintenance and action,
• but they can easily *afford to be complex*, because there is no need for detailed communication with a programmer.
• Electronic Organisms
– Have the ability to react immediately to unforeseen challenges, without the need for a programmer to recognize the situation and
– deal with it by modifying a program.
– Do so by recurrence to fundamental goals and organizing principles, just as programmers do so now
• Electronic organisms will live, grow and evolve in the rapidly growing world of installed computers and networks, just as microbes, plants and animals live in natural ecosystems.
• Strong forces are pushing technology towards electronic organisms.
• Rapidly growing installed processing power worldwide is creating expectations of novel functions of increased complexity.
• Novel computer functions produce a quadratically growing need for coordination.
• Software production is a black hole for human intellectual power, and society will soon no longer be able to satisfy that demand-results in **software crisis**
• The example of animal and human nervous systems and of organisms in general lead the way to the solution of these problems
• A general expectation that autonomous electronic organisms will create security problems is due to a fundamental misunderstanding

What are Business Processes?
• A set of logically related tasks performed to achieve a defined business outcome
• A set of processes forms a business system
• Characteristics of business processes
– Customers - recipients of outcomes
– Cross organizational boundaries

Redesign with IT - Five Steps
• Develop Business Vision and Process Objectives

• Integrated Development (ID) Processes to be redesigned
• Understand and Measure the Existing Process
• ID IT Levers
• Design and Build a Prototype of the New Design

Types of Processes:

Process Dimension & Type	*Typical Example*	*Typical IT Role*
Entities		
Interorganizational	Order from a supplier	Lower transaction costs; eliminate intermediaries
Interfunctional	Develop a new product	Work across geography; greater simultaneity
Interpersonal	Approve a bank loan	Role and task integration
Objects		
Physical	Mfg a product	Increased outcome flexibility
Informational	Create a proposal	
Activities		
Operational	Fill order	Reduce time and costs; increase output quality
Managerial	Develop a budget	Improve analysis; increase participation

Management Issues
• Management Roles - commitment even through across functional boundaries
• Processes and Organization
• Skills - new ones required
• Continual Organization Improvement
• IT Organization in Enterprise may change
• Continuous Process Improvement

3.5.3 Loose Integration vs Full Integration:
• Loose
– simple exchange of info
– no guarantee of same interpretation
– ex. Dedicated interface
• Full
– specificities are known only the one system
– two systems contribute to a common task
– two systems share the definition of items exchanged

3.5.4 Process Alignment:
• The greatest challenge facing businesses today is the unrelenting pace of change.
• Every organization wants to improve its ability to respond rapidly, dynamically and economically to market forces.
• Well-defined processes are the key to becoming a more responsive and adaptable enterprise.

• In a process-centric organization, processes are integrated holistically, are well designed and visible to management, and are measured and managed against corporate strategy and goals.
• Process-centric companies have well defined processes that are horizontally and vertically aligned, are well governed, and produce cost-effective and reliable outcomes.
• Business Process Management (BPM) provides
– a comprehensive system for improving processes,
– aligning business processes with business goals and
– assuring underlying IT applications, human competency and
– organization designs support the process performance objectives.
• Within the BPM there are two constituencies, both using the term "BPM”.“
– One constituency uses the term to refer to changing the way processes are designed, managed and measured.
– The other uses the term "BPM" to refer to IT efforts to create and install BPM systems, including the development of rule-based systems, automated process monitoring systems, and software standards.
• Presently ,IT is developing a new generation of software tools that make it easier for IT to align its applications and databases with business processes
• BPM software products have
– the potential to significantly improve a company’s performance, and
– efforts to organize IT developments in ways that align and support business processes is a major step forward in managing IT
• However, the more important issue is getting business organizations to embrace a process centric approach to management

To be a process centric company
• a company needs to master and integrate all of the process elements within the organization and
• overcome the gap that lies between those interested in the human aspects of process change and
• those interested in the automation of processes

• If your company is to become process-centric you need to develop a central business process architecture that ties corporate strategy to both human and IT implementation.
• Senior managers
– need to invest the time in developing the business process architecture,
– need to monitor the performance of the processes defined in the architecture and
– they need to set priorities and manage the processes and the people engaged in all levels of process change.
• When this occurs you will have built business process right into the fabric of your organization, assuring that your company’s people and processes are aligned to improve organizational performance
• Everyone interested in reaping the multiple proven benefits of a process centric approach to business process change needs to reach out to others and define BPM as broadly as possible.
• To assure this happens we need to work together to assure that BPM embraces all possible process change efforts.

3.5.5 Framework to manage integrated change:
• The alignment between Business Processes (BP) and Information Technologies (IT) is a major issue in most organizations
• It directly impacts on the organization's agility and flexibility to change according to business needs.
• The concepts upon which alignment is perceived are addressed in what is called today the "Enterprise Architecture", gathering business and IT together
• Many Enterprise Architecture Frameworks have been proposed, focusing on different concerns and with different approaches for guiding the development of an IT infrastructure well suited for the organization.

• Each Enterprise Architecture Framework has its own concepts, components, and methodologies to derive the component all the required artifact.
• However, when the main concern is alignment, we may consider simpler architecture concepts and simpler methodologies because the focus is not to define development artifacts but only to check their consistency
• Integration framework composed of a collection of technologies and services which form a middleware to enable integration of systems and applications across the enterprise
• Supply chain management applications (for managing inventory and shipping), customer relationship management applications (for managing current and potential customers), business intelligence applications (for finding patterns from existing data from operations), and other types of applications (for managing data such as human resources data, health care, internal communications, etc.) typically cannot communicate with one another in order to share data or business rules
• For this reason, such applications are sometimes referred to as islands of automation or information silos.
• This lack of communication leads to inefficiencies, wherein identical data are stored in multiple locations, or straightforward processes are unable to be automated
• Enterprise application integration is the process of linking such applications within a single organization together in order to simplify and automate business processes to the greatest extent possible, while at the same time avoiding having to make sweeping changes to the existing applications or data structures.
• In the word EAI is the "unrestricted sharing of data and business processes among any connected application or data sources in the enterprise."

Chapter 4: Decision Support system (DSS)

4.1 DSS, Operations research models

Learning objectives

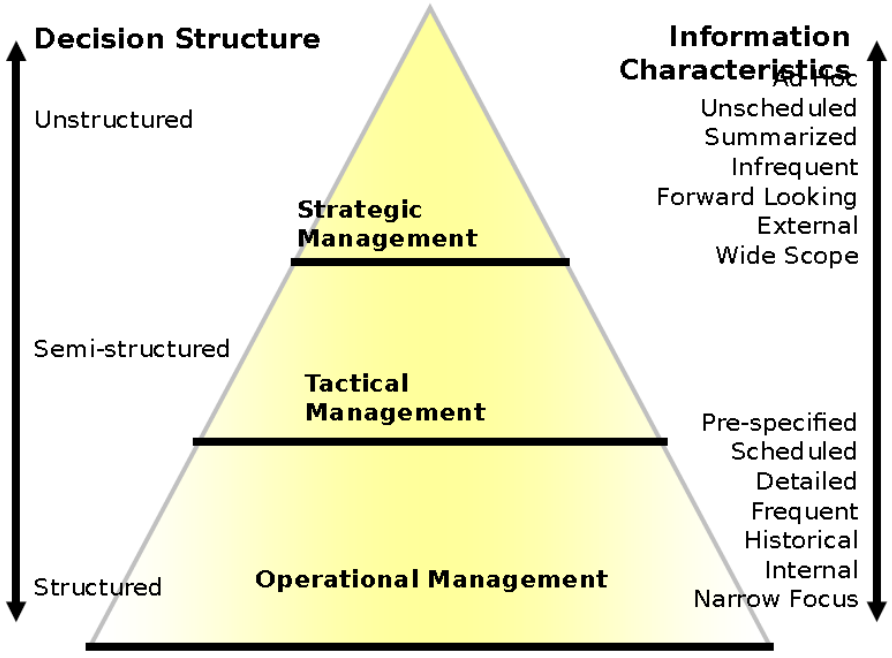
• Assess how information systems support the activities of managers and management decision making.
• Demonstrate how decision-support systems (DSS) differ from MIS and how they provide value to the business.
• Describe components of DSS & GDSS
• Differentiate DSS& GDSS
• Demonstrate how DSS & GDSS helps in decision making
• Discuss Executive support system in the Enterprise

DSS-Background

• Decision makers are faced with increasingly stressful environments– highly competitive, fast-paced, near real-time, overloaded with information, data distributed throughout the enterprise, and multinational in scope.

Types of Decisions (according to Simon)

• *Structured decisions* are repetitive, routine, and involve a definite procedure for handling (e.g., restock inventory)
• *Unstructured decisions* are non routine decisions
• Where decision maker must provide judgment, evaluation, and insights
• There is no agreed-upon procedure for making the decision (e.g., decide on corporate objectives).
• *Semi structured decisions* are ones where only part of the problem has a clear-cut answer
• provided by an acceptable procedure (e.g., develop a marketing plan)



DSS

• Assists the mgmt decision making by – combining data, sophisticated analysis models, tools & user friendly s/w into a single powerful system
• Can support semi structured & unstructured decision making
• DSS provides users with flexible set of tools & capabilities for analyzing important block of data

MIS vs. DSS

MIS	DSS
• Periodic reports	• Special reports that may only be generated once
• Pre-specified, generic reports	• May not know what kind of report to generate until the problem surfaces; specialized reports.

• In a DSS, a manager generates the report through an interactive interface
– Flexible & Adaptable reports
• DSS Reporting is produced through analytical modeling, not just computing an average, or plotting a graph.
– Business Models are programmed into a DSS

Strategic Management:

- The People
  - Board of Directors
  - Chief Executive Officer
  - President
- Decisions
  - Develop Overall Goals
  - Long-term Planning
  - Determine Direction
    - Political
    - Economic
    - Competitive

Tactical Management

- People
  - Business Unit Managers
  - Vice-President to Middle-Manager
- Decisions
  - short-medium range planning
  - schedules
  - budgets
  - policies
  - procedures
  - resource allocation

Operational Management

- People
  - Middle-Managers to
  - Supervisors
  - Self-directed teams
- Decisions
  - short-range planning
  - production schedules
  - day-to-day decisions
  - use of resources
  - enforce polices
  - follow procedures

MIS vs. DSS:

	MIS	DSS
Support	**Info** about performance	**Info and modeling** to analyze problems
Report Form	Periodic reports or On Demand	Interactive Inquiries
Format	Pre-specified Fixed format	Flexible and Adaptable
Processing	Extract and manipulate data	Analytical modeling of data

Differences in System Characteristics:

Dimensions	TPS	MIS	DSS
Type of users	Clerical and supervisory	Middle management	All levels including top mgmt and professionals
Focus	Data transactions	Information	Decision, flexibility
Applications	Payroll, sales data, inventory	Sales forecasting production control	Strategic planning, integrated problems
Ease of use	Low	Moderate	High
Processing interest	Expediency	Efficiency	Effectiveness
Reason for development	Cost saving, customer service	Reporting basic information	Improved decision making

What is Analytical Modeling?

• **Supply Chain Modeling** – Simulate what would happen if you reduced your inventory?
– How many items would go out of stock?
– How much would you save?
– Is reducing inventory a good thing?

More Modeling
Price Point Modeling – model what would happen if you lowered or raised the price of your product
• uses information about your customers income and your competitors prices
• uses well-known supply and demand Models

How is DSS reporting different?
• Modeling helps predict the outcome of a decision.
• This directly helps you make a decision – Possibly an optimal decision
• With a DSS you can explore possible alternatives.

Analytical Modeling is the key:

Type of Modeling	Example
What-if analysis	What if we cut advertising by 10% what would happen to sales?
Sensitivity analysis	Let's cut advertising by 1% repeatedly so we can see its relationship to sales
Goal-seeking analysis	Let's try increasing advertising until sales reach $1 million
Optimization analysis	What level of advertising maximizes our overall profit?

Decision Support System:
• Decision support systems are created to help people make decisions
• provides access to information & analysis tools
• Many stockbrokers now use programs that will automatically put in requests to sell shares once they reach a certain price
• Another example is a simple analytical tool that banks use to help formulate loans for perspective customers.
• DSS depends upon the accuracy of math involved in creating the model &
• The ability of users to accurately interpret the resulting data
• DSS sometimes described as evolutionary step after MIS
• For this description to be valid; MIS must be defined narrowly as automating of routine & structured task to support decision making
• Such systems solves a specific problem or a class of problems such as scheduling, planning, resource allocation & forecasting
• They allow managers to ask adhoc queries & receive customized responses
• Although DSS initially targeted to top managers & middle managers, these systems are deployed at all the levels of the organization
• They support individual decision making to group decision making

DSS characteristics:
• Supports in decision making in both structured and unstructured problem environment
• Supports decision making at all the levels of the organization

• Provides decision support for several interdependent decisions.
• Supports all aspects of decision making process
• DSS are made of people, computers, procedures, databases, interactive query facility & so on
• Thus, they are intended to be evolutionary/adaptive & easy for people to use.

Example - DSS

Typical information that a decision support application might gather and present would be:
• Comparative sales figures between one week and the next
• Projected revenue figures based on new product sales assumptions
• The consequences of different decision alternatives, given past experience in a context that is described

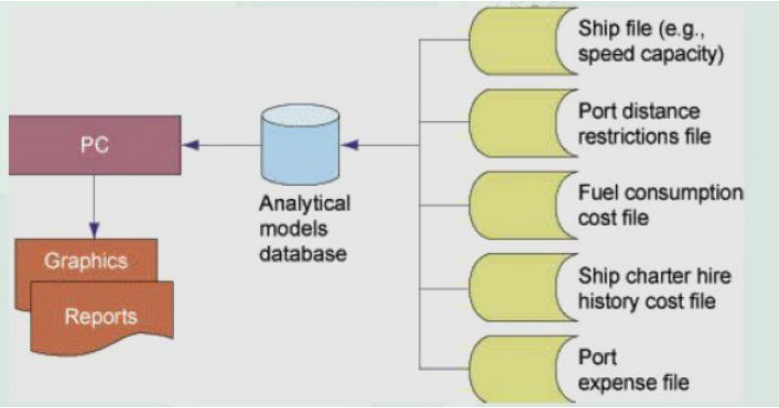


Fig. Example – DSS

Types of DSS: Two basic types: – Model driven DSS and Data driven DSS

What is a Model?

• A mathematical representation that relates variables for solving a decision problem
• Convert the decision problem into a model
• There can be multiple solutions to a model
• Use math techniques solve the model

Model driven DSS

• Analytical capabilities; Can answer 'what-if' scenarios
• Can be used for deciding which path to take (Goal seek)
• Can be used to determine what inputs will get you the desired output (Solving)
• Primarily stand alone system isolated from major organizational Information systems
• Uses some type of model to perform "what-if "analysis & other kinds of analysis
• Such systems developed by end-user divisions or groups not under the control of central IS control
• Analysis capabilities based on the strong theory or model combined with easy to use UI
• Dynamic module adjusts based on changing variables
• Pattern of behavior can become useful model
• Models often based on mathematical Equations

Data-driven DSS

• These systems analyze large pool of data found in major organizational systems
• Supports decision making by allowing users to extract useful information that previously were buried in large quantities of data
• Often data collected from TPS are collected in data warehouses

• Online analysis processing (OLAP) & data mining can then be used to analyze data
• Example: companies now build data driven DSS to mine customer data gathered from their web sites & enterprise systems

DSS – Components:
• DSS Database
• DSS software system
• Model
• User Interface

DSS Database
• A collection of current or historical data from a number of application or groups
• Can be
– a small DB residing in PC or
– massive warehouse

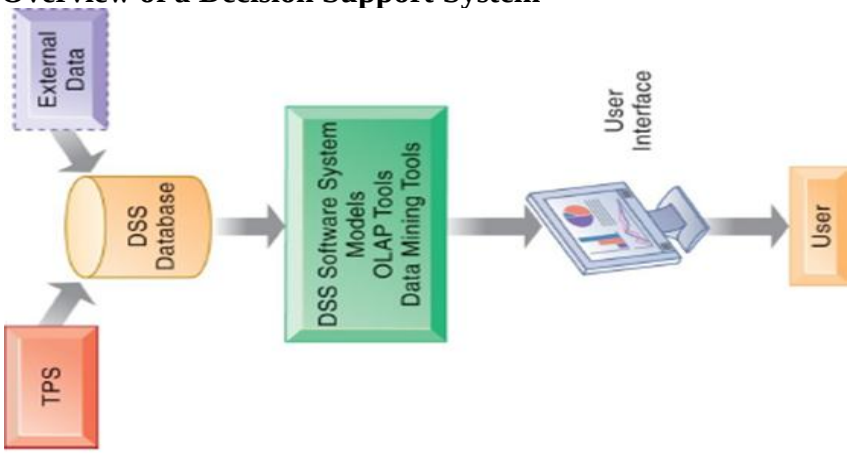
DSS software system
• Contains the software tools that are used for data analysis
• May contain various OLAP tools, data mining tools or a collection of mathematical & analytical models
• Can be made easily accessible to the DSS user

Model

•Abstract representation that illustrates components or relationships of phenomenon;
• may be physical (airplane), mathematical (equation), or verbal model (procedure for writing an order)
• Statistical models
• Optimization models
• Forecasting models

User interface
• DSS user interface controls the interaction between the user and the tools:– graphical, flexible, and easy to use (e.g., Wizards)

Overview of a Decision Support System



Example : DSS
• Statistical modeling s/w can be used to help establish relationships
• Libraries of statistical models
• Contains full range of expected statistical functions including mean, median, deviations & scatter plots

• s/w has the ability to project future outcomes by analyzing the series of data
• Optimization models helps to determine optical resource allocation to maximize/minimize cost & time etc- aims to maximize profit
• Forecasting models

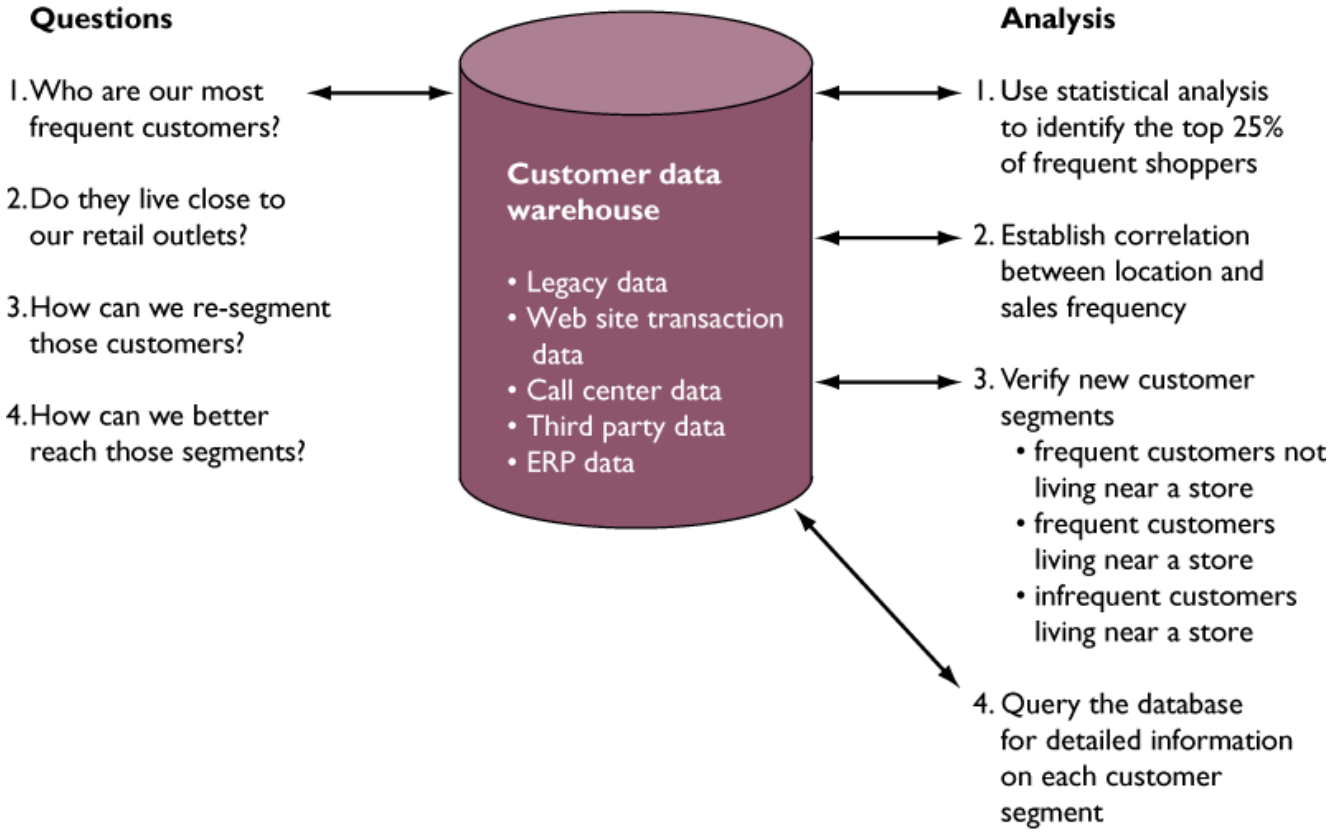
DSS-Perceived benefits: decision quality, improved communication, cost reduction, increased productivity, time savings and improved customer and employee Satisfaction

DSS- Applications:

• Business value of DSS
• Burlington Coat Factory: DSS for pricing
  - DSS manages pricing and inventory nationwide, considering complex interdependencies between initial prices, promotions, markdowns, cross-item pricing effects and item seasonality
• Syngenta: DSS for profitability analysis
  - DSS determines if freight charges, employee sales commissions, currency shifts, and other costs in proposed sale make that sale or product unprofitable
• Compass Bank: DSS for customer relationship management
  - DSS analyzes relationship between checking and savings account activity and default risk to help it minimize default risk in credit card business

DSS for Customer Relationship Management
• Uses data mining to guide decisions
• Consolidates customer information into massive data warehouses
• Uses various analytical tools to slice information into small segments

DSS for Customer Relationship Management:



DSS for Supply Chain Management

• Comprehensive examination of inventory, supplier performance, logistics data
• To help managers search alternatives and decide on the most efficient and cost effective combination
• Reduces overall costs
• Increases speed and accuracy of filling customer orders

DSS other examples……..

• Data visualization tools:
  - Help users see patterns and relationships in large amounts of data that would be difficult to discern if data were presented as traditional lists of text
• Geographic information systems (GIS):
  - Category of DSS that use data visualization technology to analyze and display data in form of digitized maps
  - Used for decisions that require knowledge about geographic distribution of people or other resources, e.g.:
    - Helping local governments calculate emergency response times to natural disasters
    - Help retail chains identify profitable new store locations

Web-based customer decision-support systems (CDSS)

• Support decision-making process of existing or potential customer
• Automobile companies that use CDSS to allow Web site visitors to configure desired car
• Financial services companies with Webbased asset-management tools for customers; Fidelity Investments: customer portfolio allocations, retirement savings plans...
• Home.com: mortgage, rent or buy...

4.2 Group DSS(GDSS):

Need of GDSS

• Early DSS focused largely on supporting individual decision making
• In recent days much work is accomplished in groups within an organization
• System developers & scholars began to focus on how computers can support group & organization decision making.

What Is a GDSS?

• Interactive, computer-based system used to facilitate solution of unstructured problems by set of decision makers working together as group
• Primary focus on communication
• Designed to improve quality and effectiveness of decision-making meetings
• Make meetings more productive by providing tools to facilitate:
• Planning, generating, organizing, and evaluating ideas
• Establishing priorities
•Documenting meeting proceedings for others

Elements that GDSS addresses:

• Improved preplanning
• Increased participation
• Open, collaborative meeting atmosphere
• Criticism free idea generation
• Idea organization & evaluation
• Setting priorities & making decisions
• Documentation of meeting

• Access to external information
• Preservation of "organizational memory"

Components of GDSS:

• **Hardware**
  - Facility: Appropriate facility, furniture, layout
  - Electronic hardware: Audiovisual, computer, networking equipment

• **Software**
  - Electronic questionnaires, electronic brainstorming tools, idea organizers
  - Tools for voting or setting priorities, stakeholder identification and analysis tools, policy formation tools
  - Tools ensure anonymity
  - Group dictionaries

• **People**
  - Participants and trained facilitator, support

Figure : Group System Tools

The sequence of activities and collaborative support tools used in an electronic meeting system facilitate communication among attendees and generate a full record of the meeting.

GDSS software tools:

• Electronic questionnaires
• Electronic brainstorming tools
• Idea organizers
• Questionnaires tools
• Tools for voting & setting priorities
• Stakeholder & analysis tools

• Policy formation tools
• Group dictionaries

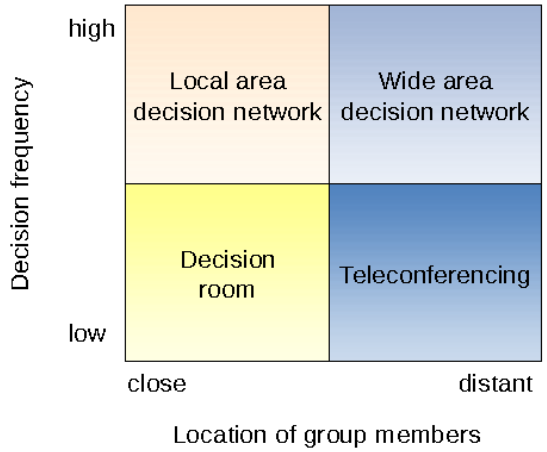
Overview of GDSS meeting:
• **Electronic meeting system (EMS)**-used to make group meetings more productive by facilitating communication & decision making
• Each attendee has workstation, networked to facilitator's workstation and meeting's file server
• Whiteboards on either side of projection screen
• Seating arrangements typically semicircular, tiered
• Facilitator controls use of tools during meeting
• All input saved to server, kept confidential
• After meeting, full record

Business value of GDSS
• Supports greater numbers of attendees
  - Without GDSS, decision-making meeting process breaks down with more than 5 attendees
•More collaborative atmosphere
  - Guarantees anonymity
• Can increase number of ideas generated and quality of decisions made
•Most useful for idea generation, complex problems, large groups
• Successful use of GDSS depends on many factors
  - Facilitator's effectiveness, culture and environment, planning, composition of group, appropriateness of tools selected, etc.
• Group Decision Support System (GDSS) contains most of the elements of DSS plus software to provide effective support in group decision-making settings

Characteristics of a GDSS
• Special design
• Ease of use
• Flexibility
• Decision-making support
• Anonymous (secret) input
• Reduction of negative group behavior
• Parallel communication
• Automated record keeping
• Cost, control, complexity factors

Components of a GDSS Software
• Database
• Model base
• Dialogue manager
• Communication capability
• Special software (also called GroupWare)
– people located around the world work on the same project, documents, and files, efficiently and at the same time

GDSS Alternatives



• **Decision Room**
– For decision makers located in the same geographic area or building
– Use of computing devices, special software, networking capabilities, display equipment, and a session leader
– Collect, coordinate, and feedback organized information to help a group make a decision
– Combines face-to-face verbal interaction with technology-aided formalization

• **Wide Area Decision Network:** Characteristics
– Location of group members is distant
– Decision frequency is high
– Virtual workgroups
    • Groups of workers located around the world working on common problems via a GDSS

4.3 Enterprise and executive decision support systems
Enterprise Decision Support Systems
• DSS to provide enterprise-wide support
• Supports Executives
• Many decision makers in different locations
• Enterprise Resource Planning (ERP) systems
• Solves unstructured problems

Executive support in the enterprise

a. The Role of Executive Support Systems in the Firm

• ESS can bring together data from all parts of the firm and enable managers to select, access, and tailor them as needed.
• Uses easy-to-use desktop analysis tools & online data display
• It tries to avoid the problem of data overload so common in paper reports.
    - Here, data can be filtered & viewed in the graphical format
• The ability to drill down (from summery to detailed) is useful not only to senior executives but also to employees at lower levels of the firm who need to analyze data.
• OLAP tools provide the ability to drill down
• Major challenge is to integrate data from systems designed for very different purposes
• So that senior executives review organizational performance from a firm wise perspective
• Inclusion of modeling and analysis tools usable with a minimum of training

b. Business Value of Executive Support Systems

• Ability to analyze, compare, and highlight trends
• Graphical interface enables users to review data more quickly and with more insight, speeding decision making.
• Timeliness and availability of data enables more timely decision making, helping businesses move toward a "sense-and-respond" strategy.
• Increases upper management span of control, better monitoring

Management opportunities, challenges, and solutions guidelines:
• **Management Opportunities:**
- DSS, GDSS and ESS provide opportunities for increasing precision, accuracy, and speed of decisions made by managers and employees

• **Management Challenges:**
- Building systems that can actually fulfill executive information requirements
- Changing management thinking to make better use of systems for decision support
- Organizational resistance

• **Solutions guidelines:**
- Flexible design and development
- Information requirements are less structured and therefore require more user involvement during development
- The system must be flexible, easy to use, and capable of supporting alternative decision options
- Training and management support
- User training, involvement, and experience; top management support; and length of use are the most important factors in the success of management support systems

Executives' Role and Their Information Needs:
• Decisional Executive Role (2 Phases)
1. Identification of problems and/or opportunities
2. The decision of what to do about them

• Flow chart and information flow
• Use phases to determine executives' information needs

Future of Executive and Enterprise Support Systems
• Toolbox for customized systems
• Multimedia support
• Better access (via PDFs and cell phones)
• Virtual Reality and 3-D Image Displays
• Merging of analytical systems (OLAP / multidimensional analysis)) with desktop publishing
• Client/server architecture
• Web-enabled EIS
• Automated support and intelligent assistance
• Integration of EIS and Group Support Systems
• Global EIS
• Integration and deployment with ERP products

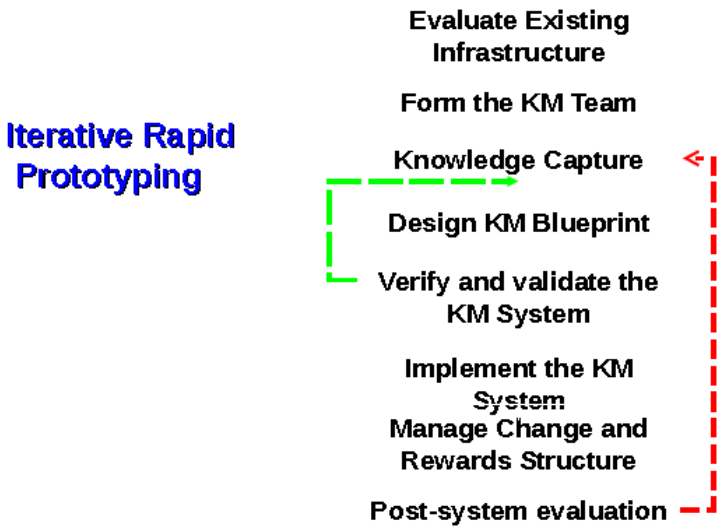
4.4 Knowledge Management, Knowledge based Expert System

Objective
• To access the **importance of knowledge management** in an organization
• Describe the applications that are **more useful for distributing, creating & sharing knowledge** in the firm

• Evaluate the **role of AI in knowledge management**
• Demonstrate **how organizations can use expert systems** & case based reasoning to capture knowledge
• Explain how organizations can **use neural networks & other intelligent techniques to improve knowledge base**

Knowledge Management(KM)- Introduction
• Refers to – The set of process developed in an org to create, gather, store & maintain & disseminate firm's knowledge
• Information Technology plays a major role - enabler for better business process
• Developing procedures & routines in business process to **optimize the creation, flow, learning, protection & sharing of knowledge** in the **firm- core management responsibility**
• **Responsible** head for knowledge management program- **Chief Knowledge officer**
• Producing unique product/service at lower cost based on superior knowledge (production process & design)
• Firm increasingly depends on digital technology these days
• Knowledge is a central productive & strategic asset,
• Organizational success depends on firm's ability to produce, gather, store & disseminate knowledge

Stages of KM System Life Cycle
KM system development life cycle is largely composed



(1) Evaluate Existing Infrastructure:
• The first stage is to evaluate existing infrastructure
• Where several questions that you need to ask for justification
System justifications:
• What knowledge will be lost through retirement, transfer, or departure to other firms?
• Is the proposed KM system needed in several locations?
• Are experts available and willing to help in building a KM system?
• Does the problem in question require years of experience and tacit reasoning to solve?

(2) Form the KM Team
• After carefully evaluating the infrastructure, the next stage –form the KM team to develop KMS
• The team success will depend on a number of factors, including those shown here.
• Identify the ***key stakeholders*** of the prospective KM system.
• Team success depends on:

- Ability of team members
- Team size
- Complexity of the project
- Leadership and team motivation
- Not promising more than can be realistically delivered

(3) Knowledge Capture

• The next stage after forming the KM team –knowledge capture
• **Explicit** knowledge captured in repositories from various media
• **Tacit** knowledge captured from company experts using various tools and methodologies
• Knowledge developers capture knowledge from experts in order to build the **knowledge base**

How do firms obtain knowledge?

• Through variety of organizational learning mechanisms

(4) Design the KM Blueprint

• Next, to develop the KM blueprint based on the knowledge captured
The KM blueprint addresses several issues:
• Finalize scope of proposed KM system with realized net benefits
• Decide on required system components
• Develop the key layers of the KM software architecture to meet company requirements
• System interoperability and scalability with existing company IT infrastructure

(5)Testing the KM System

• As the KM system is under development, it goes through a repetitive iteration of verification and validation
• Verification (functionality) procedure: ensures that the system has the right functions
• *Validation* (integrity) procedure: ensures that the system has the right output
• Validation of KM systems is not foolproof

(6) Implement the KM System

• Finally, there will be the time when the KMS will be rolled out for users to use.
• Converting a new KM system into actual operation
• includes conversion of data or files
• also includes user training
• **Quality assurance is important**, which includes checking for:
– Reasoning errors
– Ambiguity
– Incompleteness
– False representation (false positive and false negative)

- False +ve:-Eg: A perfectly legitimate transaction could trigger IDS to believe that an attack was in progress.
- False -ve:-Eg: an attack takes place and the IDS doesn't detect it

(7) Manage Change and Rewards Structure

• Finally, during implementation, encounter resistance from people
• because - fear of losing control.
• Goal is to minimize resistance to change
• Resistances via projection, avoidance, or aggression

(8) Post-system Evaluation

• Assess system impact in terms of effects on:
– People
– Procedures
– Performance of the business
• Areas of concern:
– Quality of decision making
– Attitude of end users
– Costs of Knowledge processing and update

Components of Knowledge Management Systems

• Technologies

– Communication
- Access knowledge
- Communicates with others

– Collaboration
- Perform groupwork
- Synchronous or asynchronous
- Same place/different place

– Storage and retrieval
- Capture, storing, retrieval, and management of both explicit and tacit knowledge through collaborative systems

• Supporting technologies

– **Artificial intelligence**
- Expert systems, neural networks, fuzzy logic, intelligent agents

– **Intelligent agents**
- Systems that learn how users work and provide assistance

– **Knowledge discovery in databases**
- Process used to search for and extract information
  - Internal = data and document mining
  - External = model marts and model warehouses

– **XML**
- Extensible Markup Language
- Enables standardized representations of data
- Better collaboration and communication through portals

Knowledge Management System Implementation

• Challenge to identify and integrate components
– Early systems developed with networks, groupware, databases
• Knowledge ware
– Technology tools that support knowledge management
- Collaborative computing tools
  - Groupware
- Knowledge servers
- Enterprise knowledge portals
- Document management systems
  - Content management systems
- Knowledge harvesting tools
- Search engines

- Knowledge management suites
  - Complete out-of-the-box solutions

• Implementation
– Software packages available
- Include one or more tools
– Consulting firms
– Outsourcing
- Application Service Providers

Knowledge Management System Integration
• Database– Knowledge discovery in databases
• CRM– Provide tacit knowledge to users
• Supply chain management systems– Can access combined tacit and explicit knowledge
• Corporate intranets and extranets
– Knowledge flows more freely in both directions
– Capture knowledge directly with little user involvement
– Deliver knowledge when system thinks it is needed

Human Resources
• Chief knowledge officer
– Senior level
– Sets strategic priorities
– Defines area of knowledge based on organization mission and goals
– Creates infrastructure
– Identifies knowledge champions
– Manages content produced by groups
– Adds to knowledge base
• CEO– Champion knowledge management
• Upper management– Ensures availability of resources to CKO
• Communities of practice
• Knowledge management system developers– Team members that develop system
• Knowledge management system staff– Catalog and manage knowledge

Knowledge Management Valuation
• Asset-based approaches– Identifies intellectual assets and Focuses on increasing value
• Knowledge linked to applications and business benefits approaches
– Balanced scorecard
– Economic value added
– Inclusive valuation methodology
– Return on management ratio
– Knowledge capital measure
• Estimated sale price approach

Factors Leading to Success and Failure of Systems
• Success
– Companies must assess need
– System needs technical and organizational infrastructure to build on
– System must have economic value to organization
– Senior management support

– Organization needs multiple channels for knowledge transfer
– Appropriate organizational culture
• Failure
– System does not meet organization's needs
– Lack of commitment
– No incentive to use system
– Lack of integration

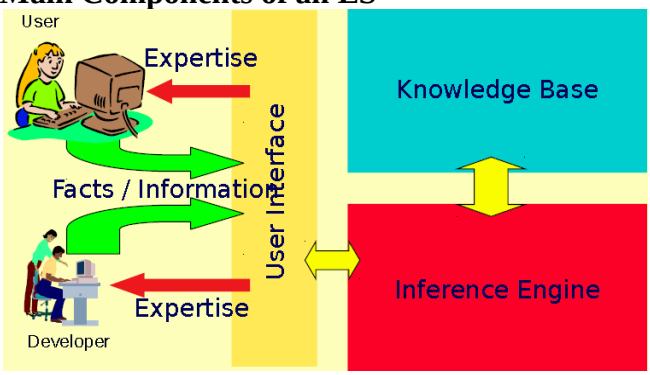
Knowledge based expert System

What is an Expert System (ES)?
• relies on internally represented knowledge to perform tasks
• utilizes reasoning methods to derive appropriate new knowledge
• usually restricted to a specific *problem domain*
• some systems try to capture commonsense knowledge
– General Problem Solver (Newell, Shaw, Simon), Cyc (Lenat)

Definitions "Expert System"
• A computer system that emulates the decision-making ability of a human expert in a restricted domain
• Edward Feigenbaum – "An intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solutions."
• The term *knowledge-based system* is often used synonymously

Expert Systems
• A computer system that uses representation of human expertise in a specialized domain to perform functions similar to those normally performed by human expert
• Well designed systems initiate the reasoning processes experts use to solve specific problems
• Such systems can be used by experts as knowledge assistants
• ES are used to propagate scarce knowledge resources for improved, consistent results.
• Ultimately such system can perform better than any single human in making judgments in a specific usually narrow, area of expertise.
• **Expert system (ES)**: emulates knowledge of human expert
– Solves problems
– Makes decision
– Makes expertise available to novices
– Scope is limited to expert's experience
– Part of **artificial intelligence (AI)** Research

Main Components of an ES



• **Knowledge base**: Used by ES
– Collection of facts and relationships among them
– Built on series of IF-THEN rules
• **Inference engine**: combines data input by user with data relationships
• **Neural networks**: Used by ES to mimic human brain learning
– Refines itself based on decision success rate
– Useful for detecting fraud

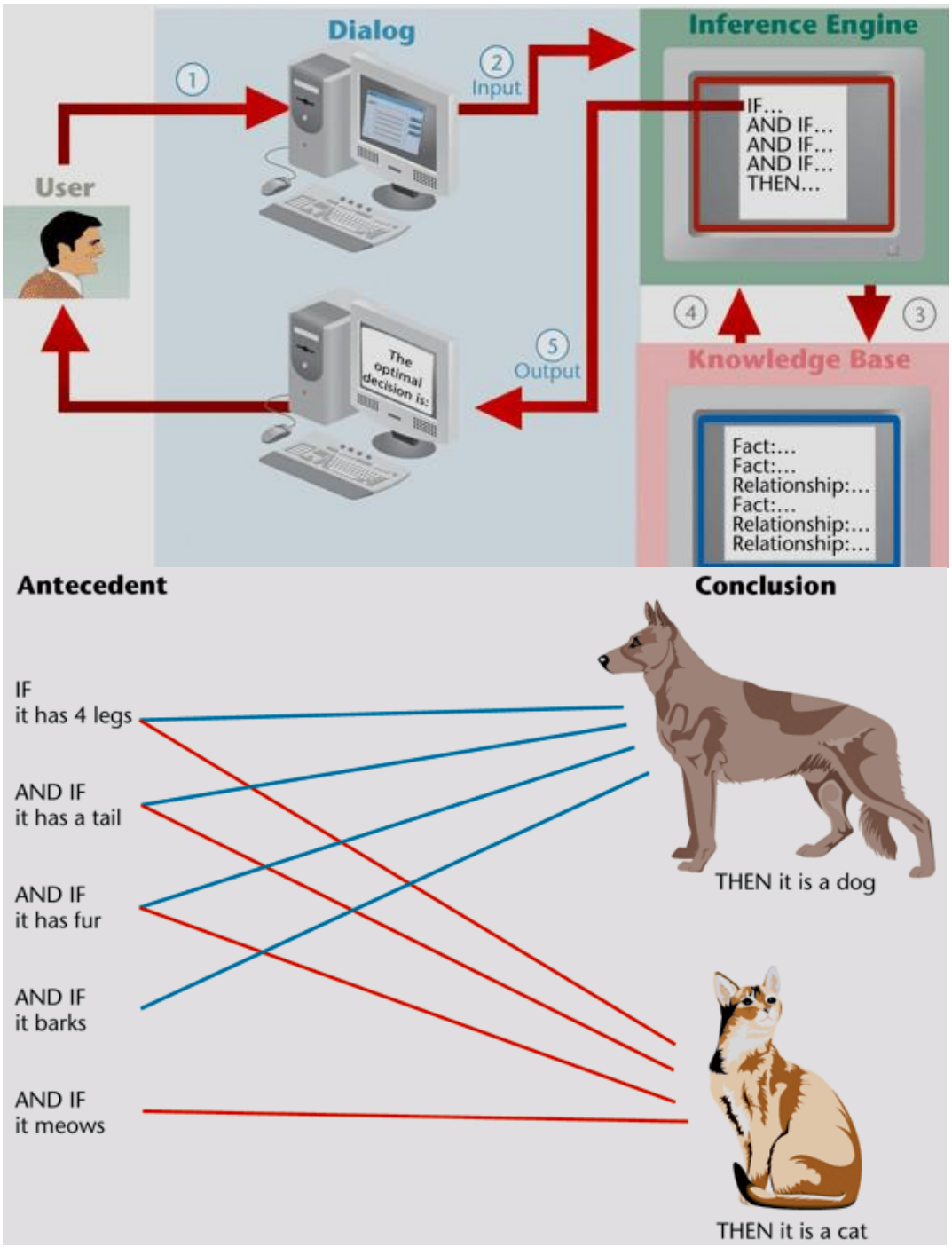


Figure : In IF-THEN rules, different combinations of conditions lead to different conclusions

Expert Systems in Action
• ESs help many industries

– Medical diagnosis
- Help doctors with treatment advice
- Diagnosis of symptoms
- Can detect disease

– Medical management
- Discern which treatment patient should receive
- Administrative decisions

– Telephone network maintenance
- Diagnose and fix network failures

Fig. Medical expert systems help diagnose patients and suggest treatments.

• ESs help many industries

– Credit evaluation
- Approve credit card charges
- Analyzes credit
- Based on financial reports
- Used by loan officers

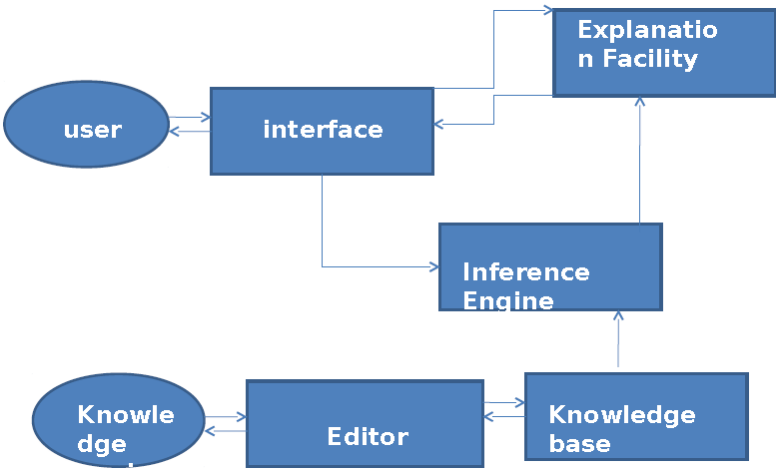
– Detection of insider securities trading
- Prevent trading of stocks based on private information
- Analyze history

– Detection of common metals
- Identify common metals and alloys
- Based on results of chemical tests

– Irrigation and pest management
- Indicate crop market value
- Detect toxins

– Diagnosis and prediction of mechanical failure
- Diagnose cause of failure

Architecture of Typical ES:



Components of ES:

1. Knowledge acquisition subsystem

– The accumulation, transfer & transformation of problem solving expertise from experts or document knowledge sources to a computer program for constructing & expanding the knowledge base

– Subparts

- Knowledge base
- Inference engine (brain control structure)
- User interface

2. Blackboard (workplace)– Working memory area that records plan(how to attach a problem), agenda (potential action awaiting execution) & solutions.
3. Explanation System (justifier)– Trace responsibilities for conclusions & explain ES behavior
4. Knowledge refining system– Explanation & feedbacks
5. Users– Non expert human who needs advice & training

Benefits of ES

1. Increased o/p & productivity
2. Decreased decision making time
3. Increased process & product quality
4. Reduce downtime
5. Capture of scarce expertise
6. Flexibility
7. Easier equipment operation
8. Elimination of need for expensive equipment
9. Operation in hazardous environment.
10. Ability to work with incomplete & uncertain information
11. Knowledge transfer to remote locations
12. Enhancement of other IS

The Key to ES Success

• convincing ideas– rules, cognitive models
• Practical applications– medicine, computer technology
• Separation of knowledge and inference (deduce)– expert system *shell*
• allows the re-use of the "machinery" for different domains
• Concentration on domain knowledge– general reasoning is too complicated

When (Not) to Use ESs

• expert systems are not suitable for all types of domains and tasks
– conventional algorithms are known and efficient
– the main challenge is computation, not knowledge
– knowledge cannot be captured easily
– users may be reluctant to apply an expert system to a critical task

ES Tools

• ES languages
– higher-level languages specifically designed for knowledge representation and reasoning
– SAIL, KRL, KQML, DAML, OWL (Web Ontology Language)
• ES shells
– an ES development tool/environment where the user provides the knowledge base
– CLIPS, JESS, Mycin, Babylon, G2 etc.

ES Advantages

• economical: lower cost per user
• availability: accessible anytime, almost anywhere
• response time: often faster than human experts
• reliability: can be greater than that of human experts, no distraction, fatigue(tiredness), emotional involvement
• explanation: reasoning steps that lead to a particular conclusion
• intellectual property: can't walk out of the door

ES Problems

• limited knowledge: "shallow" knowledge
• no "deep" understanding of the concepts and their relationships
– no "common-sense" knowledge
– no knowledge from possibly relevant related domains
– "closed world"
• the ES knows only what it has been explicitly "told"
• it doesn't know what it doesn't know
• mechanical reasoning
– may not have or select the most appropriate method for a particular problem
– some "easy" problems are computationally very expensive
• lack of trust: users may not want to leave critical decisions to machines

Artificial Intelligence

• An effort to develop computer based systems that behaves as humans
• Systems would be able to learn
– Natural languages
– Accomplish coordinated physical task

4.5 Data Warehousing and OLAP Technology:

1. What is a data warehouse?
2. A multi-dimensional data model
3. Data warehouse architecture
4. Data warehouse implementation
5. From data warehousing to data mining

Data Mining & Warehousing

Data, Data everywhere yet ...

- I can't find the data I need
  - data is scattered over the network
  - many versions, subtle differences
- I can't get the data I need
  - need an expert to get the data
- I can't understand the data I found
  - available data poorly documented
- I can't use the data I found
  - results are unexpected
  - data needs to be transformed from one form to other

1. What is a Data Warehouse?

A single, complete and consistent store of data obtained from a variety of different sources made available to end users in a what they can understand and use in a business context.

• Defined in many different ways, but not rigorously.

– A decision support database that is maintained separately from the organization's operational database

– Support information processing by providing a solid platform of consolidated, historical data for analysis.

• "A data warehouse is a subject-oriented, integrated, time-variant, and nonvolatile collection of data in support of management's decision-making process."

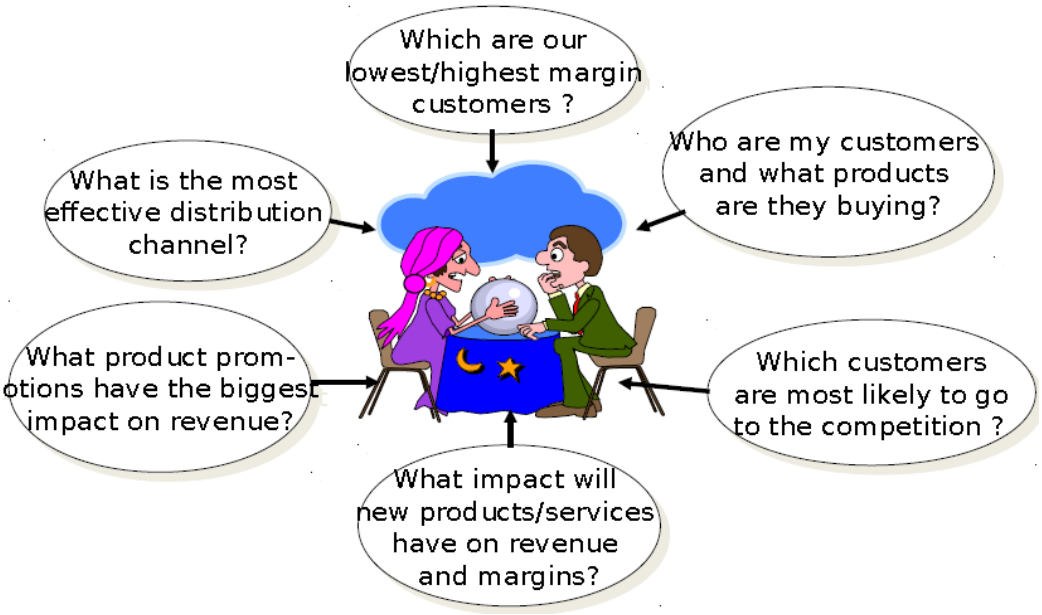


Fig. Why Data Warehousing?

Data Warehouse—Subject-Oriented

• Organized around major subjects, such as customer, product, sales

• Focusing on the modeling and analysis of data for decision makers, not on daily operations or transaction processing

• Provide a simple and concise view around particular subject issues by excluding data that are not useful in the decision support process

Data Warehouse—Integrated

• Constructed by integrating multiple, heterogeneous data sources

– relational databases, flat files, on-line transaction records

• Data cleaning and data integration techniques are applied.
– Ensure consistency in naming conventions, encoding structures, attribute measures, etc. among different data sources
• E.g., Hotel price: currency, tax, breakfast covered, etc.
– When data is moved to the warehouse, it is converted.

Data Warehouse—Time Variant
• The time horizon for the data warehouse is significantly longer than that of operational systems
– Operational database: current value data
– Data warehouse data: provide information from a historical perspective (e.g., past 5-10 years)
• Every key structure in the data warehouse
– Contains an element of time, explicitly or implicitly
– But the key of operational data may or may not contain "time element"

Data Warehouse—Nonvolatile
• A physically separate store of data transformed from the operational environment
• Operational update of data does not occur in the data warehouse environment
– Does not require transaction processing, recovery, and concurrency control mechanisms
– Requires only two operations in data accessing: *initial loading of data* and *access of data*

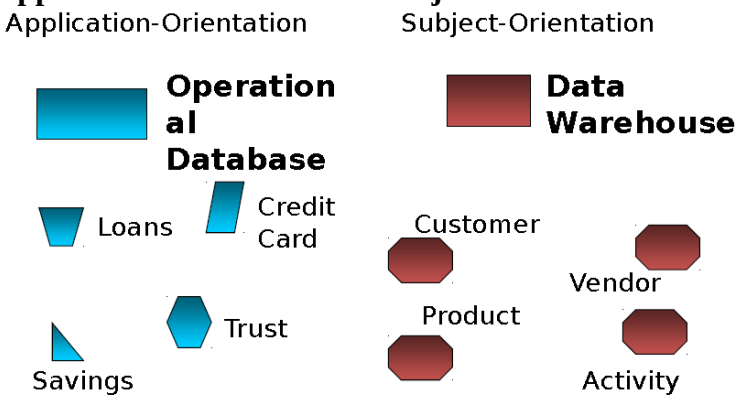
Data Warehouse vs. Heterogeneous DBMS
• Traditional heterogeneous DB integration: A query driven approach
– Build wrappers/mediators on top of heterogeneous databases
– When a query is posed to a client site, a meta-dictionary is used to translate the query into queries appropriate for individual heterogeneous sites involved, and the results are integrated into a global answer set
– Complex information filtering, compete for resources
• Data warehouse: update-driven, high performance
– Information from heterogeneous sources is integrated in advance and stored in warehouses for direct query and analysis

Data Warehouse vs. Operational DBMS
• OLTP (on-line transaction processing)
– Major task of traditional relational DBMS
– Day-to-day operations: purchasing, inventory, banking, manufacturing, payroll, registration, accounting, etc.
• OLAP (on-line analytical processing)
– Major task of data warehouse system
– Data analysis and decision making
• Distinct features (OLTP vs. OLAP):
– User and system orientation: customer vs. market
– Data contents: current, detailed vs. historical, consolidated
– Database design: ER + application vs. star + subject
– View: current, local vs. evolutionary, integrated
– Access patterns: update vs. read-only but complex queries

So, what's diffffferent? OLTP vs. OLAP

	OLTP	OLAP
users	clerk, IT professional	knowledge worker
function	day to day operations	decision support
DB design	application-oriented	subject-oriented
data	current, up-to-date detailed, flat relational isolated	historical, summarized, multidimensional integrated, consolidated
usage	repetitive	ad-hoc
access	read/write index/hash on prim. key	lots of scans
unit of work	short, simple transaction	complex query
# records accessed	tens	millions
#users	thousands	hundreds
DB size	100MB-GB	100GB-TB
metric	transaction throughput	query throughput, response

Application-Orientation vs. Subject-Orientation



Why Separate Data Warehouse?

• High performance for both systems
– DBMS— tuned for OLTP: access methods, indexing, concurrency control, recovery
– Warehouse—tuned for OLAP: complex OLAP queries, multidimensional view, consolidation
• Different functions and different data:
– missing data: Decision support requires historical data which operational DBs do not typically maintain
– data consolidation: DS requires consolidation (aggregation, summarization) of data from heterogeneous sources
– data quality: different sources typically use inconsistent data representations, codes and formats which have to be reconciled

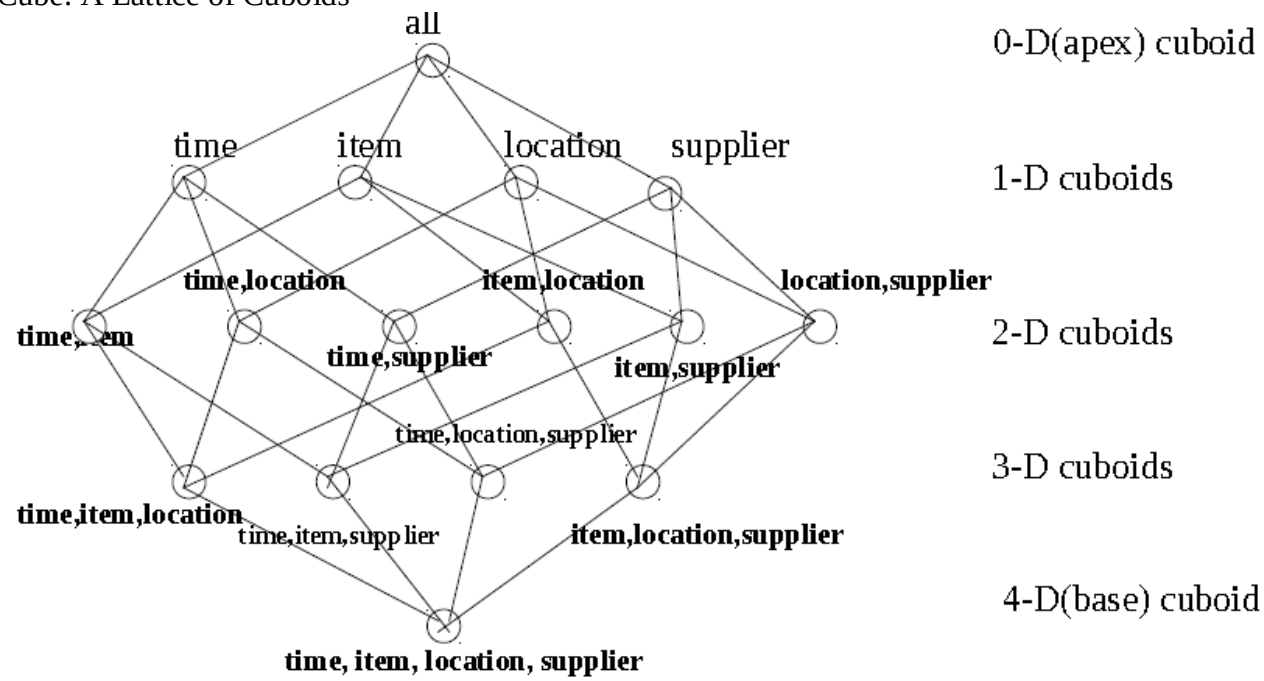
OLTP Systems are used to "run" a business but The Data Warehouse helps to "optimize" the business

2. A multi-dimensional data model

From Tables and Spreadsheets to Data Cubes

• A data warehouse is based on a multidimensional data model which views data in the form of a data cube
• A data cube, such as sales, allows data to be modeled and viewed in multiple dimensions
– Dimension tables, such as item (item_name, brand, type), or time(day, week, month, quarter, year)
– Fact table contains measures (such as dollars_sold) and keys to each of the related dimension tables
• In data warehousing literature, an n-D base cube is called a base cuboid. The top most 0-D cuboid, which holds the highest-level of summarization, is called the apex cuboid. The lattice(network) of cuboids forms a data cube.

Cube: A Lattice of Cuboids



Conceptual Modeling of Data Warehouses
• Modeling data warehouses: dimensions & measures
– Star schema: A fact table in the middle connected to a set of dimension tables
– Snowflake schema: A refinement of star schema where some dimensional hierarchy is normalized into a set of smaller dimension tables, forming a shape similar to snowflake
– Fact constellations: A schema which contains multiple fact tables shares dimensions. It is a collection of star schemas which shares their dimension. So it is also called as a galaxy schema.

Example of Star Schema
• The most common modeling paradigm is the star schema, in which the data warehouse contains
(1) a large central table (fact table) containing the bulk of the data, with no redundancy
(2) a set of smaller attendant tables (dimension tables), one for each dimension.
• **Example** Sales are considered along four dimensions, namely, *time*, *item*, *branch*, and *location*. The schema contains a central fact table for *sales* that contains keys to each of the four dimensions, along with two measures: *dollars sold* and *units sold*. To minimize the size of the fact table, dimension identifiers (such as *time key* and *item key*) are system-generated identifiers.
• Notice that in the star schema, each dimension is represented by only one table, and each table contains a set of attributes.

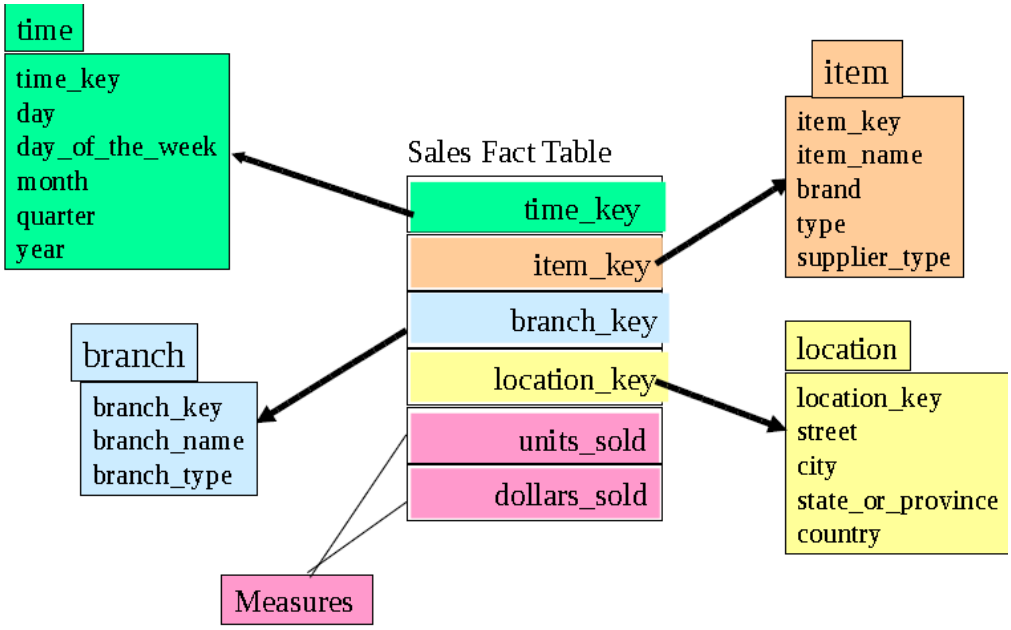
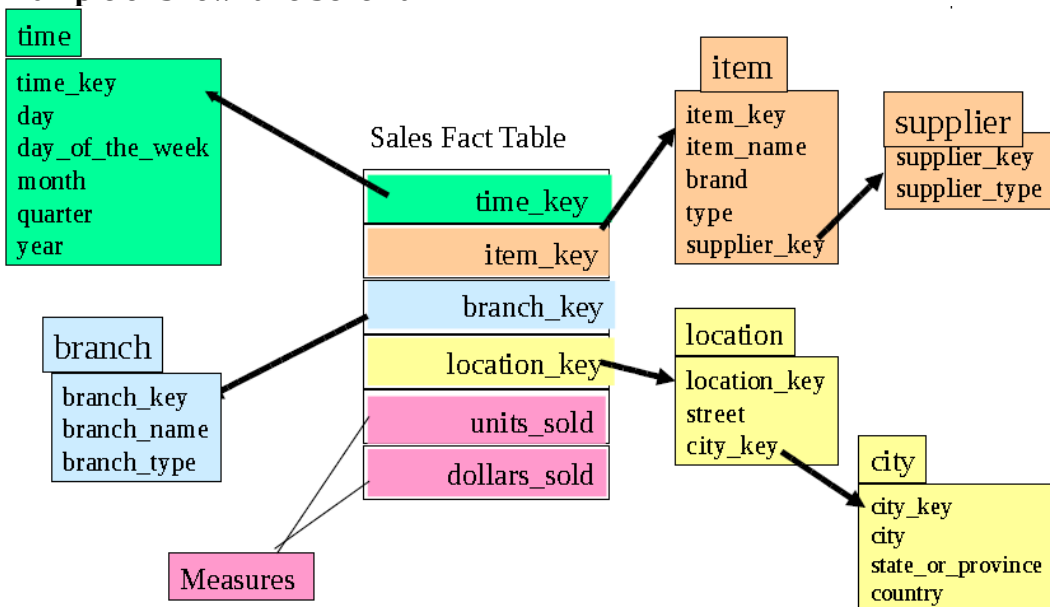


Fig. star schema

Example of Snowflake Schema



Snowflake Schema

Here,

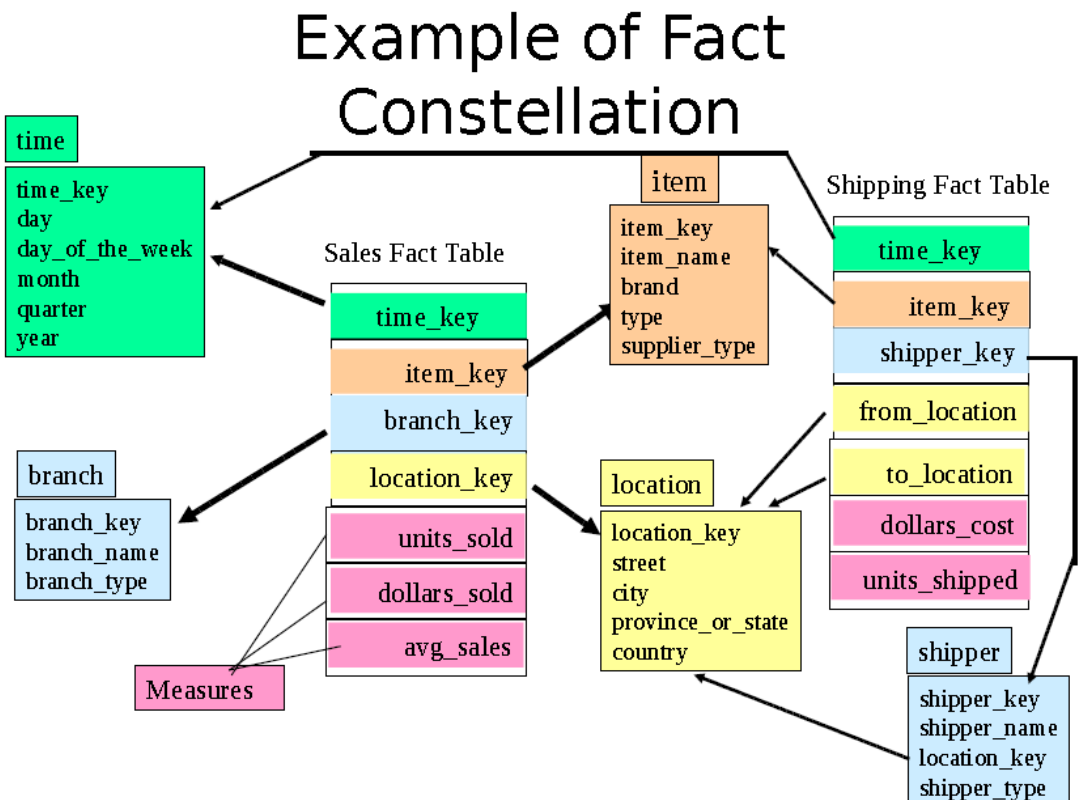
• some dimension tables are *normalized*, thereby further splitting the data into additional tables. The resulting schema graph forms a shape similar to a snowflake.

• The major difference between the snowflake and star schema models is that the dimension tables of the snowflake model may be kept in normalized form to reduce redundancies.

• Such a table is easy to maintain and saves storage space. However, this saving of space is negligible in comparison to the typical magnitude of the fact table.

• Snowflake structure can reduce the effectiveness of browsing, since more joins will be needed to execute a query.

• The system performance may be adversely impacted. Hence, although the snowflake schema reduces redundancy, it is not as popular as the star schema in data warehouse design.



Fact Constellation

• Sophisticated applications may require multiple fact tables to *share* dimension tables. This kind of schema can be viewed as a collection of stars, and hence is called a galaxy schema or a fact constellation.

• This schema specifies two fact tables, *sales* and *shipping*. The *sales* table definition is identical to that of the star schema .

• The *shipping* table has five dimensions, or keys: *item key, time key, shipper key, from location*, and *to location*, and two measures: *dollars cost* and *units shipped*.

• A fact constellation schema allows dimension tables to be shared between fact tables.

• For example, the dimensions tables for *time, item*, and *location* are shared between both the *sales* and *shipping* fact tables. The fact constellation schema is commonly used, since it can model multiple, interrelated subjects.

3. Data warehouse architecture

Design of Data Warehouse: A Business Analysis Framework

• Four views regarding the design of a data warehouse

– Top-down view

• allows selection of the relevant information necessary for the data warehouse

– Data source view

• exposes the information being captured, stored, and managed by operational systems

– Data warehouse view

• consists of fact tables and dimension tables

– Business query view

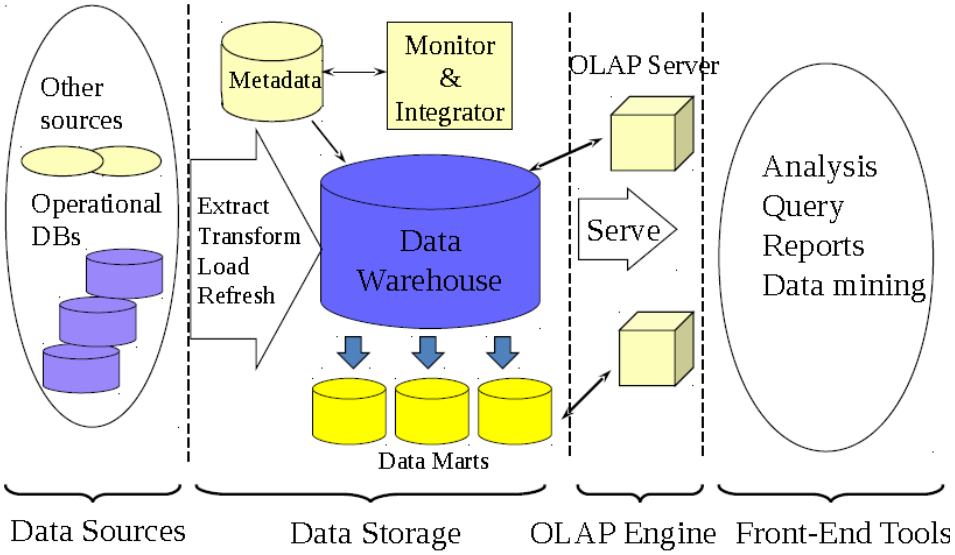
• sees the perspectives of data in the warehouse from the view of end-user

Data Warehouse Design Process

• Top-down, bottom-up approaches or a combination of both

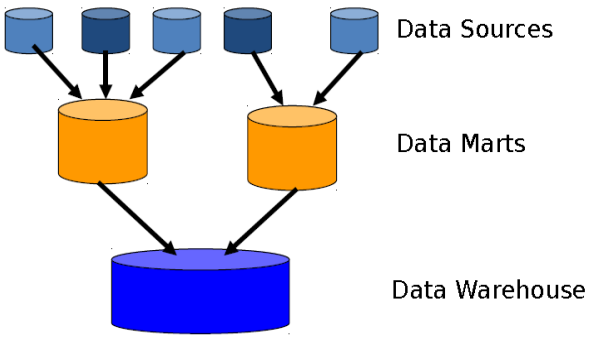
– Top-down: Starts with overall design and planning (mature)
– Bottom-up: Starts with experiments and prototypes (rapid)
• From software engineering point of view
– Waterfall: structured and systematic analysis at each step before proceeding to the next
– Spiral: rapid generation of increasingly functional systems, short turnaround time, quick turn around
• Typical data warehouse design process
– Choose a business process to model, e.g., orders, invoices, etc.
– Choose the *grain* (*atomic level of data*) of the business process
– Choose the dimensions that will apply to each fact table record
– Choose the measure that will populate each fact table record

Data Warehouse: A Multi-Tiered Architecture

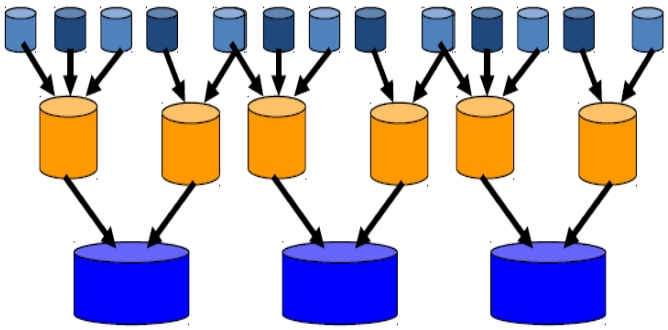


Three Data Warehouse Models
• Enterprise warehouse
– collects all of the information about subjects spanning the entire organization
• Data Mart
– a subset of corporate-wide data that is of value to a specific groups of users. Its scope is confined to specific, selected groups, such as marketing data mart
• Independent vs. dependent (directly from warehouse) data mart
• Virtual warehouse
– A set of views over operational databases
– Only some of the possible summary views may be materialized

Data Mart Centric

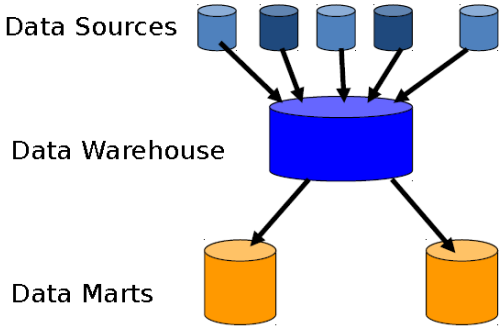


Problems with Data Mart Centric Solution



If you end up creating multiple warehouses, integrating them is a problem

True Warehouse



Data Warehouse Back-End Tools and Utilities
• Data extraction: get data from multiple, heterogeneous, and external sources
• Data cleaning: detect errors in the data and rectify them when possible
• Data transformation: convert data from legacy or host format to warehouse format
• Load: sort, summarize, consolidate, compute views, check integrity, and build indicies and partitions
• Refresh: propagate the updates from the data sources to the Warehouse

Metadata Repository

Meta data is the data defining warehouse objects. It stores:
• Description of the structure of the data warehouse
  - schema, view, dimensions, hierarchies, derived data defn, data mart locations and contents
• Operational meta-data
  - data lineage (history of migrated data and transformation path), currency of data (active, archived, or purged), monitoring information (warehouse usage statistics, error reports, audit trails)
• The algorithms used for summarization
• The mapping from operational environment to the data warehouse
• Data related to system performance
  - warehouse schema, view and derived data definitions
• Business data
  - business terms and definitions, ownership of data, charging Policies

OLAP Server Architectures
• Relational OLAP (ROLAP)
– Use relational or extended-relational DBMS to store and manage warehouse data and OLAP middle ware
– Include optimization of DBMS backend, implementation of aggregation navigation logic, and additional tools and services
– Greater scalability
• Multidimensional OLAP (MOLAP)
– Sparse array-based multidimensional storage engine
– Fast indexing to pre-computed summarized data
• Hybrid OLAP (HOLAP) (e.g., Microsoft SQLServer)
– Flexibility, e.g., low level: relational, high-level: array
• Specialized SQL servers (e.g., Redbricks)
– Specialized support for SQL queries over star/snowflake Schemas

4. Data warehouse implementation

Efficient Data Cube Computation

• Data cube can be viewed as a lattice (network) of cuboids
– The bottom-most cuboid is the base cuboid
– The top-most cuboid (apex) contains only one cell
• Materialization (appearance) of data cube
– Materialize every (cuboids) (full materialization), none (no materialization), or some (partial materialization)
– Selection of which cuboids to materialize
• Based on size, sharing, access frequency, etc.

Cube Operation

- Cube definition and computation in DMQL

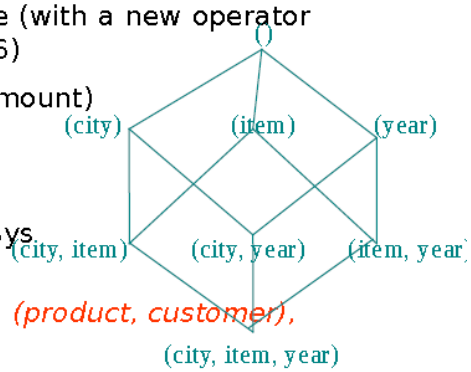
```
define cube sales[item, city, year]:
  sum(sales_in_dollars)

compute cube sales
```

- Transform it into a SQL-like language (with a new operator cube by, introduced by Gray et al.'96)

```
SELECT item, city, year, SUM (amount)
FROM SALES
CUBE BY item, city, year
```

- Need compute the following Group-Bys

(date, product, customer),
(date,product),(date, customer), (product, customer),
(date), (product), (customer)
()



5. From data warehousing to data mining

Data Warehouse Usage

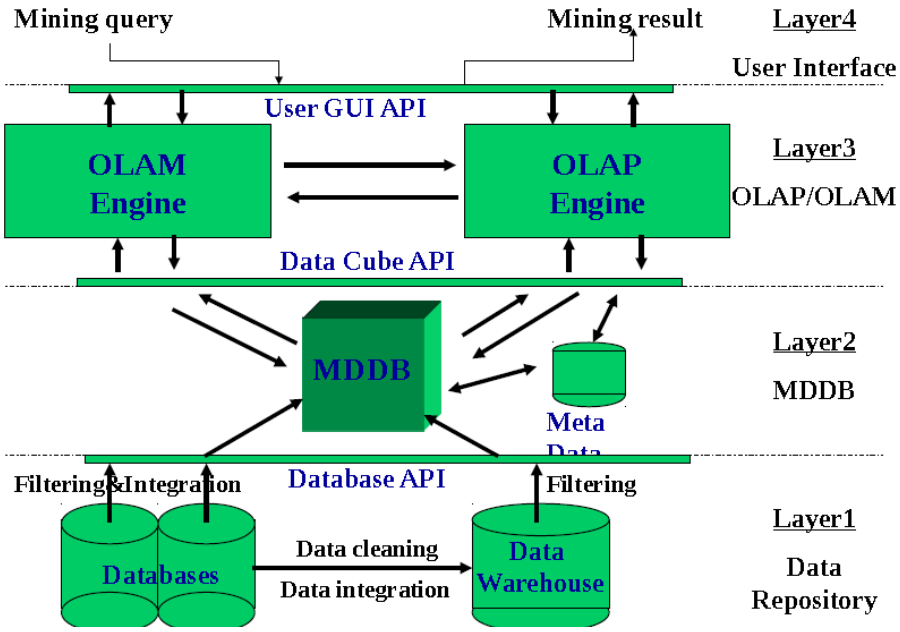
• Three kinds of data warehouse applications
– Information processing
• supports querying, basic statistical analysis, and reporting using crosstabs, tables, charts and graphs
– Analytical processing
• multi-dimensional analysis of data warehouse data
• supports basic OLAP operations, slice-dice, drilling, pivoting
– Data mining
• knowledge discovery from hidden patterns
• supports associations, constructing analytical models, performing classification and prediction, and presenting the mining results using visualization tools

From On-Line Analytical Processing (OLAP) to On Line Analytical Mining (OLAM)

• Why online analytical mining?
– High quality of data in data warehouses
• DW contains integrated, consistent, cleaned data
– Available information processing structure surrounding data warehouses
• ODBC, OLEDB, Web accessing, service facilities, reporting and OLAP tools
– OLAP-based exploratory data analysis

• Mining with drilling, dicing, pivoting, etc.
– On-line selection of data mining functions
• Integration and swapping of multiple mining functions, algorithms, and tasks

An OLAM System Architecture



Data Mining: (Applications)

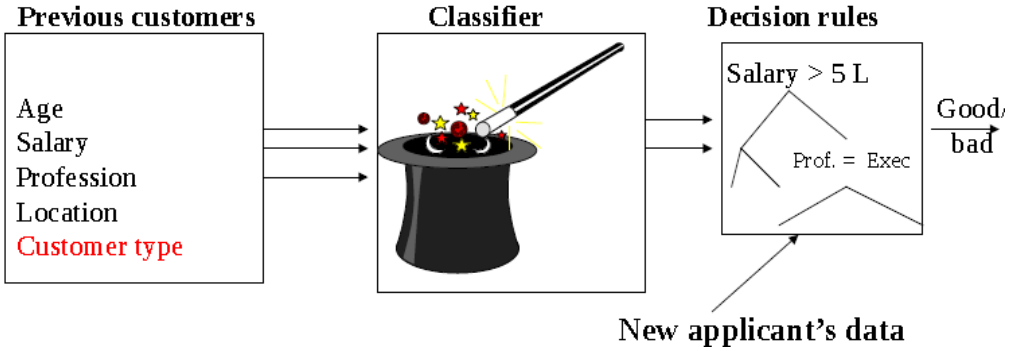
• Banking: loan/credit card approval– predict good customers based on old customers
• Customer relationship management:– identify those who are likely to leave for a competitor.
• Targeted marketing:– identify likely responders to promotions
• Fraud detection: telecommunications, financial transactions:– from an online stream of event identify fraudulent events
• Manufacturing and production:– automatically adjust knobs when process parameter changes
• Medicine: disease outcome, effectiveness of treatments:– analyze patient disease history: find relationship between diseases
• Molecular/Pharmaceutical: identify new drugs
• Scientific data analysis:– identify new galaxies by searching for sub clusters
• Web site/store design and promotion:– find affinity of visitor to pages and modify layout

Definition- Data mining
• Process of semi-automatically analyzing large databases to find interesting and useful patterns
• Overlaps with machine learning, statistics, artificial intelligence and databases but
– more scalable in number of features and instances
– more automated to handle heterogeneous data
• Also known as Knowledge Discovery in Databases (KDD)

Some basic operations
• Predictive: – Regression and Classification
• Descriptive:
– Clustering / similarity matching
– Association rules and variants
– Deviation detection

Classification

• Given old data about customers and payments, predict new applicant's loan eligibility.



Classification methods

Goal: Predict class Ci = f(x1, x2, .. Xn)

• Regression: (linear or any other polynomial)

– a*x1 + b*x2 + c = Ci.

• Nearest neighbor

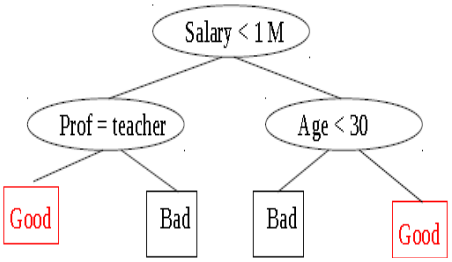
• Decision tree classifier: divide decision space into piecewise constant regions.

• Probabilistic/generative models

• Neural networks: partition by non-linear Boundaries

Decision trees

• Tree where internal nodes are simple decision rules on one or more attributes and leaf nodes are predicted class labels.

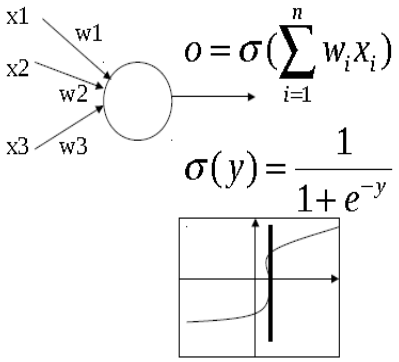


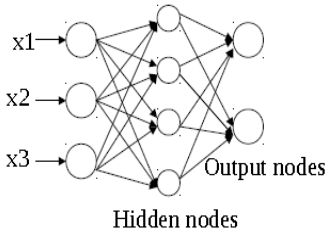
• Pros
  + Reasonable training time
  + Fast application
  + Easy to interpret
  + Easy to implement
  + Can handle large number of features

• Cons
  – Cannot handle complicated relationship between features
  – simple decision boundaries
  – problems with lots of missing data

Neural Network: Set of nodes connected by directed weighted edges

Basic NN unit





• Pros
  + Can learn more complicated class boundaries
  + Fast application
  + Can handle large number of features

• Cons
  – Slow training time
  – Hard to interpret
  – Hard to implement: trial and error for choosing number of nodes

Chapter 5: Planning of Information System

What Is an Information Systems Plan?

– Business planning – the process of identifying the firm's goals, objectives, and priorities + developing action plans for accomplishing them.

– Information systems planning – the part of business planning concerned with developing the firm's information systems resources.

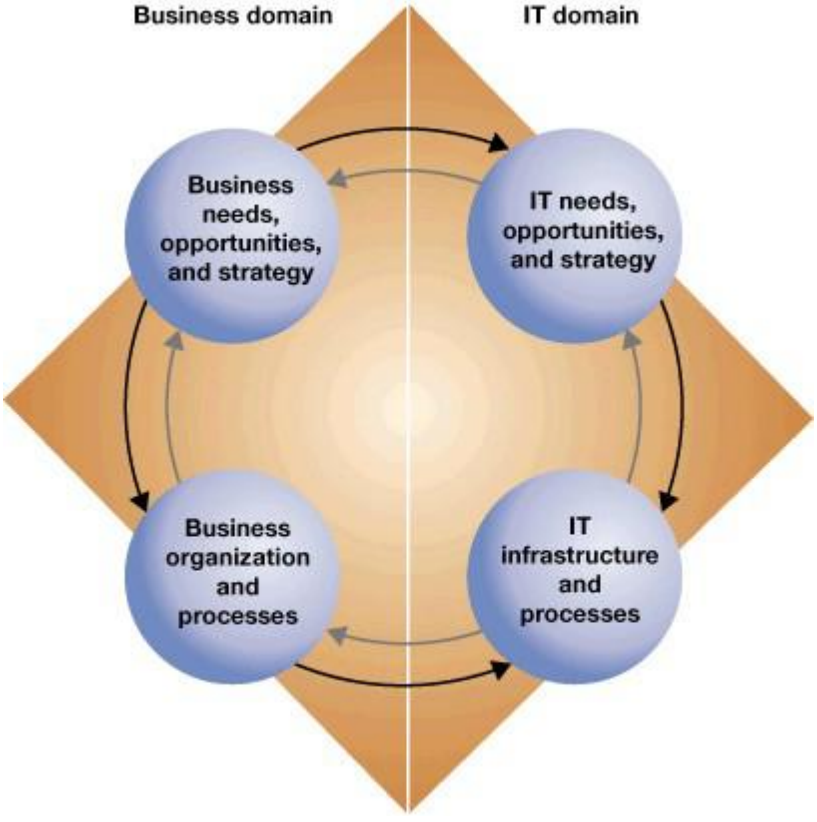
Challenges in Business Planning

• Foreseeing and assessing opportunities
• Assuring consistency with organizational plans and objectives
• Building systems
• Maintaining information system performance
• Collaborating with IT professionals

Principles for IS Planning

• Support the firm's business strategy with appropriate technical architecture
• Evaluate technology as a component of a larger system
• Recognize life cycle costs, not just acquisition costs

Strategic Alignment of Business and IT



Consistency with business priorities

• Critical success factors (CSF) – the things that must go right for a business to succeed
– Identify the firm's primary mission and the objectives that determine satisfactory overall performance
– Executives identify a relatively small number of CSFs

Reengineering and Downsizing

• Business process reengineering (BPR) – the fundamental rethinking and radical redesign of business processes to achieve *dramatic improvements in critical measures of performance*
– Difficult and risky
• Downsizing – improve process efficiency by reducing the number of people involved in the process

The Planning Process

• Management structure provides a way of distinguishing the levels of planning and organizational development which are typical of most businesses (profit or non-profit).

Management Structure & Planning Process:

The planning process - Introduction

• Three main types of plans that a manager uses: **operational, tactical and strategic**
• These three types of plans are stepping stones &
• Their relationship to one another aid in the achievement of organizational goals

Evolution of planning process

• Early stages of IS - applications initiated & implemented as a separate entity without integration
• Using such approach did not satisfy the customer
• Result - Projects were behind schedule & required more resources than expected
• Soon it was clear that identifying project milestones, getting user sign off on specification &
• Scheduling management throughout the cycle produces better result

Why plans are integrated?

• Operational plans are necessary to attain tactical plans and tactical plans lead to the achievement of strategic plans
• Managers use these three main types of plans in their pursuit of company goals
• The elements of organizational planning and development are interrelated, dynamic, and interdependent as shown in the following illustration:

Elements of Organizational Planning:

Strategic & Tactical Thinking

• Strategic thinking gives direction and builds systems for product & service delivery
• Tactical thinking links strategy (a long-term business plan) to operations (an immediate work plan)
• builds systems for day-to-day work supervision, guidance, quality control, and customer relationship
• Thinking is the engine of both long-term and short-term planning
• It provides the intellectual and motivational framework for creativity, vision, problem solving and opportunistic initiative
• It is therefore, possible to identify two kinds of thinking:
– **Strategic / Systems Thinking** and **Tactical / Operational Thinking**.
• In both instances the thinking is typically not a product of one individual, but reflects a group process of collaborative thinking, planning, and decision-making of an executive or operational team.
• Therefore, the skills of collaboration, communication, and shared vision are integral components of the creative thinking process

What is Strategic Planning?

• Process to establish priorities on what you will accomplish in the future
• Forces you to make choices on what you will do and what you will not do
• Pulls the entire organization together around a single game plan for execution
• Broad outline on where resources will get allocated
• Deals with the development of an org's missions, goals, strategies, policies
• Corporation may begin the process by developing a shared vision
• Uses a variety of techniques including team building, scenario modeling & consensus creating exercises
• Most strategic planning done using medium to long range planning horizon
• However, if foreseen incidents occur, an org use strategic planning for a short time frame to confront a crisis

Why do Strategic Planning?

• If you fail to plan, then you plan to fail: Be proactive about the future
• Strategic planning improves performance
• Counter excessive inward and short-term thinking
• Solve major issues at a macro level
• Communicate to everyone what is most important

A Good Strategic Plan should:

• Address critical performance issues
• Create the right balance between what the organization is capable of doing vs. what the organization would like to do
• Cover a sufficient time period to close the performance gap
• Visionary – convey a desired future end state
• Flexible – allow and accommodate change
• Guide decision making at lower levels –operational, tactical, individual

Objectives-Strategic IS plan

• Four main objectives:
1. Business alignment: Aligning investment in IT with company's vision with strategic business goals
2. Competitive advantage: Exploiting IT to create innovative & strategic IS for competitive advantage
3. Resource Management: Developing plans for efficient & effective management of IT resources
4. Technology Architecture: Developing technology policies & designing an IT architecture for the organization

Strategic Planning Model: A B C D E

Where we are, where we want to be, how we will do it and how are we doing

1. **Assessment**
2. **Baseline**
3. **Components**
4. **Down to Specifics**
5. **Evaluate**

Pre-Requisites to Planning:

Before you begin, make sure the groundwork has been done to make the planning process work like
• Senior leadership commitment
• Who will do what?
• What will each group do?
• How will we do it?
• When is the best time?

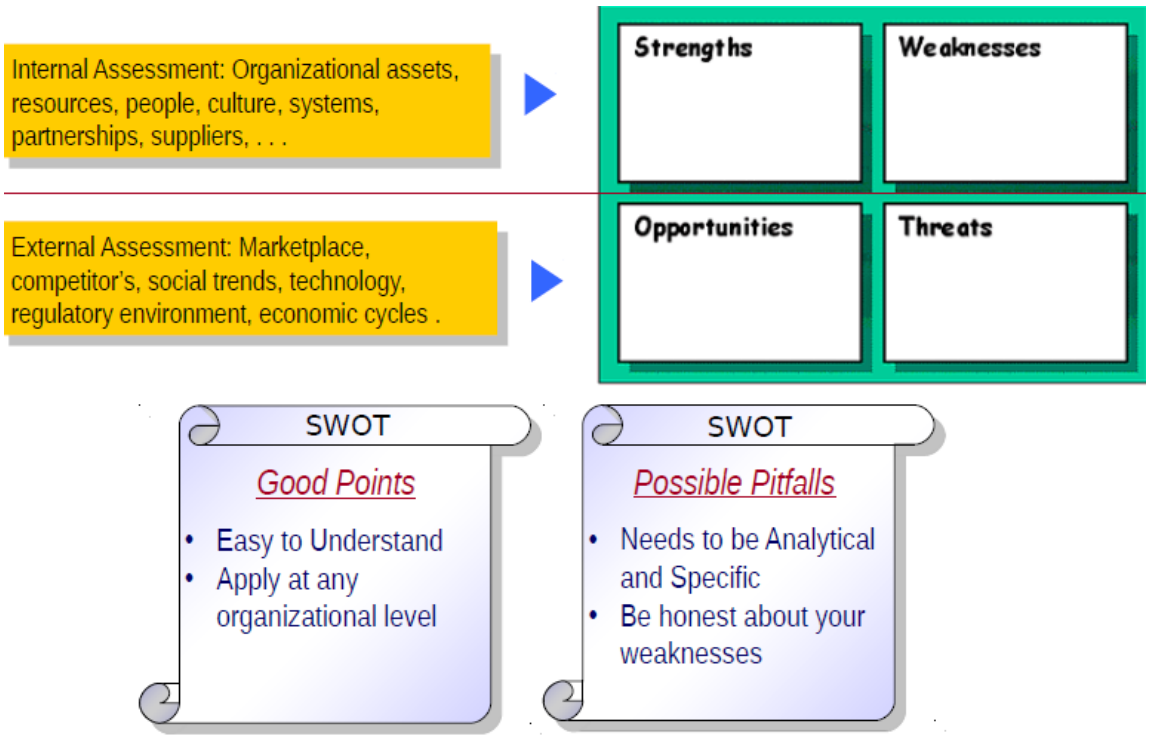
1. **Assessment Model: S W O T**

a. **Strength's Assessment**

• Strength's – Those things that you do well, the high value or performance points
• Strengths can be tangible: Loyal customers, efficient distribution channels, very high quality products, excellent financial condition
• Strengths can be intangible: Good leadership, strategic insights, customer intelligence, solid reputation, high skilled workforce
• Often considered "Core Competencies"

b. **Weaknesses Assessment**

• Weaknesses – Those things that prevent you from doing what you really need to do
• Since weaknesses are internal, they are within your control
• Weaknesses include: Bad leadership, unskilled workforce, insufficient resources, poor product quality, slow distribution and delivery channels, outdated technologies, lack of planning



c. Opportunities Assessment

• Opportunities – Potential areas for growth and higher performance
• External in nature – marketplace, unhappy customers with competitor's, better economic conditions, more open trading policies
• Internal opportunities should be classified as Strength's
• Timing may be important for capitalizing on opportunities

d. Threats Assessment

• Threats – Challenges confronting the organization, external in nature
• Threats can take a wide range – bad press coverage, shifts in consumer behavior, substitute products, new regulations
• May be useful to classify or assign probabilities to threats
• The more accurate you are in identifying threats, the better position you are for dealing with

2. BASELINE: Why create a baseline?

• Puts everything about the organization into a single context for comparability and planning
• Descriptive about the company as well as the overall environment
• Include information about relationships –customers, suppliers, partners
• Preferred format is the Organizational Profile

Organizational Profile

a. Operating Environment

• Products and Services – Suppliers, Delivery Channels, Contracts, Arrangements
• Organizational Culture – Barriers, Leadership, Communication, Cohesiveness
• Workforce Productivity – Skill levels, diversity, contractor's, aging workforce
• Infrastructure – Systems, technology, facilities
• Regulatory – Product / Service Regulation, ISO Quality Standards, Safety, Environmental

b. Business Relationships

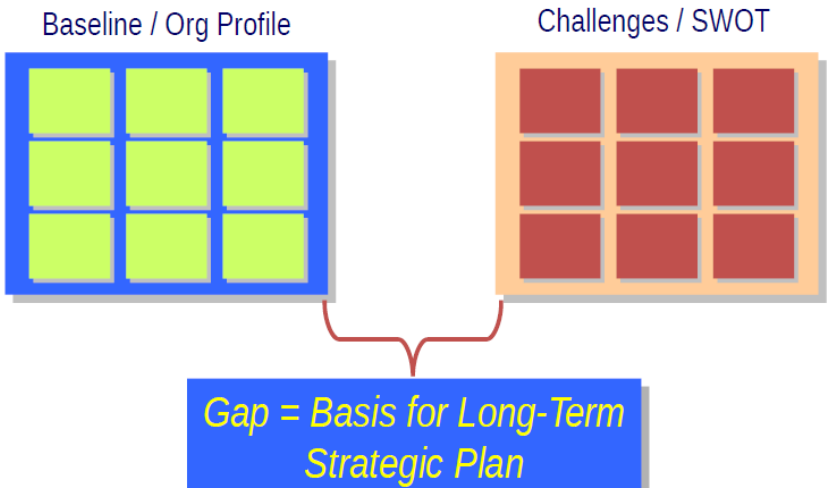
• Organizational Structure – Business Units, Functions, Board, Management Layers

• Customer Relationships – Requirements, Satisfaction, Loyalty, Expectations
• Value Chain – Relationship between everyone in the value chain
• Partner Relationships – Alliances, long-term suppliers, customer partnerships

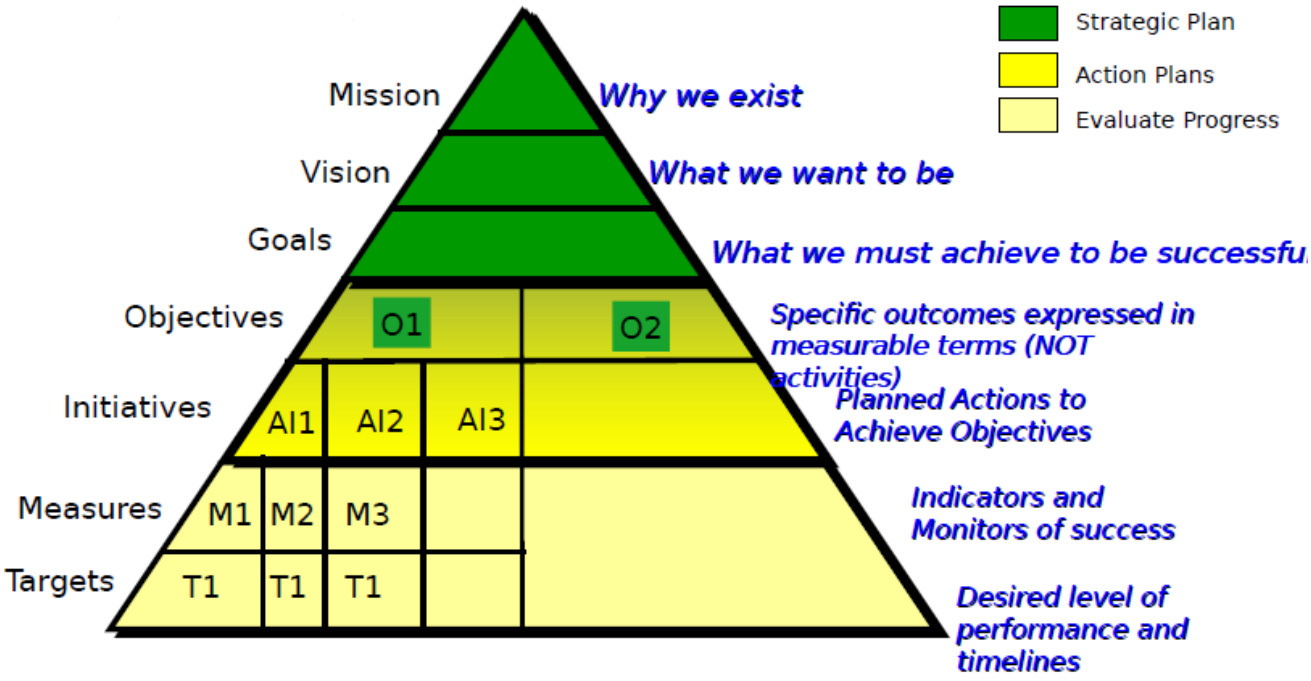
c. Key Performance Categories

• Customer
• Products and Services
• Financial
• Human Capital
• Operational
• External (Regulatory Compliance, Social Responsibility)

GAP Analysis:



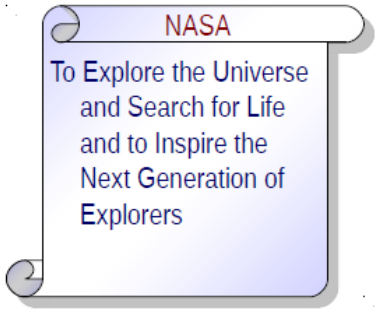
Major Components of the Strategic Plan / Down to Action



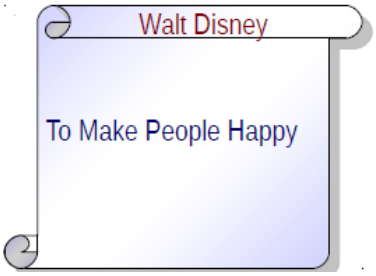
3. Components: Mission Statement

• Captures the essence of why the organization exists – Who we are, what we do
• Explains the basic needs that you fulfill
• Expresses the core values of the organization
• Should be brief and to the point
• Easy to understand
• If possible, try to convey the unique nature of your organization and the role it plays that differentiates it from others

Examples – Good and Bad Mission Statements



Does a good job of expressing the core values of the organization. Also conveys unique qualities about the organization.



Too vague and and unclear. Need more descriptive information about what makes the organization special.

Vision

• How the organization wants to be perceived in the future – what success looks like
• An expression of the desired end state
• Challenges everyone to reach for something significant – inspires a compelling future
• Provides a long-term focus for the entire Organization

Guiding Principles and Values

• Every organization should be guided by a set of values and beliefs
• Provides an underlying framework for making decisions – part of the organization's culture
• Values are often rooted in ethical themes, such as honesty, trust, integrity, respect, fairness
• Values should be applicable across the entire organization
• Values may be appropriate for certain best

Examples of Guiding Principles and Values

- We obey the law and do not compromise moral or ethical principles – ever. We expect to be measured by what we do, as well as what we say.
- We treat everyone with respect and appreciate individual differences. We carefully consider the impact of business decisions on our people and we recognize exceptional contributions.
- We are strategically entrepreneurial in the pursuit of excellence, encouraging original thought and its application, and willing to take risks based on sound business judgment.
- We are committed to forging public and private partnerships that combine diverse strengths, skills and resources.

Goals:
• Describes a future end-state – desired outcome that is supportive of the mission and vision.
• Shapes the way ahead in actionable terms.
• Best applied where there are clear choices about the future.
• Puts strategic focus into the organization –specific ownership of the goal should be assigned to someone within the organization.
• May not work well where things are changing fast – goals tend to be long-term for

Developing Goals
• Cascade from the top of the Strategic Plan –Mission, Vision, Guiding Principles.
• Look at your strategic analysis – SWOT, Environmental Scan, Past Performance, Gaps
• Limit to a critical few – such as five to eight goals.
• Broad participation in the development of goals: Consensus from above – buy-in at the execution level.
• Should drive higher levels of performance

Examples of Goals

- Reorganize the entire organization for better responsiveness to customers
- We will partner with other businesses, industry leaders, and government agencies in order to better meet the needs of stakeholders across the entire value stream.
- Manage our resources with fiscal responsibility and efficiency through a single comprehensive process that is aligned to our strategic plan.
- Improve the quality and accuracy of service support information provided to our internal customers.
- Establish a means by which our decision making process is market and customer focus.
- Maintain and enhance the physical conditions of our public facilities.

Objectives
• Relevant - directly supports the goal
• Compels the organization into action
• Specific enough so we can quantify and measure the results
• Simple and easy to understand
• Realistic and attainable
• Conveys responsibility and ownership
• Acceptable to those who must execute
• May need several objectives to meet a goal

Goals vs. Objectives

GOALS	OBJECTIVES
Very short statement, few words	Longer statement, more descriptive
Broad in scope	Narrow in scope
Directly relates to the Mission Statement	Indirectly relates to the Mission Statement
Covers long time period (such as 10 years)	Covers short time period (such 1 year budget cycle)

Examples of Objectives

- Develop a customer intelligence database system to capture and analyze patterns in purchasing behavior across our product line.
- Launch at least three value stream pilot projects to kick-off our transformation to a leaner organization.
- Centralize the procurement process for improvements in enterprise-wide purchasing power.
- Consolidate payable processing through a P-Card System over the next two years.
- Monitor and address employee morale issues through an annual employee satisfaction survey across all business functions.

4. Down to Specifies: What are Action Plans?

• The Action Plan identifies the specific steps that will be taken to achieve the initiatives and strategic objectives – where the rubber meets the road
• Each Initiative has a supporting Action Plan(s) attached to it
• Action Plans are geared toward operations, procedures, and processes
• They describe who does what, when it will be completed, and how the organization knows when steps are completed
• Like Initiatives, Action Plans require the monitoring of progress on Objectives, for which measures are needed

Objectives → Initiatives → Action Plans

Characteristics of Action Plans

• Assign responsibility for the successful completion of the Action Plan. Who is responsible? What are the roles and responsibilities?
• Detail all required steps to achieve the Initiative that the Action Plan is supporting. Where will the actions be taken?
• Establish a time frame for the completion each steps. When will we need to take these actions?
• Establish the resources required to complete the steps. How much will it take to execute these actions?
• Define the specific actions (steps) that must be taken to implement the initiative. Determine the deliverables (in measurable terms) that should result from completion of individual steps. Identify in-process measures to ensure the processes used to carry out the action are working as intended. Define the expected results and milestones of the action plan.
• Provide a brief status report on each step, whether completed or not. What communication process will we follow? How well are we doing in executing our action plan?
• Based on the above criteria, you should be able to clearly define your action plan. If you have several action plans, you may have to prioritize.

Action Plan Execution

• Requires that you have answered the Who, What, How, Where, and When questions related to the project or initiative that drives strategic execution
• Coordinate with lower level sections, administrative and operating personnel since they will execute the Action Plan in the form of specific work plans
• Assign action responsibility and set timelines –Develop working plans and schedules that have specific action steps
• Resource the project or initiative and document in the form of detail budgets (may require reallocation prior to execution)
• Monitor progress against milestones and measurements
• Correct and revise action plans per comparison of actual results against original action plan

Quantify from Action Level Up in terms of Measurements

• Measure your milestones – short-term outcomes at the Action Item level.
• Measure the outcomes of your objectives.
• Try to keep your measures one per objective.
• May want to include lead and lag measures to depict cause-effect relationships if you are uncertain about driving (leading) the desired outcome.
• Establish measures using a template to capture

Criteria for Good Measures

Integrity – Complete; useful; inclusive of several types of measure; designed to measure the most important activities of the organization
Reliable: Consistent
Accurate - Correct
Timely – Available when needed: designed to use and report data in a usable timeframe
Confidential and Secure: Free from inappropriate release or attack

Examples of Measurements Lead Indicators

• Average time to initiate customer contact: shorter time should lead to better customer service
• Average response time to incident: below average response times should lead to increased effectiveness in dealing with incident
• Facilities that meet facility quality A1 rating: should lead to improved operational readiness for meeting customer needs

Examples of Measurements Lag Indicators:

• Overall customer satisfaction rating: how well you are doing looking back
• Business Units met budgeted service hour targets: after the fact reporting of service delivery volume
• Number of category C safety accidents at construction sites: historical report of what has already taken place

Targets:

Average Time to Process New Employee Setups in DB	65 days Year 2007	60 days Year 2008	55 days Year 2009
Utilization Rate for Rental Housing Units	90% for Year 2007	92% for Year 2008	95% for Year 2009
Toxic Sites meeting in-service compliance	55% for Year 2007	70% for Year 2008	95% for Year 2009
Personnel Fully Trained in Safety and Emergency	65% by 2rd Quarter	75% by 3th Quarter	90% by 4th Quarter
Open Positions Filled after 30 day promotion period	75 positions Sept 2007	100 positions Jan 2008	135 positions July 2008
% Reduction in Orders Filled Short in 1st Cycle	50% by Year 2008	65% by Year 2009	85% by Year 2010

Fig. Examples of target

• For each measurement, you should have at least one target
• Targets should stretch the organization to higher levels of performance
• Incremental improvements over current performance can be used to establish your targets
• Targets put focus on your strategy
• When you reach your targets, you have successfully executed your strategy

5. **Evaluate: Continuous Feedback through the Balanced Scorecard**

• Cascade and align from the top to create a Strategic Management System.
• Use the Balanced Scorecard framework to organize and report actionable components.
• Use the Scorecard for managing the execution of your strategy.
• Scorecard "forces" you to look at different perspectives and take into account cause effect relationships (lead and lag indicators)
• Improves how you communicate your strategy – critical to execution.

D2-D5: Build the Balanced Scorecard Performance Management

• Establish a regular review cycle using your balanced scorecard.
• Analyze and compare trends using graphs for rapid communication of performance.
• Don't be afraid to change your metrics – life cycle (inputs to outputs to outcomes)
• Work back upstream to revise your plans: Action Plans > Operating Plans > Strategic Plans
• Planning is very dynamic – must be flexible to change.
• Recognize and reward good performance results
• Brainstorm and change – take corrective action on poor performance results.

D2-D5: Build the Balanced Scorecard Automating the Process:

1. Active Strategy (www.activestrategy.com)
2. Corda (www.corda.com)
3. Corporater (www.corporater.com)
4. Rocket CorVu (www.corvu.com)
5. Cockpit Communicator (www.4ghi.com)
6. Biz Score (www.efmsoftware.nl)
7. Executive Dashboard (www.iexecutivedashboard.com)
8. PM Express (www.pm-express.com)
9. Strategy 2 Act (www.strategy2act.com)
10. 20 20 Software (www.cashfocus.com)

Link Budgets to Strategic Plan

• The world's best Strategic Plan will fail if it is not adequately resourced through the budgeting process
• Strategic Plans cannot succeed without people, time, money, and other key resources
• Aligning resources validates that initiatives and action plans comprising the strategic plan support the strategic objectives

What Resources? How to Link?

Every Action Plan should identify the following:
•The **people** resources needed to succeed
•The **time** resources needed to succeed
•The **money** resources needed to succeed
•The **physical resources** (facilities, technology, etc.) needed to succeed

Resource information is gathered by Objective Owners which is provided to the Budget Coordinators for each Business Unit.
Resources identified for each Action Plan are used to establish the total cost of the Initiative.
Cost-bundling of Initiatives at the Objective level is used by our Business Unit Budget Coordinators to create the Operating Plan Budget

Some Final Thoughts
• Integrate all components from the top to the bottom: Vision > Mission > Goals > Objectives > Measures > Targets > Initiatives > Action Plans > Budgets.
• Get Early Wins (Quick Kills) to create some momentum
• Seek external expertise (where possible and permissible)
• Articulate your requirements to senior leadership if they are really serious about strategic execution

Tactical Plans
• **Tactical plans** support strategic plans by translating them into specific plans relevant to a distinct area of the organization.
• Tactical plans are concerned with the-responsibility and functionality of lower-level departments to fulfill their parts of the strategic plan
• It involves the design of tactics, setting of objectives, development of procedures, rules, schedules & budgets
• Most tactical IS planning is done on a short range basis
• It builds on business/IT strategies developed in strategic IS planning stages
• It produces project proposals for development of new or Improved IS that implement IT architecture created during strategic IS planning
• These projects are then evaluated, ranked & fitted in multilayer development plan
• Finally, a resource allocation plan is developed to specify the IS resources, financial commitments & org. change needed to implement strategic IT development plan of firm
• Budget acts as a control mechanism

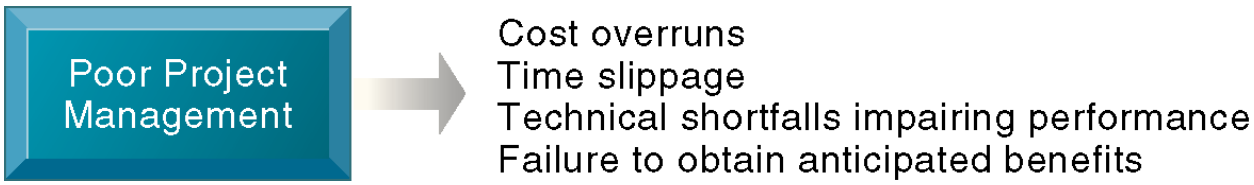
Example: Tactical Plans
• For example, when Martha, the middle-level manager at Nino's, learns about Tommy's strategic plan for increasing productivity, Martha immediately begins to think about possible tactical plans to ensure that happens.
• Tactical planning for Martha might include things like testing a new process in making pizzas that has been proven to shorten the amount of time it takes for prepping the pizza to be cooked or perhaps looking into purchasing a better oven that can speed up the amount of time it takes to cook a pizza or even considering ways to better map out delivery routes and drivers
• As a tactical planner, Martha needs to create a set of calculated actions that take a shorter amount of time and are narrower in scope than the strategic plan is but still help to bring the organization closer to the longterm goal.

Operational Plans
• **Operational plans** sit at the bottom of the totem pole;
• They are the plans that are made by frontline, or low-level, managers.
• All operational plans are focused on the specific procedures and processes that occur within the lowest levels of the organization
• Managers must plan the routine tasks of the department using a high level of detail.

• It involves detailed planning for accomplishment of new IS development projects
• Includes preparation of operating budgets
• Annual operating budgets specify allocation of financial & other resources needed in support the organization's information service operations & system development & maintenance activities
• Project planning is an important operational planning function
• It involves development of plans, procedures & schedule for IS development project
• Such planning is an important part of the project management effort that plans & controls the implementation of business projects
• Frank, the frontline manager at Nino's Pizzeria, is responsible for operational planning.
• Operational planning activities for Frank would include things like scheduling employees each week; assessing, ordering and stocking inventory; creating a monthly budget; developing a promotional advertisement for the quarter to increase the sales of a certain product (such as the Hawaiian pizza) or outlining an employee's performance goals for the year.
• Operational plans can be either single-use or ongoing plans.
• **Single-use plans** are those plans that are intended to be used only once.
• They include activities that would not be repeated and often have an expiration.
• Creating a monthly budget and developing a promotional advertisement for the quarter to increase the sales of a certain product are examples of how Frank would utilize single-use planning.
• Operational plans are made by low-level managers.
• **Ongoing plans** are those plans that are built to withstand the test of time.
• They are created with the intent to be used several times and undergo changes when necessary.
• Outlining an employee's performance goals for the year would be considered an ongoing plan that Frank must develop, assess and update, if necessary. Ongoing plans are typically a **policy**, **procedure** or **rule**.
• Policies are general statements, or guidelines, that aid a manager in understanding routine responsibilities of his or her role as a manager.
• Examples of policies include things such as hiring, training, outlining and assessing performance appraisals and disciplining and terminating subordinates.
• A procedure details the step-by-step process of carrying out a certain task, such as assessing, ordering and stocking inventory.
• A rule provides managers and employees with specific and explicit guidelines of behavior that is what they should and should not do as a member of the organization

Chapter 6 – implementation of Information system

Consequences of Poor Project Management



Without proper management:
• A systems development project takes longer to complete and
• Most often exceeds the allocated budget
• The resulting information system most likely is technically inferior and
•May not be able to demonstrate any benefits to the organization.

Selecting Projects

Enterprise analysis and critical success factors:

• To develop effective information systems plan, organization must have clear understanding of:
- Long-term information requirements
- Short-term information requirements

• Two principal methodologies for establishing essential information requirements of organization as a whole
- Enterprise analysis
- Critical success factors

Enterprise analysis (business systems planning):

• Seeks to understand information requirements by examining entire organization in terms of organizational units, functions, processes, and data elements
• Helps identify key entities and attributes of firm's data
• Central method is large survey of managers on how they use information
• Results analyzed and data elements organized into logical application groups
• Disadvantages: Produces enormous amount of data; expensive; time-consuming

Critical Success Factors

• Developed to help senior executives
• defines the information needs for the purpose of managing their organizations
• CSFs gathers information needs for management decision-making
• CSFs gets sharpened by the industry, the firm, the manager & the broader environment

CSF characteristics:

– CSF hierarchy,
– types,
– uniqueness, and
– stability over time.

CSF Hierarchy:

1. Industry

2. Organizational
3. Division and
4. Individual

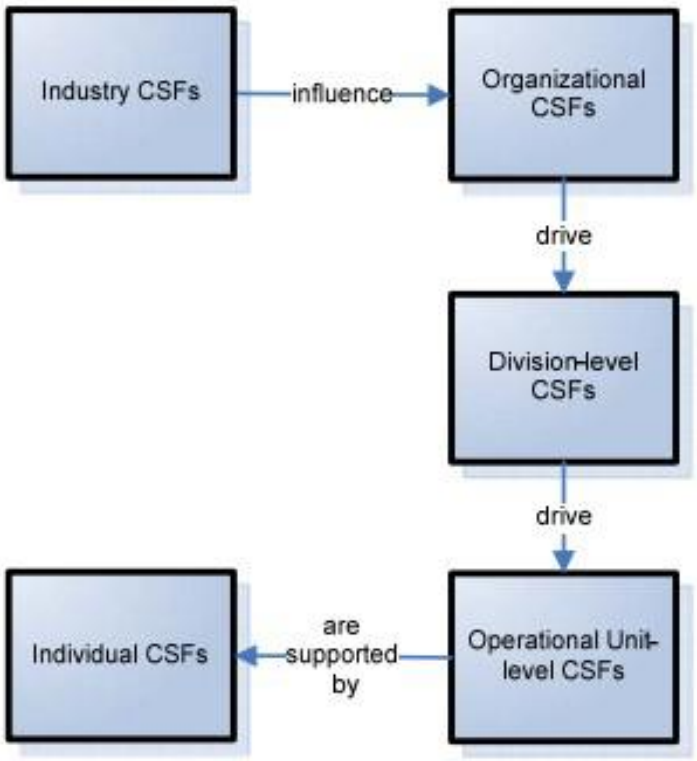


Figure: CSF Hierarchy

CSF Types:
Five types of CSFs that reflect the way in which they contribute to the achievement of the mission:

1. The structure of the particular industry (industry CSFs)
2. Competitive strategy, industry position, and geographical location (strategy CSFs)
3. The macro environment (environmental CSFs)
4. Problems or challenges to the organization (temporal CSFs)
5. Management perspective (management CSFs)

Critical Success Factor Method:
1. Define scope
2. Collect data
3. Analyze data
4. Derive CSFs
5. Analyze CSFs

CSF advantages
• CSF produces smaller data set to analyze than enterprise analysis
– E.g. Auto industry CSFs might include styling, quality, cost
• Central method:

- Only Interviews with top managers to identify goals and resulting CSFs
- Personal CSFs aggregated to develop firm CSFs

• The question focuses on small number of CSFs rather than broad inquiry
• Especially suitable for top level management &
• For development of DSS & ESS
• Unlike enterprise analysis, the CSF method focuses org attention on how information should be handled

CSF disadvantages

•No clear methods for aggregation of personal CSFs into firm CSFs
•Confusion between
- Interviewers & interviewees &
- individual CSFs and organizational CSFs

Using CSFs to Develop Systems

• The CSF approach relies on interviews with key managers to identify their CSFs.
• Individual CSFs are aggregated to develop CSFs for the entire firm.
• Systems can then be built to deliver information on these CSFs
• Furthermore, managers are asked to identify their goals& CSF that would ensure attainment of goals
• This includes consideration of
– competitive strategies of the firm,
– its position in the industry &
– its economic & political environment
• Managers are also asked to a limited number of key areas where things have to go right
• CSF can be used to develop an IT architecture & IS architecture portfolio

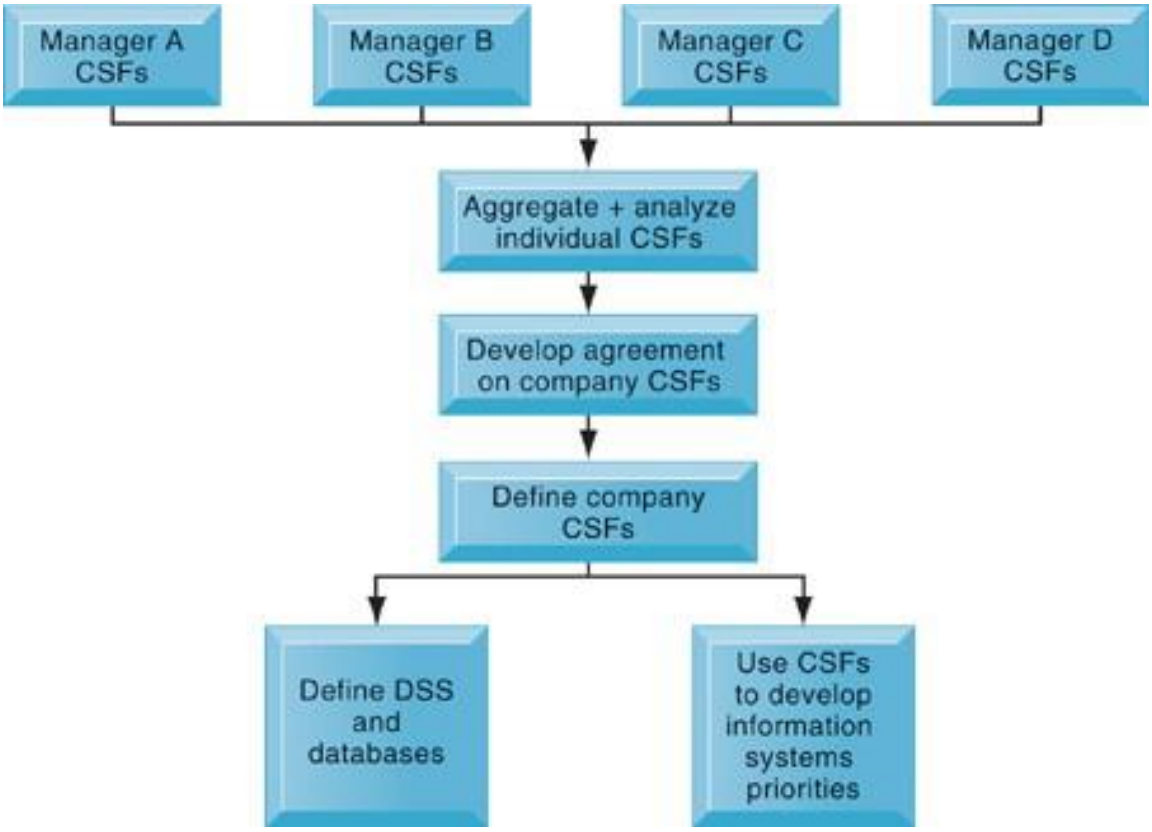


Figure: Using CSFs to Develop Systems

Portfolio Analysis

• Once an association has adopted a strategic plan, the next step is to convert the goals and objectives in that plan to a work plan and budget

• Portfolio analysis has been devised to help associations bridge the gap between strategy formulation and strategy implementation.

• Portfolio analysis is *a systematic way to analyze the products and services* that make up an association's business portfolio

• Portfolio analysis helps you decide which of these products and services should be emphasized and which should be phased out, based on objective criteria.

• Used to evaluate alternative system projects

• Inventories all of the organization's information systems projects and assets are collected

• Each system has profile of risk and benefit

- High-benefit, low risk
- High-benefit, high risk
- Low-benefit, low risk
- Low-benefit, high risk

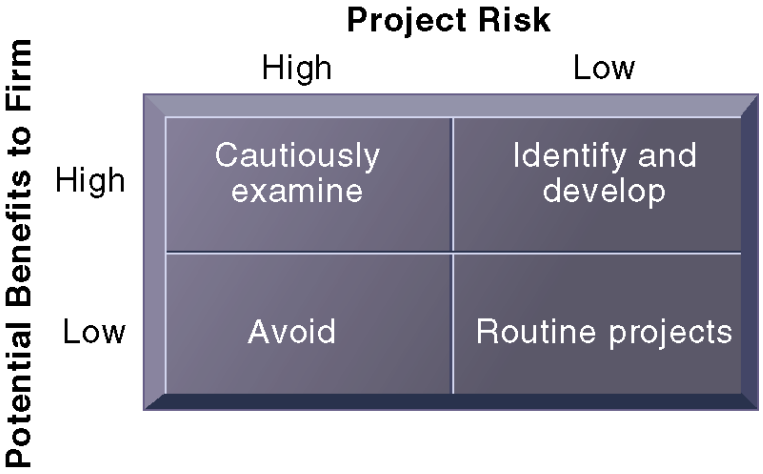
• To improve return on portfolio, balance Portfolio analysis

A System Portfolio

• Companies should examine their portfolio of projects in terms of potential benefits and likely risks

• Certain kinds of projects should be avoided altogether and others developed rapidly

• There is no ideal mix. Companies in different industries have different profiles



Portfolio analysis: advantages:

1. It encourages management to evaluate each of the organization's businesses individually and to set objectives and allocate resources for each.
2. It stimulates the use of externally oriented data to supplement management's intuitive judgment.
3. It raises the issue of cash flow availability for use in expansion and growth.

Portfolio analysis: disadvantages:

1.It is not easy to define product/market segments.
2. It provides an illusion of scientific rigor when some subjective judgments are involved.

Dimensions of project risk: Level of project risk influenced by:

1. **Project size**

- Indicated by cost, time, number of organizational units affected

- Organizational complexity also an issue

2. **Project structure:** Structured, defined requirements run lower risk
3. **Experience with technology**

System implementation benefits from:

• **High levels of user involvement**
- System more likely to conform to requirements
- Users more likely to accept system

• **Management support**
- Positive perception by both users and technical staff
- Ensures sufficient funding and resources
- Enforcement of required organizational changes

User-designer communication gap

• Users and information systems specialists tend to have different backgrounds, interests, and priorities
• Leads to divergent organizational loyalties, approaches to problem solving, and vocabularies
- **User concerns:**
  - Will the system deliver the information I need for work?
  - How quickly can I access the data?
- **Designer concerns:**
  - How much disk storage space will the master file require?

Controlling risk factors

• First step in managing project risk involves identifying nature and level of risk of project
• Each project can then be managed with tools and risk-management approaches geared to level of risk
• Managing technical complexity: Internal integration tools
- Project leaders with technical and administrative experience
- Highly experienced team members
- Frequent team meetings
- Securing of technical experience outside

Formal planning and control tools

• Document and monitor project plans
• Help identify bottlenecks and determine impact of problems on project completion times
• Chart progress of project against budgets and target dates
• **Gantt chart**
- Lists project activities and corresponding start and completion dates
- Visual representation of timing of tasks and resources required

• **PERT chart**
- Portrays project as network diagram
- Nodes represent tasks

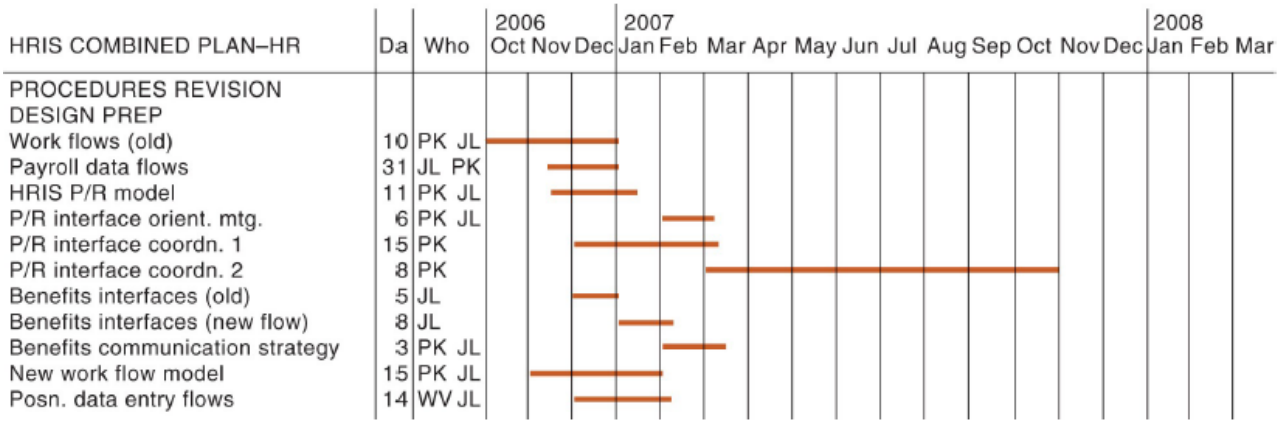


Fig. Gantt Chart

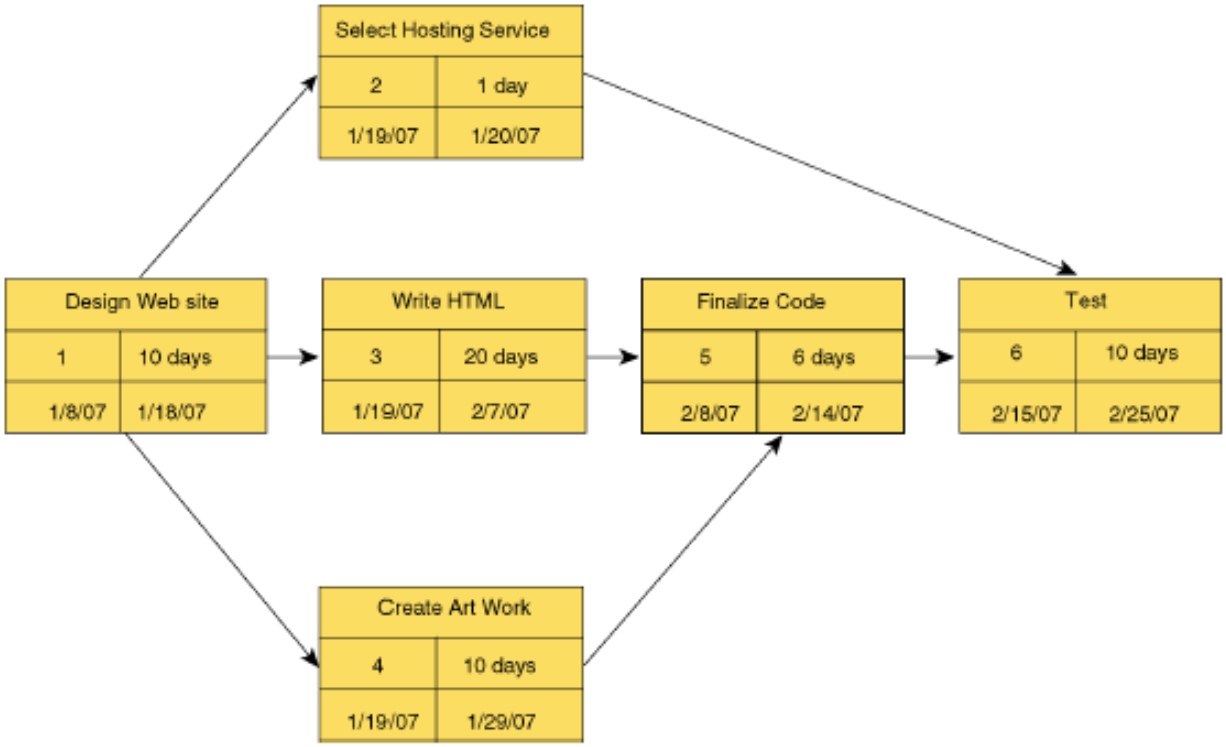


Fig. PERT Chart

This is a simplified PERT chart for creating a small Web site. It shows the ordering of project tasks and the relationship of a task with preceding and succeeding tasks.

Increasing user involvement and overcoming user

• **External integration tools** consist of ways to link work of implementation team to users at all organizational levels

- Active involvement of users
- Implementation team's responsiveness to users

• User resistance to organizational change

- Users may believe change is detrimental to their interests
- **Counter implementation**: Deliberate strategy to thwart implementation of an information system or an innovation in an organization
  • E.g. increased error rates, disruptions, turnover, sabotage

Strategies to overcome user resistance

• User education and training

• Management edicts and policies
• Incentives for cooperation
• Improvement of end-user interface
• Resolution of organizational problems prior to introduction of new system

Designing for the organization

• Information system projects must address ways in which organization changes when new system installed

- Procedural changes
- Job functions
- Organizational structure
- Power relationships
- Work structure

• **Ergonomics:** Interaction of people and machines in work environment

- Design of jobs
- Health issues
- End-user interfaces

• **Organizational impact analysis**

- How system will affect organizational structure, attitudes, decision making, operations

• **Sociotechnical design**

- Addresses human and organizational issues
- Separate sets of technical and social design solutions
- Final design is solution that best meets both technical and social objectives

• **Project management software**

- Can automate many aspects of project management
- Capabilities for
  - Defining, ordering, editing tasks
  - Assigning resources to tasks
  - Tracking progress
- Microsoft Project
  - Most widely used project management software
  - PERT, Gantt charts
  - Critical path analysis
  - Product Guide wizards
  - Enterprise Project Management Solution version

System Development & Organizational change

• IT promotes various degrees of organizational change,
• Ranging from incremental to far reaching
• Most common form of organizational change ranges from:
1. Automation
2. Rationalization-streamlining standard operating procedures
3. Reengineering- business process analyzed, simplified & redesigned
4. Paradigm shifts – in order to get very high productivity

The importance of change management information systems success and failure

Change Management and the Concept of Implementation

Change Management:

- Changes in the way that information is defined, accessed and used to manage the organizations resources for various benefits.
- Changes breeds resistance and opposition.
- Managing it is called change management

Implementation:

- All organizational activities working toward the adoption, management, and routinization of a new system change agent
- The systems analyst who develops technical solutions and redefines the configurations, interactions, job activities, and power relationships of various organizational groups
- Acts as catalyst for the entire change process and is responsible for ensuring that all parties involved accept the changes created by a new system

Causes of Implementation Success and Failure

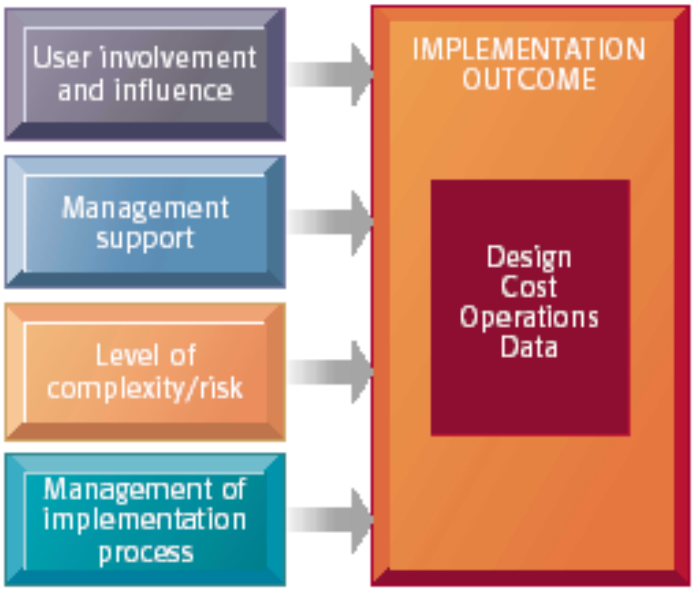


Fig. Information Systems Success or Failure Factors

User Involvement and Influence

- If users are heavily involved in systems design, they have more opportunities to mold the system according to their priorities and business requirements and control the outcome.
- Involved users are more likely to react positively to the completed system.
- User-Designer Communications Gap: Users can have limited understanding of other issues and solutions.
- Management Support and Commitment: Commitment of management to An information systems project usually results in a more positive perception and acceptance by users and the technical services staff.
- Management backing also ensures that a systems project receives sufficient funding and resources to be successful All the changes in work habits and procedures and any organizational realignment associated with a new system depend on management backing

Level of Complexity and Risk

- The level of project risk is influenced by:
- Project size

- Project structure
- Level of technical expertise of the information systems team

Likely Consequences of Poor Project Management:

- Costs that vastly exceed budgets
- Unexpected time slippage
- Technical shortfalls resulting in performance that is significantly below the estimated level
- Failure to obtain anticipated benefits. Possible reasons for poor management:
- Ignorance and optimism
- Mythical man-month
- Falling behind: Bad news travels slowly upward

Change Management Challenges for Business Process Reengineering, Enterprise Applications, and Mergers and Acquisitions

Successful implementation includes addressing employees' concerns about change:

- Resistance by key managers
- Changing job functions, career paths, recruitment practices
- Managing training

System Implications of Mergers and Acquisitions (M&As):

- As are major growth engines for businesses, enabling firms to
- Gain market share and expertise very quickly
- Critical issues include the organizational characteristics of the merging companies and IT infrastructures
- Realistic costs of integration
- Estimated benefits of economies in operation, scope, knowledge, and time
- Problematic systems that require major investments to integrate
- More than 70 percent of all M&As result in a decline in shareholder value

Controlling Risk Factors

1. Managing technical complexity:

- Formal planning and control tools
- Increasing user involvement and overcoming user resistance
- External integration tools: Ways to link the work of the implementation team to users at all organizational levels
- Counter implementation: Deliberate strategy to thwart the implementation of an information system or an innovation in an organization

2. Designing for the Organization

- Systems development must address how the organization will change when the new system is installed, including installation of intranets, extranets, and Web applications
- Organizational impact analysis
- Allowing for the human factor User performance standards
- Ergonomics

3. Sociotechnical Design:

- Explores workgroup organization and impacts from technical solutions
- Blends technical efficiency with sensitivity to human and organizational needs
- Raises productivity without sacrificing human and social goals

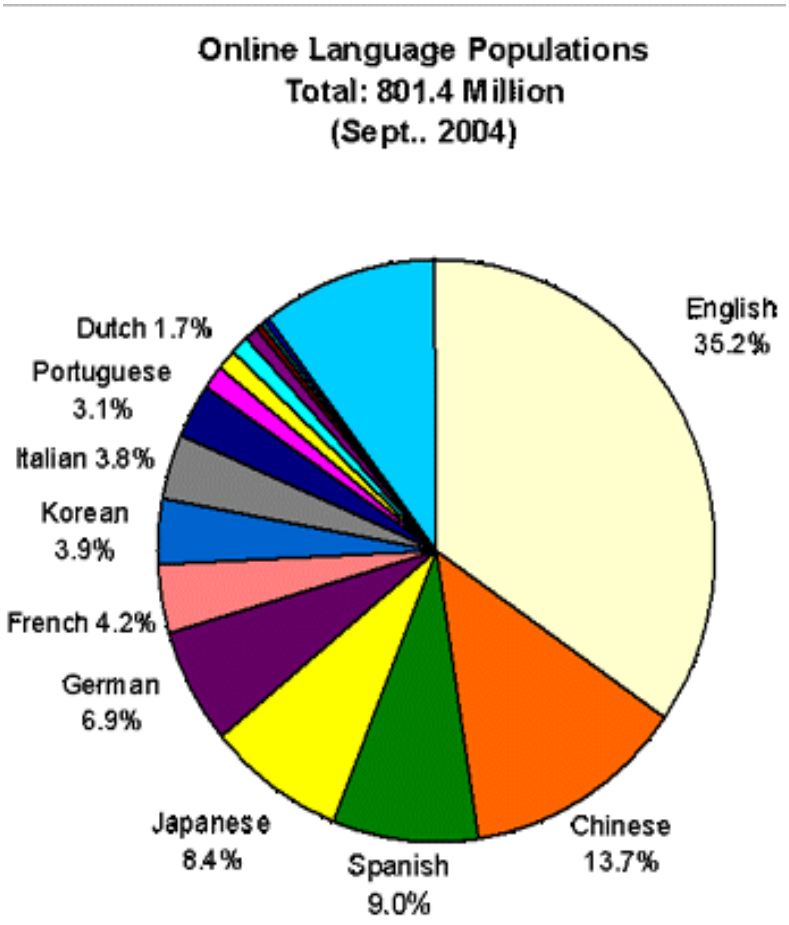
Chapter 7- Web based information system and navigation

7.1 The Structure of the Web:

- How many people use the web?
- What is the size of the web?
- How many web sites are there?
- How many searches per day?
- How do web pages change?
- What is the graph structure of the web?
- How could the structure arise?
- What can we do with link analysis?

Global Internet Statistics:

- 25% of world population is online as of mid-2009
- 51.4% online in the UK
- 92.5% online is Sweden
- 58.4% online in Europe
- 77.4% online in USA
- 10.9% online in Africa
- 21.5% online in Asia

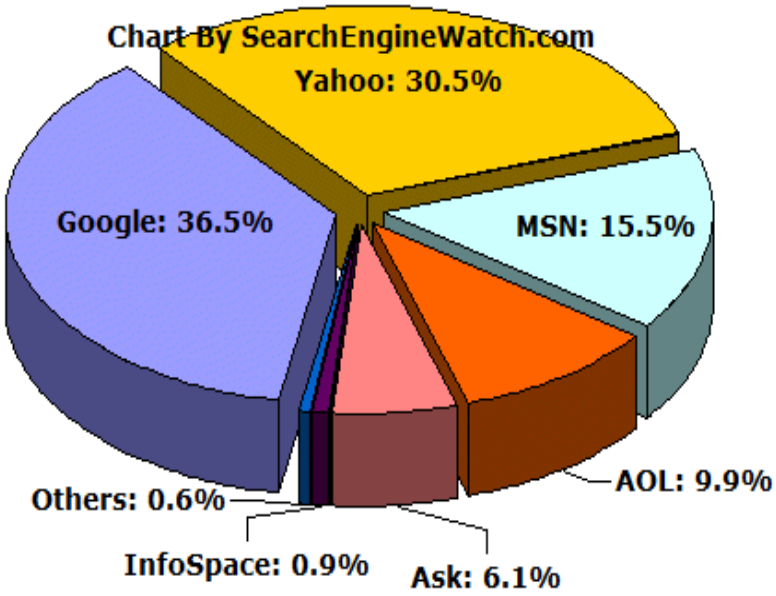


The Size of the Web

- Lawrence and Giles 1999 – 800 million
- Over 11.5 billion in 2005 (Google indexed over 8 billion at the time)
- About 600 billion in 2010, approaching 1 trillion
- The *deep* (or *hidden* or *invisible)* web contains 400-550 times more information.

Search Engine Statistics

- Google has over 40,000 searches a second.
- In 2005 Google has 36.5% searches but as of 2010 Google dominates with Bing and Yahoo far behind.
- In China and Korea local engines are more popular.
- Users are spending more time on the web (over 34 hours a month, Feb. 2009).



How do Web Pages Change:

- Most pages do not change much.
- Larger pages change more often.
- Commercial pages change more often.
- Past change to a web page is a good indicator of future change.
- About 30% of pages are very similar to other pages, and being a near-duplicate is fairly stable.

7.2 Link analysis:

Meta-Search Engines

- Search engine that passes query to several other search engines and integrate results.
  - Submit queries to host sites.
  - Parse resulting HTML pages to extract search results.
  - Integrate multiple rankings into a "consensus" ranking.
  - Present integrated results to user.
- Examples:
  - Metacrawler
  - SavvySearch
  - Dogpile

HTML Structure & Feature Weighting

- Weight tokens under particular HTML tags more heavily:
  - <TITLE> tokens (Google seems to like title matches)
  - <H1>,<H2>… tokens
  - <META> keyword tokens
- Parse page into conceptual sections (e.g. navigation links vs. page content) and weight tokens differently based on section.

Bibliometrics: Citation Analysis

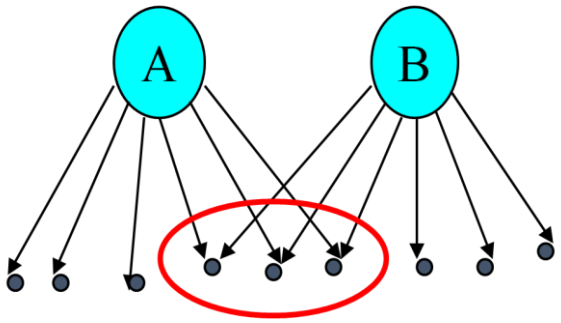
- Many standard documents include *bibliographies* (or *references*), explicit *citations* to other previously published documents.
- Using citations as links, standard corpora can be viewed as a graph.
- The structure of this graph, independent of content, can provide interesting information about the similarity of documents and the structure of information.
- CF corpus includes citation information.

Impact Factor

- Developed by Garfield in 1972 to measure the importance (quality, influence) of scientific journals.
- Measure of how often papers in the journal are cited by other scientists.
- Computed and published annually by the Institute for Scientific Information (ISI).
- The *impact factor* of a journal *J* in year *Y* is the average number of citations (from indexed documents published in year *Y*) to a paper published in *J* in year *Y*–1 or *Y*–2.
- Does not account for the quality of the citing article.

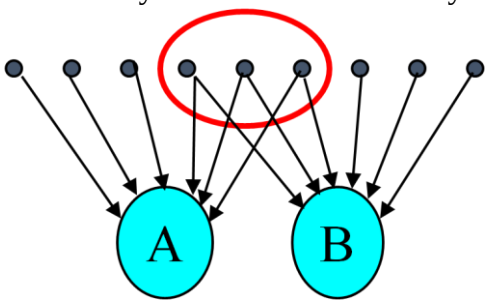
Bibliographic Coupling

- Measure of similarity of documents introduced by Kessler in 1963.
- The bibliographic coupling of two documents *A* and *B* is the number of documents cited by *both A* and *B*.
- Size of the intersection of their bibliographies.
- Maybe want to normalize by size of bibliographies?



Co-Citation

- An alternate citation-based measure of similarity introduced by Small in 1973.
- Number of documents that cite both *A* and *B*.
- Maybe want to normalize by total number of documents citing either *A* or *B* ?



Citations vs. Links

- Web links are a bit different than citations:
  - Many links are navigational.
  - Many pages with high in-degree are portals not content providers.
  - Not all links are endorsements.
  - Company websites don't point to their competitors.
  - Citations to relevant literature is enforced by peer-review.

Authorities

- *Authorities* are pages that are recognized as providing significant, trustworthy, and useful information on a topic.
- *In-degree* (number of pointers to a page) is one simple measure of authority.
- However in-degree treats all links as equal.
- Should links from pages that are themselves authoritative count more?

Hubs

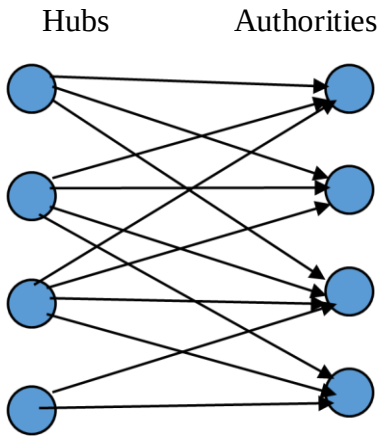
- *Hubs* are index pages that provide lots of useful links to relevant content pages (topic authorities).
  - Hub pages for IR are included in the course home page: http://www.cs.utexas.edu/users/mooney/ir-course

HITS

- Algorithm developed by Kleinberg in 1998.
- Attempts to computationally determine hubs and authorities on a particular topic through analysis of a relevant subgraph of the web.
- Based on mutually recursive facts:
  - Hubs point to lots of authorities.
  - Authorities are pointed to by lots of hubs.

Hubs and Authorities

- Together they tend to form a bipartite graph:

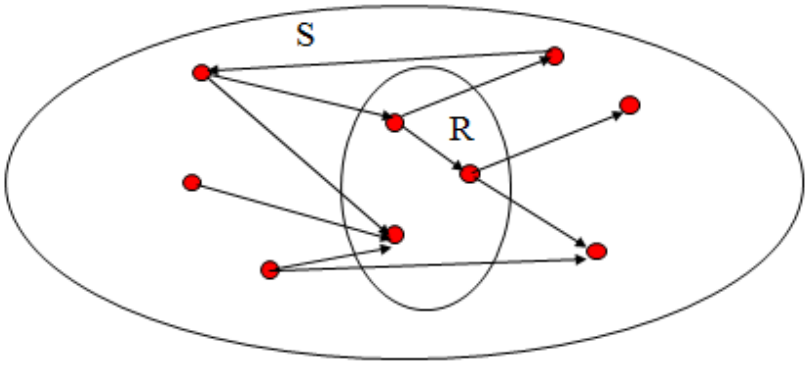


HITS Algorithm

- Computes hubs and authorities for a particular topic specified by a normal query.
- First determines a set of relevant pages for the query called the *base* set *S*.
- Analyze the link structure of the web subgraph defined by *S* to find authority and hub pages in this set.

Constructing a Base Subgraph

- For a specific query *Q*, let the set of documents returned by a standard search engine be called the *root* set *R*.
- Initialize *S* to *R*.
- Add to *S* all pages pointed to by any page in *R*.
- Add to *S* all pages that point to any page in *R*.



Base Limitations

- To limit computational expense:
  - Limit number of root pages to the top 200 pages retrieved for the query.
  - Limit number of "back-pointer" pages to a random set of at most 50 pages returned by a "reverse link" query.
- To eliminate purely navigational links:
  - Eliminate links between two pages on the same host.
- To eliminate "non-authority-conveying" links:
  - Allow only *m* (*m* ≅ 4–8) pages from a given host as pointers to any individual page.

Authorities and In-Degree

- Even within the base set *S* for a given query, the nodes with highest in-degree are not necessarily authorities (may just be generally popular pages like Yahoo or Amazon).

- True authority pages are pointed to by a number of hubs (i.e. pages that point to lots of authorities).

Iterative Algorithm

- Use an iterative algorithm to slowly converge on a mutually reinforcing set of hubs and authorities.
- Maintain for each page $p \in S$:
  - Authority score: a_p (vector $\boldsymbol{a}$)
  - Hub score: h_p (vector $\boldsymbol{h}$)
- Initialize all $a_p = h_p = 1$
- Maintain normalized scores:

$$\sum_{p \in S} (a_p)^2 = 1 \qquad \sum_{p \in S} (h_p)^2 = 1$$

HITS Update Rules

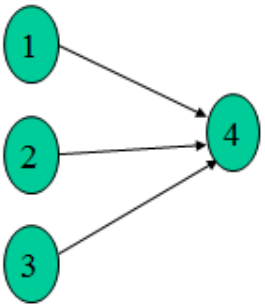
Authorities are pointed to by lots of good hubs:

$$a_p = \sum_{q:q \to p} h_q$$

Hubs point to lots of good authorities:

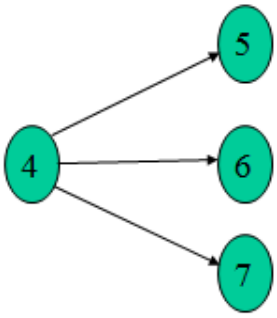
$$h_p = \sum_{q:p \to q} a_q$$

Illustrated Update Rules



$a_4 = h_1 + h_2 + h_3$

$h_4 = a_5 + a_6 + a_7$



HITS Iterative Algorithm

Initialize for all $p \in S$: $a_p = h_p = 1$

For i = 1 to k:

For all $p \in S$: $a_p = \sum_{q:q \to p} h_q$ *(update auth. scores)*

For all $p \in S$: $h_p = \sum_{q:p \to q} a_q$ *(update hub scores)*

For all $p \in S$: $a_p = a_p / c$ c: $\sum_{p \in S} (a_p / c)^2 = 1$ *(normalize* $\boldsymbol{a}$*)*

For all $p \in S$: $h_p = h_p / c$ c: $\sum_{p \in S} (h_p / c)^2 = 1$ *(normalize* $\boldsymbol{h}$*)*

Convergence

- Algorithm converges to a *fix-point* if iterated indefinitely.
- Define *A* to be the adjacency matrix for the subgraph defined by *S*.

- $A_{ij} = 1$ for $i \in S, j \in S$ iff $i \rightarrow j$

- Authority vector, $\boldsymbol{a}$, converges to the principal eigenvector of $A^T A$
- Hub vector, $\boldsymbol{h}$, converges to the principal eigenvector of AA^T
- In practice, 20 iterations produces fairly stable results.

Finding Similar Pages Using Link Structure

- Given a page, *P*, let *R* (the root set) be *t* (e.g. 200) pages that point to *P*.
- Grow a base set *S* from *R*.
- Run HITS on *S*.
- Return the best authorities in *S* as the best similar-pages for *P*.
- Finds authorities in the "link neighbor-hood" of *P*.

Similar Page Results

- Given "honda.com"
  - toyota.com
  - ford.com
  - bmwusa.com
  - saturncars.com
  - nissanmotors.com
  - audi.com
  - volvocars.com

HITS for Clustering

- An ambiguous query can result in the principal eigenvector only covering one of the possible meanings.
- Non-principal eigenvectors may contain hubs & authorities for other meanings.
- Example: "jaguar":
  - Atari video game (principal eigenvector)
  - NFL Football team (2^{nd} non-princ. eigenvector)
  - Automobile (3^{rd} non-princ. eigenvector)

PageRank

- Alternative link-analysis method used by Google (Brin & Page, 1998).
- Does not attempt to capture the distinction between hubs and authorities.
- Ranks pages just by authority.
- Applied to the entire web rather than a local neighborhood of pages surrounding the results of a query.

Google Ranking

- Complete Google ranking includes (based on university publications prior to commercialization).
  - Vector-space similarity component.
  - Keyword proximity component.

  - HTML-tag weight component (e.g. title preference).
  - PageRank component.
- Details of current commercial ranking functions are trade secrets.

Personalized PageRank

- PageRank can be biased (personalized) by changing **E** to a non-uniform distribution.
- Restrict "random jumps" to a set of specified relevant pages.
- For example, let $E(p) = 0$ except for one's own home page, for which $E(p) = \alpha$
- This results in a bias towards pages that are closer in the web graph to your own homepage.

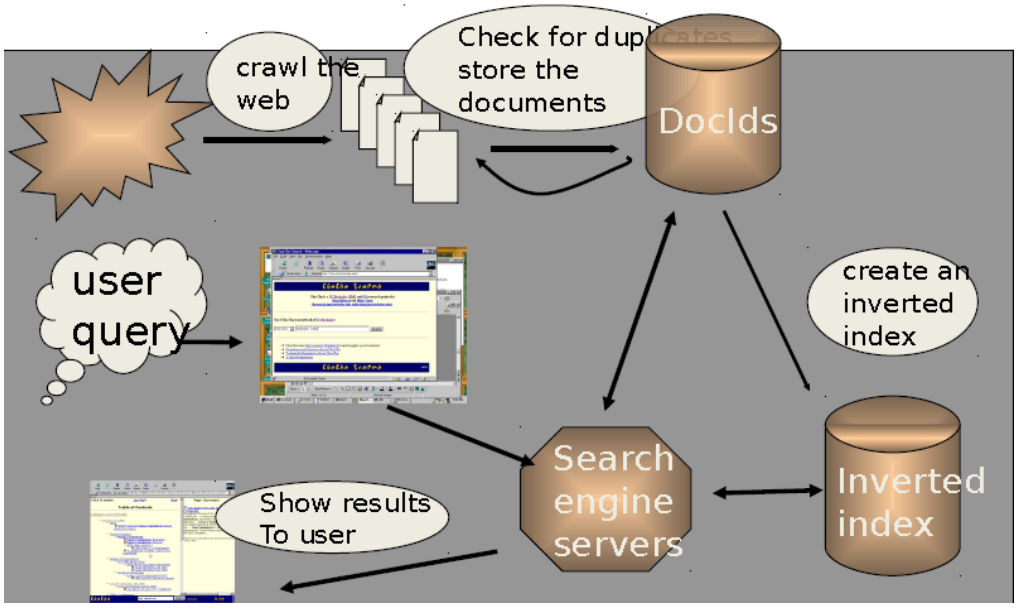
Link Analysis Conclusions

- Link analysis uses information about the structure of the web graph to aid search.
- It is one of the major innovations in web search.
- It was one of the primary reasons for Google's initial success.

7.3 Searching the web

• Search engines determine how to rank pages using automated methods that look at the Web itself,
• Without any help from external source of knowledge
• There are enough information intrinsic to the Web and its structure to figure this out
• Example: go to Google and type "u2," the first result it shows you is www.u2.com the home page of rock band U2.
• How did Google "know" that this was the best answer?

Standard Web Search Engine Architecture



Historical Background –Searching the web

• Precursor: automated information retrieval systems of 1960s – designed to search repositories of newspaper articles, scientific papers, patents, legal abstracts, and other document collections in response to keyword queries.
• Information retrieval systems have always had to deal with the problem that keywords are a very limited way to express a complex information need;

• it suffered from the problems of:
– Synonymy (multiple ways to say the same thing)
– Polysemy (multiple meanings for the same term)

• With the arrival of the Web, where everyone is an author and everyone is a searcher, the problems surrounding information retrieval exploded in scale and complexity
• Diversity in authoring styles makes it much harder to rank documents according to a common criterion
• a single topic, one can easily find pages written by experts, novices, children, conspiracy theorists—and not necessarily be able to tell which is which
• There is a correspondingly rich diversity in the set of people issuing queries, and the problem of multiple meanings becomes particularly severe
• But the Web also introduces new kinds of problems:
1. Dynamic and constantly-changing nature of Web content&
2. The genuinely of the content

Web Crawling
How do the web search engines get all of the items they index?
Main idea:
– Start with known sites
– Record information for these sites
– Follow the links from each site
– Record information found at new sites
– Repeat

Web Crawlers
How do the web search engines get all of the items they index?
More precisely:
– Put a set of known sites on a queue
– Repeat the following until the queue is empty:
    • Take the first page off of the queue
    • If this page has not yet been processed:
    - Record the information found on this page: Positions of words, links going out, etc
    - Add each link on the current page to the queue
    - Record that this page has been processed
• In what order should the links be followed?

    - Page Visit Order: breadth first or depth first search on trees

Web Crawling Issues
• Keep out signs: A file called robots.txt tells the crawler which directories are off limits
• Freshness:
– Figure out which pages change often
– Recrawl these often
• Duplicates, virtual hosts, etc
– Convert page contents with a hash function
– Compare new pages to the hash table
• Lots of problems
– Server unavailable
– Incorrect html
– Missing links

– Infinite loops
• Web crawling is difficult to do robustly!

Directories vs. Search Engines
• Web Directories:
– Hand-selected sites
– Search over the contents of the descriptions of the pages
– Organized in advance into categories

• Search Engines:
– All pages in all sites
– Search over the contents of the pages themselves
– Organized after the query by relevance rankings or other scores

7.4 Navigating the web

Web navigation refers to the process of navigating a network of information resources in the World Wide Web, which is organized as a hypertext or hypermedia. The user interface that is used to do so is called a web browser. A central theme in web design is the development of a web navigation interface that maximizes usability. The first and most important way of navigation is using a search engine. A search engine is a web site that allows you to find web pages based on the words you enter or 'search query' this can be a question, or just a collection of words that you think will be on the web page you're after.

Navigation
Principles for good navigation design: A site must:
1. Let me know where I am at all times
2. Clearly differentiate hyperlinks from content
3. Let me know clearly where I can go from here
4. Let me see where I've already been
5. Make it obvious what to do to get somewhere
6. Indicate what clicking a link will do

Good navigation
• Easily learned
• Consistent: – In terms of their placement, offerings, and appearance
• Provides feedback
• Requires an "economy of action and time"
• Users understandable labels
• Is appropriate to site's purpose
• Supports users' goals
• Provide contextual clues and flexibility

Where should you put navigation?
• Depends on the type of navigation
• The golden rules are:
– Put the most useful navigation where it's closest to hand
– Put navigation where the user is likely to look for it.

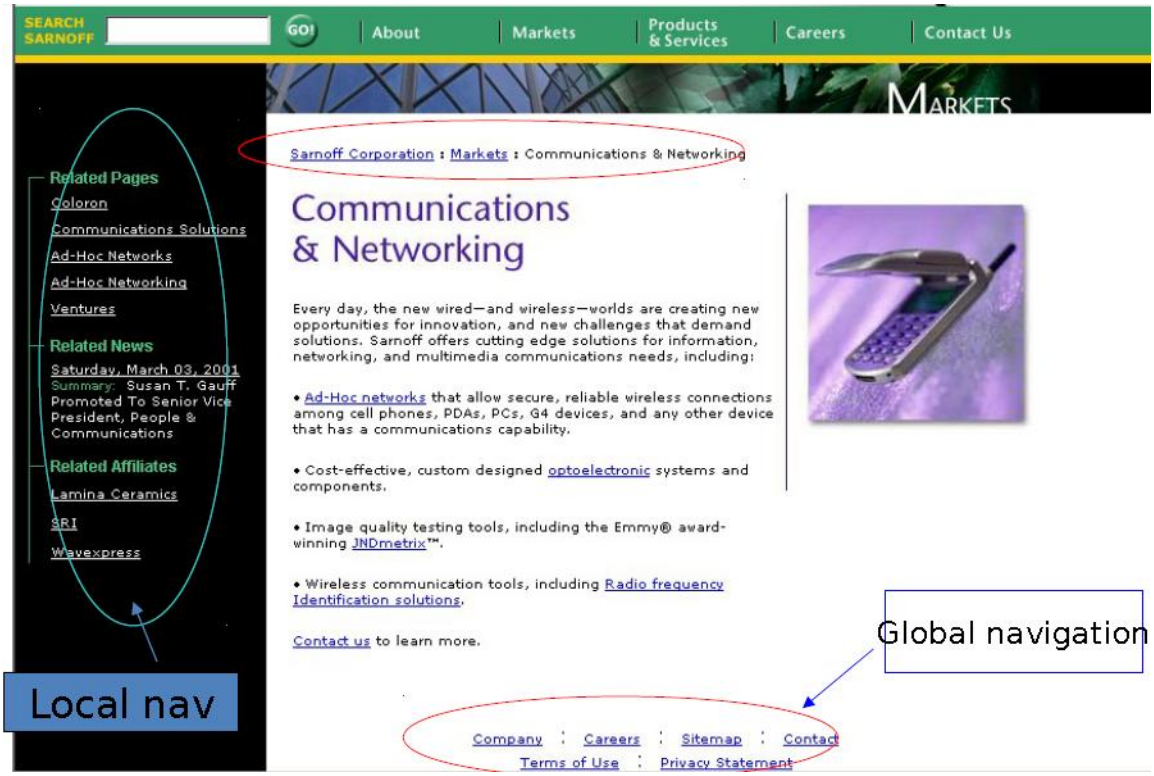
Navigation Models: common navigation conventions.
– List of contents

– Breadcrumb trail
– Horizontal top bar
– Tabs
– 2-level top (bar or tabs)
– Top and side bars
– Buttons bar with revealed drop-down
– Multiple-level tree navigation
– Paging

Types of navigation systems

• Hierarchical navigation systems
– Similar to the information hierarchy
– Require additional navigation systems

• Global navigation systems (site-wide navigation systems)
– Able to jump back to the main page
– Important to extend the global navigation system throughout the sub-site
• Local navigation systems
• Ad hoc navigation: – Embedded links



Breadcrumb trail

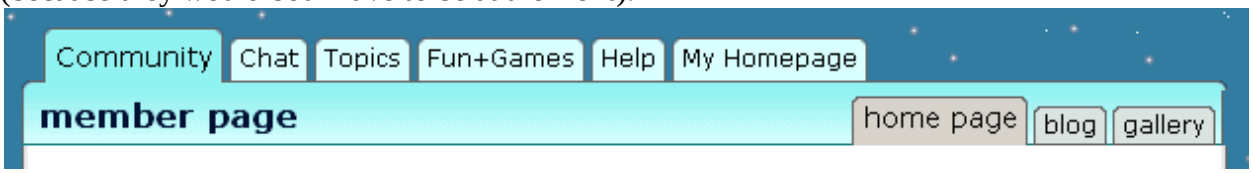
• The breadcrumb trail is the familiar navigation device that:
– Shows you where you in are in a hierarchy, and
– Lets you click back to any point above where you are now
• Breadcrumb trails are great in situations where:
– You've got a particularly deep hierarchy, say four levels or more
– The possible flow is such that a typical user might want or need to get back to a specific previous place

Horizontal top bar

Tabs

• Some extra advantages over a line of links:
– They serve to show the active section/selection very clearly
– They naturally have a working visual hierarchy, with a real-world connection that makes them extremely clear. A tab is normally attached to (part of) a folder or sheet in a binder, and physically labels everything in the folder, or on the sheet.

– They are unambiguously mutually-exclusive. It's physically impossible to have two tabs selected (because they would both have to be at the front).



Top and side bars

• Very common. The top bar is used for the site-level navigation/tools and often first-level navigation, because these are more fixed.

Navigation bar with revealed dropdowns

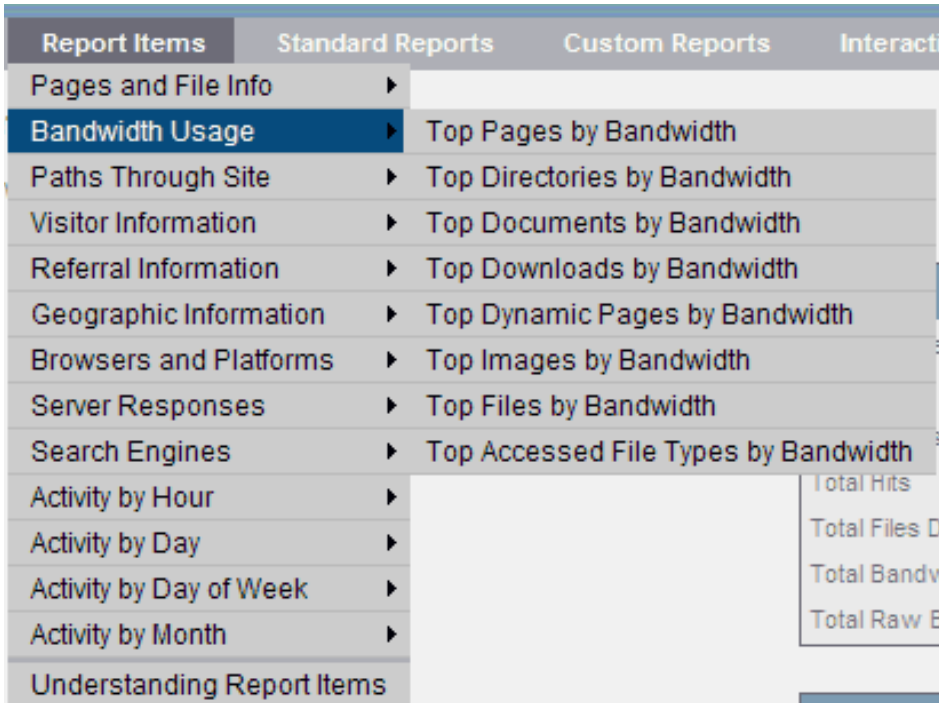


Fig. Buttons bar with revealed drop-down

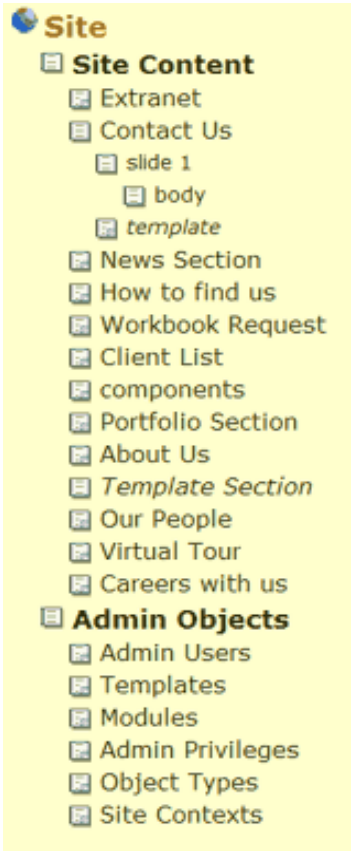


Fig. Multiple-level tree navigation

Multiple-level tree navigation

• Benefits

– It is relatively familiar and intuitive (provided it is presented in a conventional format).

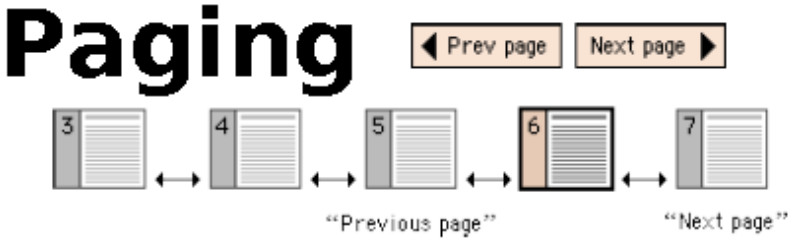
– It can provide relatively simple access to a complex structure.

`

Paging

Where you get a piece of content that spans several pages (typically long articles, long indexes, forums, or search results). By providing paging buttons and links to local home pages and contents pages you give users the tools to understand how you have organized your Web site information, even if they have not

entered your Web of pages through a home page or contents page. The buttons don't prevent people from reading the information in whatever order they choose, but they do allow readers to follow the sequence of pages you have laid out.



Showing results **91 to 100** of **107**

Previous | 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - **10** | Next

Remote Navigation Elements

• Provide an alternative bird's-eye view of the site's content
– Tables of contents
– Site Map
– Index

Book marking

When you find a web page that you find interesting, instead of having to remember the address of the webpage, you can simply save the address as a 'bookmark' in your browser. This means you can come back and find that exact web page again. There are websites that allow you to save your bookmarks to an account on the website, instead of to your browser. This means that you can access your bookmarks from anywhere–you don't have to be on the computer that you first found the site on. Also, you can make your bookmarks public, meaning other people can see the interesting stuff you've found –called 'social bookmarking'

When you find something interesting on the internet, you can also share it with your friends.

- One way to do this is to email your friends with a link to the particular web page, however this is increasingly being replaced by social media services.
- When lots of people share something with their friends, and then their friends share it on, and so on, this is called a viral or in some cases an internet meme

Hyperlinks: these links provide a connection between Web pages that allows for amazingly easy access to other Web pages. A link or hyperlink can be text, an icon, a picture, or an icon that moves a user from one Web page or Web site to another. A hyperlink (often underlined, colored etc.) has an unseen Web address imbedded in it.

Navigating Within a Web Page

• There are convenient ways to move around that particular page itself.
• Often a Web page holds more information than can fit on one screen.
• A Web page appears aligned to the upper left hand corner of your screen.
• There is often information that you cannot see farther down after the last line on the screen.
• Sometimes there is also more information to the right of the screen.

Slider & Arrows

• Scrolling is an easy way to navigate on a Web page.
• You can scroll up and down and side to side by using either the horizontal or vertical onscreen scroll bars on the bottom and right side of the screen.

• To scroll using the onscreen scroll bars, simply position your cursor on the slider on the scroll bar.
• Hold the mouse button down and drag the slider up and/or down on the vertical scroll bar (or side to side on the horizontal scroll bar).
• You can also position your cursor over the arrows at the top and the bottom of the vertical scroll bar (left and right sides of the horizontal scroll bar) to move one line at a time.
• Using Arrow Keys
– The keyboard holds some other choices for helping you move around a Web page.
– The first are the Page Up and Page Down keys on your keyboard.
• Using Mouse
• Other important components:
– URL Bar/Address Bar/Location Bar
– Home
– Back and Forward
– Refresh
– sitemap
– text links
– Tabs for more navigation etc.

• **The purpose of navigation is to:**
1. Present readers with the most user-friendly path through the classification so that they can find the content they want quickly.
2. Ensure readers always know where they are on the site.
3. Allow readers to move quickly and logically through the web site.
4. Give readers the proper context of the document they are reading.
5. Highlight for the reader parts of the classification that the organization wants to promote.

• a golden rule of navigation design:
– Start your design from the reader's point of view.
– Get the people who will actually use the web site involved in the design from the earliest point possible.
• Website navigation is important to the success of website visitor's experience to the website.
• The website's navigation system is like a road map to all the different areas and information contained within the website.

7.5 Web Uses mining

• **Web mining -** is the application of data mining techniques to discover patterns from the web. Web mining can be divided into three different types: –**Web usage mining, Web content mining** and **Web structure mining.** Process of extracting useful information from server logs. e.g. use Web usage mining is the process of finding out what users are looking for on the Internet.
• **Goal**: analyze the behavioral patterns and profiles of users interacting with a Web site
The discovered patterns are usually represented as collections of pages, objects, or resources that are frequently accessed by groups of users with common interests. Some users might be looking at only textual data whereas some others might be interested in multimedia data. Web Usage Mining is the application of data mining techniques to discover interesting usage patterns from web data. Understands and better serve the needs of Web-based applications. Usage data captures the identity or origin of Web users along with their browsing behavior at a Web site
• Web usage mining itself can be classified further depending on the kind of usage data considered:
1. Web Server Data:
• The user logs are collected by the Web server
• Typical data includes IP address, page reference and access time

2. Application Server Data:
• Commercial application servers have significant features to enable e-commerce applications to be built on top of them with little effort
• A key feature is the ability to track various kinds of business events and log them in application server logs
3. Application Level Data:
• New kinds of events can be defined in an application, and logging can be turned on for them thus generating histories of these specially defined events
• It must be noted, however, that many end applications require a combination of one or more of the techniques applied in the categories above

Why Web Usage Mining?
• Explosive growth of E-commerce
– Provides an cost-efficient way doing business
– Amazon.com: "online Wal-Mart"
• Hidden Useful information
– Visitors' profiles can be discovered
– Measuring online marketing efforts, launching marketing campaigns, etc.

How to perform Web Usage Mining

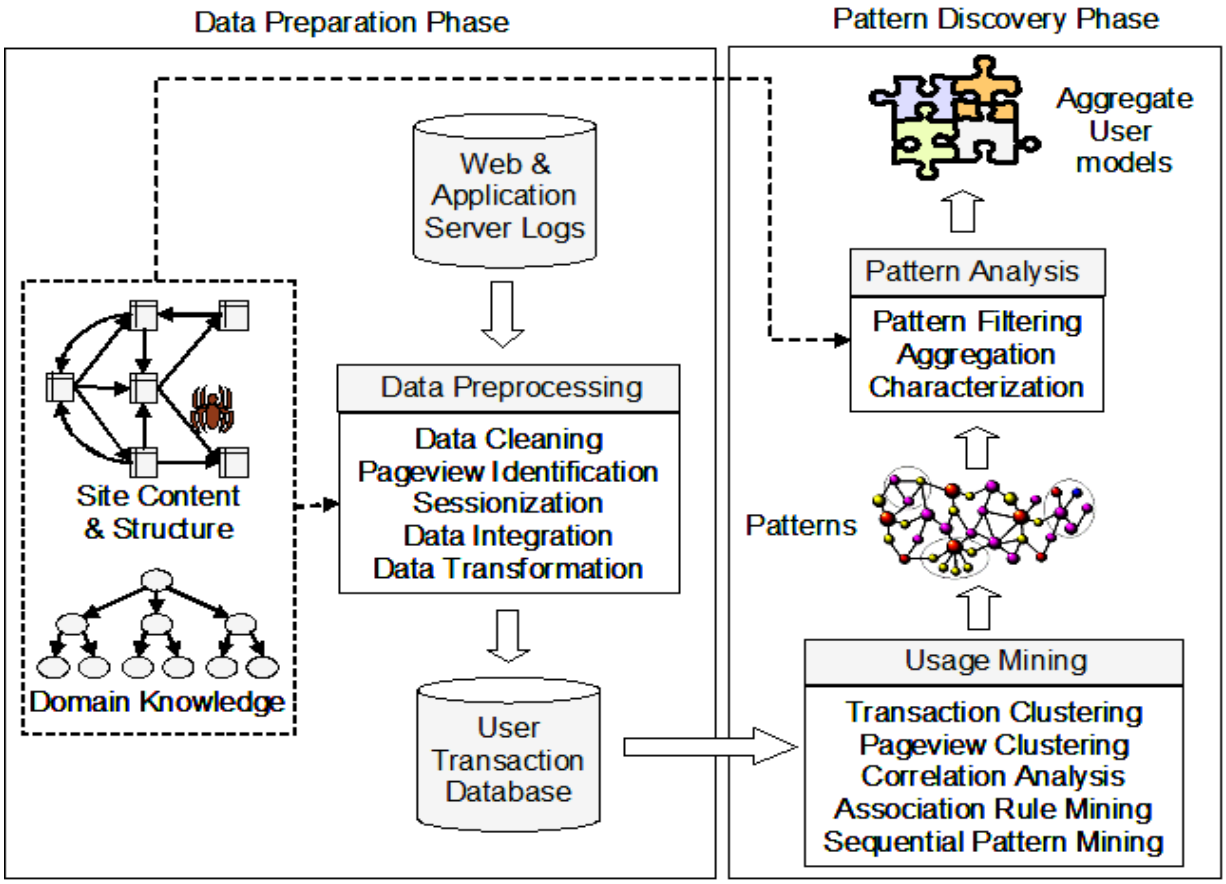
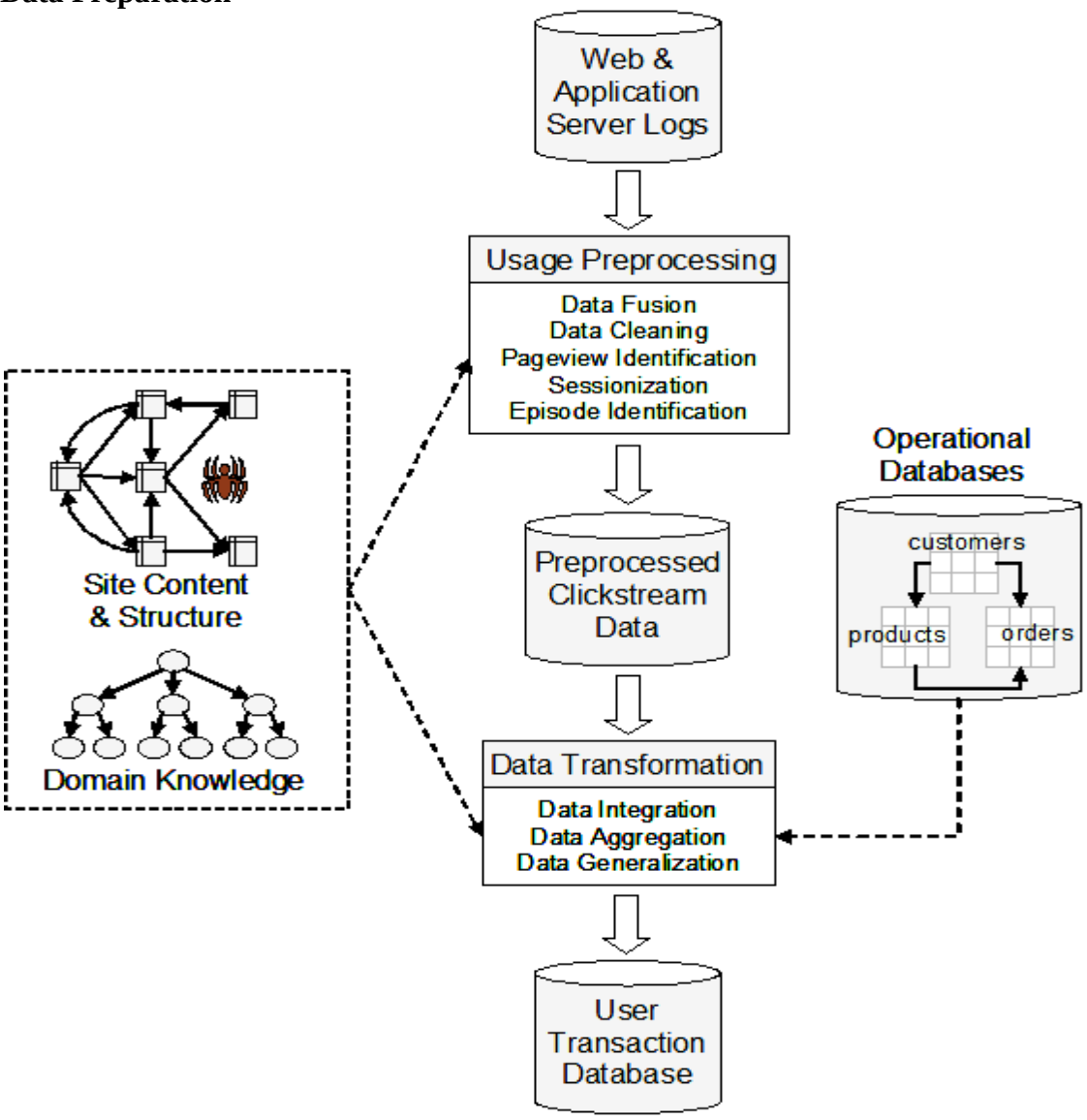


Figure: Web usage mining Process

• Obtain web traffic data from
– Web server log files
– Corporate relational databases

– Registration forms
• Apply data mining techniques and other Web mining techniques
• Two categories:– Pattern Discovery Tools and Pattern Analysis Tools

Data Preparation



Data cleaning
– remove irrelevant references and fields in server logs
– remove references due to spider navigation
– remove erroneous references
– add missing references due to caching

Pageview: A pageview is an aggregate representation of a collection of Web objects contributing to the display on a user's browser resulting from a single user action (such as a click-through). Conceptually, each pageview can be viewed as a collection of Web objects or resources representing a specific "user event," e.g., reading an article, viewing a product page, or adding a product to the shopping cart.

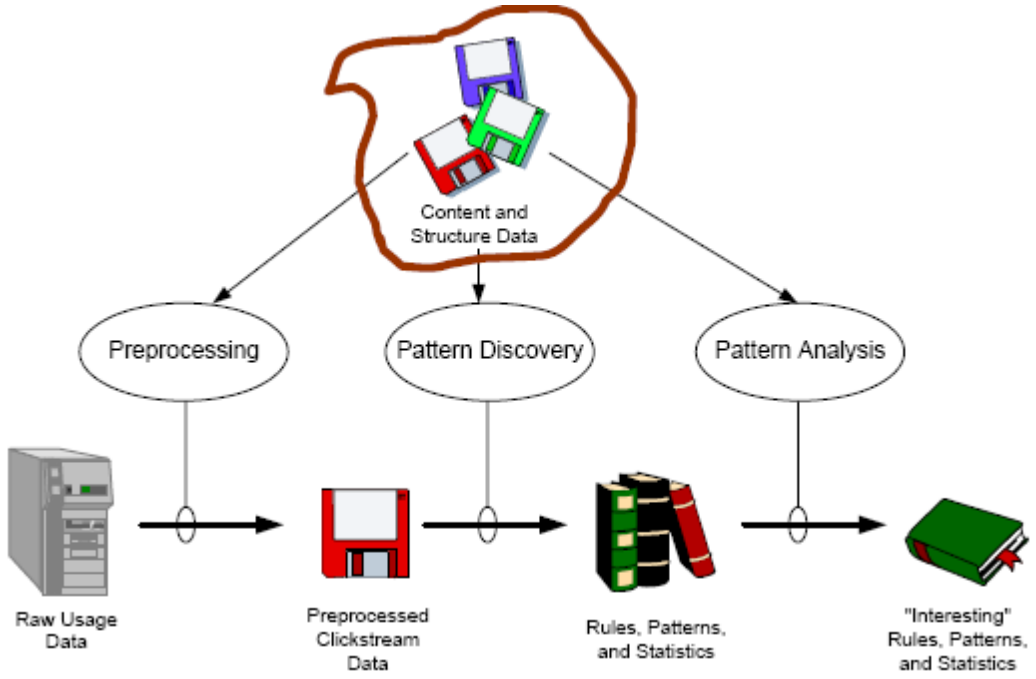
Integrating with e-commerce events

• Either product oriented or visit oriented

• Used to track and analyze conversion of browsers to buyers. Major difficulty for E-commerce events is defining and implementing the events for a site, however, in contrast to click stream data, getting reliable preprocessed data is not a problem.

• Another major challenge is the successful integration with click stream data

Web usage mining process



Advantages : Web Usage mining

• Allows companies to produce productive information pertaining to the future of their business function ability

• Some of this information can be derived from the collective information of lifetime user value, product cross marketing strategies and promotional campaign effectiveness

• The usage data that is gathered provides the companies with the ability to produce results more effective to their businesses and increasing of sales

• To develop marketing skills that will out-sell the competitors and promote the company's services or product on a higher level

• To aid in e-businesses whose business is based solely on the traffic provided through search engines

• The use of this type of web mining helps to gather the important information from customers visiting the site.

• This enables an in-depth log to complete analysis of a company's productivity flow

• E-businesses depend on this information to direct the company to the most effective Web server for promotion of their product or service

• To provide the best access routes to services or other advertisements

• When a company advertises for services provided by other companies, the usage mining data allows for the most effective access paths to these portals

• To provide the companies with the information needed to provide an effective presence to their customers

• This collection of information may include user registration, access logs and information leading to better Web site structure, proving to be most valuable to company online marketing

• These present some of the benefits for external marketing of the company's products , services and overall management
• To provide information for improvement of communication through intranet communications
• Developing strategies through this type of mining will allow for intranet based company databases to be more effective through the provision of easier access paths
• The projection of these paths helps to log the user registration information giving commonly used paths the forefront to its access
• To keep a record of fraudulent payments which can all be researched and studied through data mining
• This information can help develop more advanced and protective methods that can be undertaken to prevent such events from happening
• To foster marketing of businesses and a direct impact to the success of their promotional strategies and internet traffic.
• This information is gathered on a daily basis and continues to be analyzed consistently
• Analysis of this pertinent information will help companies to develop promotions that are more effective, internet accessibility, inter-company communication and structure, and productive marketing skills through web usage mining

Disadvantages
• Invasion of privacy
• Misuse of data (collected for different purpose n misused for different purpose)
• Some mining algorithms might use controversial attributes like sex, race, religion, or sexual orientation to categorize individuals

Applications: Web Usage mining: Major application areas for web usage mining

1. Personalization
2. System improvement
3. Site modification
4. Business intelligence
5. Usage characterization

7.6 Collaborative filtering: CF

In everyday life, we rely on recommendations from other people either by word of mouth, recommendation letters, movie and book reviews printed in newspapers. Recommendation systems are a subclass of information filtering system that seek to predict the 'rating' or 'preference' that user would give to an item (such as music, books, or movies) or social element (e.g. people or groups) they had not yet considered, using a model built from the characteristics of an item (content-based approaches) or the user's social environment (collaborative filtering approaches). Recommender systems have become extremely common in recent years. Example: – When viewing a product on Amazon.com, the store will recommend additional items based on a matrix of what other shoppers bought along with the currently selected item.
Recommender systems typically produce a list of recommendations in one of two ways:

1. Collaborative filtering approaches to build a model from a user's past behavior (items previously purchased or selected and/or numerical ratings given to those items) as well as similar decisions made by other users, then use that model to predict items (or ratings for items) that the user may have an interest in.
2. Content-based filtering approaches utilize a series of discrete characteristics of an item in order to recommend additional items with similar properties.

3. These approaches 1 and 2 are often combined (called as Hybrid Recommender Systems).

Recommender systems are a useful alternative to search algorithms since they help users discover items they might not have found by themselves. Interestingly enough, recommender systems are often implemented using search engines indexing non-traditional data. Recommender system is an active research area in the data mining and machine learning areas. They are costly to maintain

7.7 Collective Intelligence

Most complex systems, not only can be, but need to be viewed as collectives. Examples include:
– Control of a constellation of communication satellites
– Routing data/vehicles over a communication network/highway
– Dynamic data migration over large distributed databases
– Dynamic job scheduling across a (very) large computer grid
– Coordination of rovers/submersibles on Mars/ Europa
– Control of the elements of an amorphous computer/telescope

Collective Intelligence – the aggregated knowledge, insight and expertise of a diverse group –has become a reality. As individuals become more adept and comfortable sharing thoughts and ideas in virtual spaces, companies can use these insights to address critical business challenges. Harnessing Collective Intelligence can play an important role in generating new ideas, solving age-old problems, disaggregating and distributing work in new and innovative ways, and making better, more informed decisions about the future.

• COLLECTIVE: refers to any entity constituted by other entities. In this case, it usually refers to human social entities such as groups, organizations and communities.
• INTELLIGENCE AS A CAPACITY: Intelligence is variously defined as "the capacity to acquire and apply knowledge," "the ability to effectively adapt," or simply "the ability to solve problems." Collective intelligence is an emergent property of collective social systems. It is a natural product of the independent opinions or behaviors of diverse individuals or groups in a decentralized system (flock, market, guessing game) that aggregates those opinions or behaviors.

Collective intelligence is a theory that describes a type of shared or group intelligence that emerges from the collaboration and competition of many individuals and appears in consensus decision making in bacteria [clarification needed], animals, and computer networks [citation needed]. The term appears in sociobiology, political science and in context of mass peer review and crowd sourcing applications. Everything from a political party to a public wiki can reasonably be described as this loose form of collective intelligence.
• Primitive examples: families, companies, and countries etc.
• A precursor of the concept is found in entomologist –ant colony optimization.
• Collective intelligence is mass collaboration. In order for this concept to happen, four principles need to exist:
– **Openness**: Sharing ideas and intellectual property: though these resources provide the edge over competitors more benefits accrue from allowing others to share ideas and gain significant improvement and scrutiny through collaboration
– **Peering:** Horizontal organization as with the ‘opening up’ of the Linux program where users are free to modify and develop it provided that they make it available for others. Peering succeeds because it encourages self-organization – a style of

– **Sharing:** Companies have started to share some ideas while maintaining some degree of control over others, like potential and critical patent rights. Limiting all intellectual property shuts out opportunities, while sharing some expands markets and brings out products faster.

– Acting Globally: The advancement in communication technology has prompted the rise of global companies at low overhead costs. The internet is widespread, therefore a globally integrated company has no geographical boundaries and may access new markets, ideas and technology. In this context collective intelligence is often confused with shared knowledge. The former is knowledge that is generally available to all members of a community while the latter is information known by all members of a community.

Examples:

• Google takes the collective knowledge created by millions of people making websites for other purposes and harnesses that collective knowledge—using some very clever algorithms and sophisticated technology—to produce amazingly intelligent answers to the questions we type in.

• Wikipedia, at another extreme, uses much less sophisticated technology, but some very clever organizational principles and motivational techniques, to get thousands of people all over the world to volunteer their time to create an amazing online collection of knowledge.

• Social bookmarking: In social bookmarking (also called collaborative tagging), users assign tags to resources shared with other users, which gives rise to a type of information organization that emerges from this crowd sourcing process.

• Stock market predictions: Because of the Internet's ability to rapidly convey large amounts of information throughout the world, the use of collective intelligence to predict stock prices and stock price direction has become increasingly viable. Websites aggregate stock market information that is as current as possible so professional or amateur stock analysts can publish their viewpoints, enabling amateur investors to submit their financial opinions and create an aggregate opinion.

Benefits of collective intelligence to business

• Talent Utilization: At the rate technology is changing, no firm can fully keep up in the innovations needed to compete. Instead, smart firms are drawing on the power of mass collaboration to involve participation of the people they could not employ.

• Demand Creation: Firms can create a new market for complementary goods by engaging in open source community.

• Costs Reduction: Mass collaboration can help to reduce costs dramatically. Firms can release a specific software or product to be evaluated or debugged by online communities. The results will be more personal, robust and error-free products created in a short amount of time and costs.

The key question: How can people and computers be connected so that collectively they act more intelligently than any individual, group, or computer has ever done before?

• Programming Collective Intelligence takes you into the world of machine learning and statistics, and explains how to draw conclusions about user experience, marketing, personal tastes, and human behavior in general -- all from information that you and others collect every day. Each algorithm is described clearly and concisely with code that can immediately be used on your web site, blog, Wiki, or specialized application.

• Popular methods:

– Collaborative filtering techniques that enable online retailers to recommend products or media

– Methods of clustering to detect groups of similar items in a large dataset

– Search engine features -- crawlers, indexers, query engines, and the PageRank algorithm etc.