Information Systems Engineering & Information Retrieval Handbook

Renay Maher

Lakenya Lai

Revised Edition: 2014

ISBN 978-81-323-0943-7

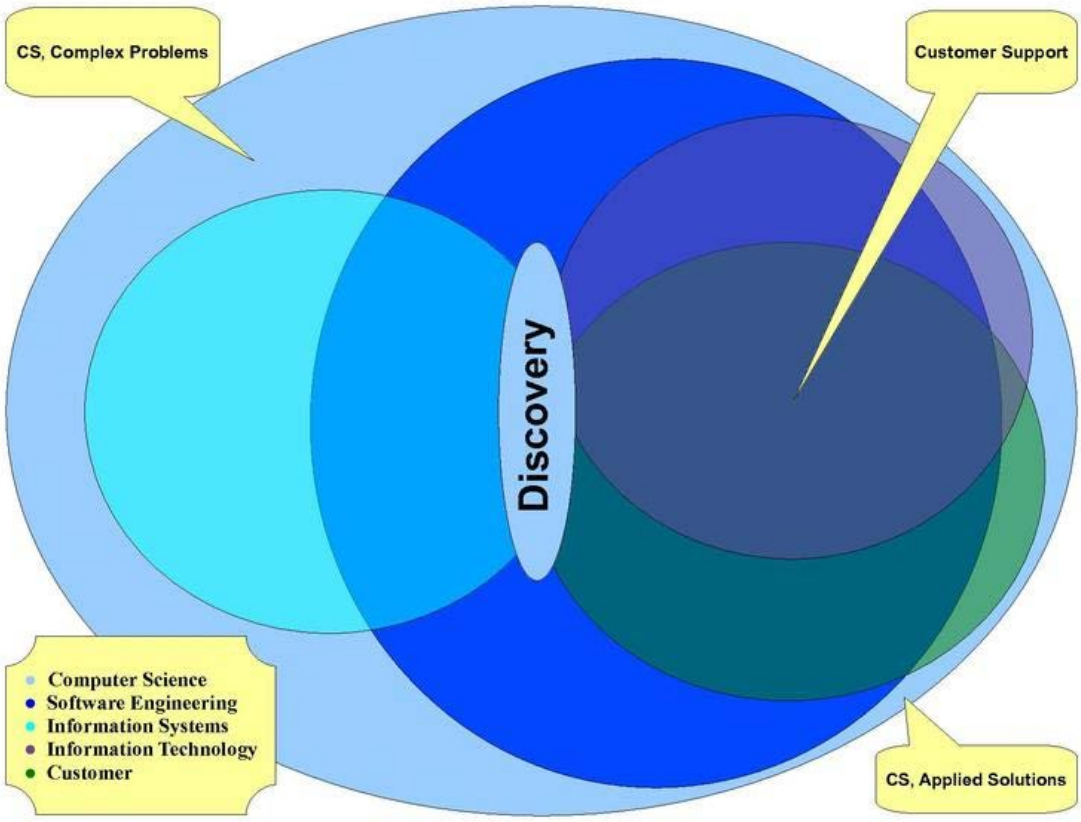
Published by:
Academic Studio
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter-1

Information Systems & Information Engineering

Information Systems



CS, SE, IS, IT, & Customer Venn Diagram where functionality spans left and design spans right stemming from discovery.

Information Systems (IS) is an academic/professional discipline bridging the business field and the well-defined computer science field that is evolving toward a new scientific area of study. An information systems discipline therefore is supported by the theoretical foundations of information and computations such that learned scholars have unique opportunities to explore the academics of various business models as well as related algorithmic processes within a computer science discipline. Typically, information systems or the more common *legacy* information systems include people, procedures, data, software, and hardware (by degree) that are used to gather and analyze digital information. Specifically computer-based information systems are complementary networks of hardware/software that people and organizations use to collect, filter, process, create, & distribute data (computing). *Computer* Information System(s) (CIS) is often a track within the computer science field studying computers and algorithmic processes, including their principles, their software & hardware designs, their applications, and their impact on society. Overall, an IS discipline emphasizes functionality over design.

As illustrated by the Venn Diagram on the right, the history of **information systems** coincides with the history of computer science that began long before the modern discipline of computer science emerged in the twentieth century. Regarding the circulation of information and ideas, numerous legacy information systems still exist today that are continuously updated to promote ethnographic approaches, to ensure data integrity, and to improve the social effectiveness & efficiency of the whole process. In general, information systems are focused upon processing information within organizations, especially within business enterprises, and sharing the benefits with modern society.

Overview

Silver et al. (1995) provided two views on (IS) and IS-centered view that includes software, hardware, data, people, and procedures. A second managerial view includes people, business processes and Information Systems.

There are various types of information systems, for example: transaction processing systems, office systems, decision support systems, knowledge management systems, database management systems, and office information systems. Critical to most information systems are information technologies, which are typically designed to enable humans to perform tasks for which the human brain is not well suited, such as: handling large amounts of information, performing complex calculations, and controlling many simultaneous processes.

Information technologies are a very important and malleable resource available to executives. Many companies have created a position of Chief Information Officer (CIO) that sits on the executive board with the Chief Executive Officer (CEO), Chief Financial Officer (CFO), Chief Operating Officer (COO) and Chief Technical Officer (CTO).The CTO may also serve as CIO, and vice versa. The Chief Information Security Officer (CISO), who focuses on information security management.

Definition

Silver et al. defined Information Systems as follows:

Information systems are implemented within an organization for the purpose of improving the effectiveness and efficiency of that organization. Capabilities of the information system and characteristics of the organization, its work systems, its people, and its development and implementation methodologies together determine the extent to which that purpose is achieved

The Discipline of Information Systems

Several IS scholars have debated the nature and foundations of Information Systems which has its roots in other reference disciplines such as Computer Science, Engineering, Mathematics, Management Science, Cybernetics, and others

The Impact on Economic Models

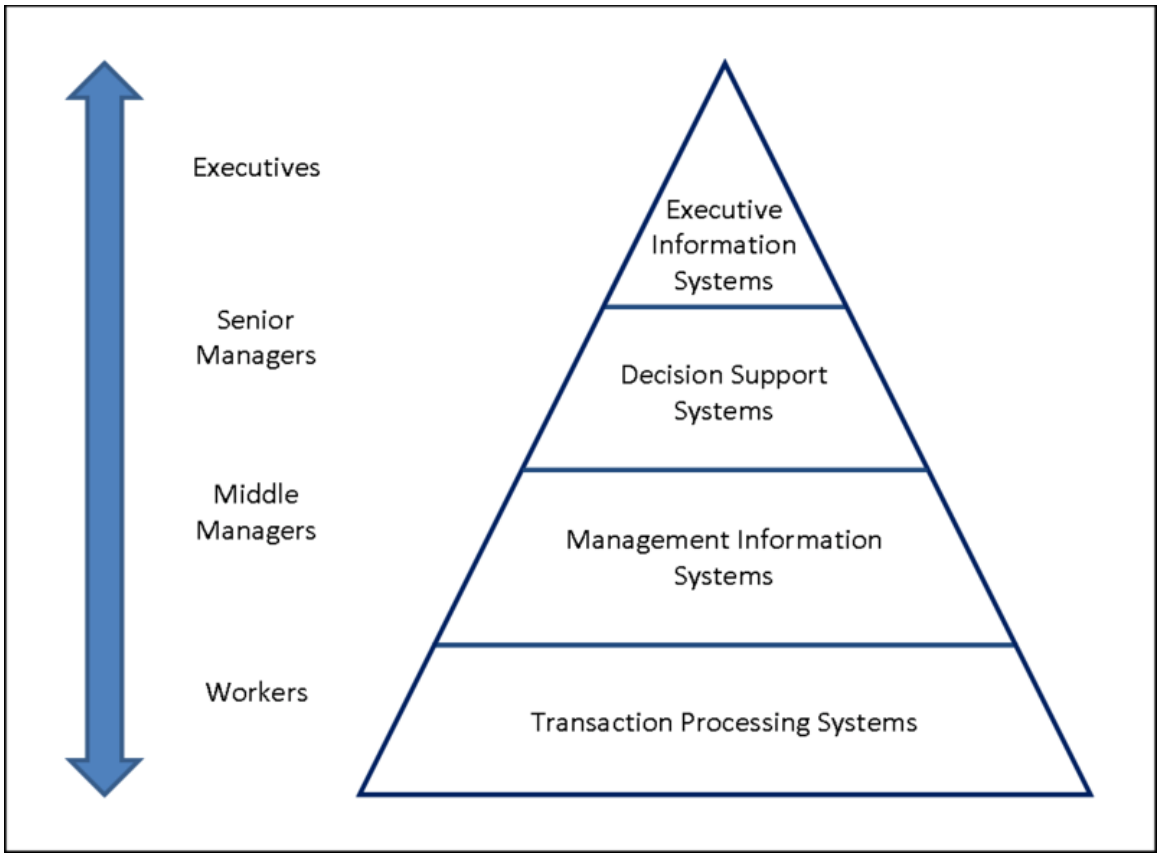
- Microeconomic theory model
- Transaction Cost theory
- Agency Theory

Differentiating IS from Related Disciplines

Similar to computer science, other disciplines can be seen as both related disciplines and foundation disciplines of IS. But, while there may be considerable overlap of the disciplines at the boundaries, the disciplines are still differentiated by the focus, purpose and orientation of their activities.

In a broad scope, the term **Information Systems** (IS) is a scientific field of study that addresses the range of strategic, managerial and operational activities involved in the gathering, processing, storing, distributing and use of information, and its associated technologies, in society and organizations. The term information systems is also used to describe an organizational function that applies IS knowledge in industry, government agencies and not-for-profit organizations. ***Information Systems*** often refers to the interaction between algorithmic processes and technology. This interaction can occur within or across organizational boundaries. An information system is not only the technology an organization uses, but also the way in which the organizations interact with the technology and the way in which the technology works with the organization's business processes. Information systems are distinct from information technology (IT) in that an information system has an information technology component that interacts with the processes components.

Types of information systems



A four level pyramid model of different types of Information Systems based on the different levels of hierarchy in an organization

The 'classic' view of Information systems found in the textbooks of the 1980s was of a pyramid of systems that reflected the hierarchy of the organization, usually Transaction processing systems at the bottom of the pyramid, followed by Management information systems, Decision support systems and ending with Executive information systems at the top. Although the pyramid model remains useful, since it was first formulated a number of new technologies have been developed and new categories of information systems have emerged, some of which no longer fit easily into the original pyramid model.

Some examples of such systems are:

- Data warehouses
- Enterprise resource planning
- Enterprise systems
- Expert systems
- Geographic information system
- Global information system
- Office Automation

Information systems career pathways

Information Systems have a number of different areas of work:

- Information systems strategy
- Information systems management
- Information systems development
- Information systems security
- Information systems iteration
- Information system organization

There are a wide variety of career paths in the information systems discipline. "Workers with specialized technical knowledge and strong communications skills will have the best prospects. Workers with management skills and an understanding of business practices and principles will have excellent opportunities, as companies are increasingly looking to technology to drive their revenue."

Information systems development

Information technology departments in larger organizations tend to strongly influence information technology development, use, and application in the organizations, which may be a business or corporation. A series of methodologies and processes can be used in order to develop and use an information system. Many developers have turned and used a more engineering approach such as the System Development Life Cycle (SDLC) which is a systematic procedure of developing an information system through stages that occur in sequence. An Information system can be developed in house (within the organization) or outsourced. This can be accomplished by outsourcing certain components or the entire system. A specific case is the geographical distribution of the development team (Offshoring, Global Information System).

A computer based information system, following a definition of Langefors, is:

- a technologically implemented medium for recording, storing, and disseminating linguistic expressions,
- as well as for drawing conclusions from such expressions.

which can be formulated as a generalized information systems design mathematical program

Geographic Information Systems, Land Information systems and Disaster Information Systems are also some of the emerging information systems but they can be broadly considered as Spatial Information Systems. System development is done in stages which include:

- Problem recognition and specification
- Information gathering

- Requirements specification for the new system
- System design
- System construction
- System implementation
- Review and maintenance

Information systems research

Information systems research is generally interdisciplinary concerned with the study of the effects of information systems on the behavior of individuals, groups, and organizations. Hevner et al. (2004) categorized research in IS into two scientific paradigms including *behavioral science* which is to develop and verify theories that explain or predict human or organizational behavior and *design science* which extends the boundaries of human and organizational capabilities by creating new and innovative artifacts.

Salvatore March and Gerald Smith proposed a framework for researching different aspects of Information Technology including outputs of the research (research outputs) and activities to carry out this research (research activities). They identified research outputs as follows:

1. *Constructs* which are concepts that form the vocabulary of a domain. They constitute a conceptualization used to describe problems within the domain and to specify their solutions.
2. A *model* which is a set of propositions or statements expressing relationships among constructs.
3. A *method* which is a set of steps (an algorithm or guideline) used to perform a task. Methods are based on a set of underlying constructs and a representation (model) of the solution space.
4. An *instantiation* is the realization of an artifact in its environment.

Also research activities including:

1. *Build* an artifact to perform a specific task.
2. *Evaluate* the artifact to determine if any progress has been achieved.
3. Given an artifact whose performance has been evaluated, it is important to determine why and how the artifact worked or did not work within its environment. Therefore *theorize* and *justify* theories about IT artifacts.

Although Information Systems as a discipline has been evolving for over 30 years now, the core focus or identity of IS research is still subject to debate among scholars such as. There are two main views around this debate: a narrow view focusing on the IT artifact as the core subject matter of IS research, and a broad view that focuses on the interplay between social and technical aspects of IT that is embedded into a dynamic evolving context. A third view provided by calling IS scholars to take a balanced attention for both the IT artifact and its context.

Since information systems is an applied field, industry practitioners expect information systems research to generate findings that are immediately applicable in practice. However, that is not always the case. Often information systems researchers explore behavioral issues in much more depth than practitioners would expect them to do. This may render information systems research results difficult to understand, and has led to criticism.

To study an information system itself, rather than its effects, information systems models are used, such as EATPUT.

The international body of Information Systems researchers, the Association for Information Systems (AIS), and its Senior Scholars Forum Subcommittee on Journals (23 April 2007), proposed a 'basket' of journals that the AIS deems as 'excellent', and nominated: Management Information Systems Quarterly (MISQ), Information Systems Research (ISR), Journal of Association of Information Systems (JAIS), Journal of Management Information Systems (JMIS), European Journal of Information Systems (EJIS), and Information Systems Journal (ISJ).

Information Engineering

Information engineering (IE) or *information engineering methodology* (IEM) in software engineering is an approach to designing and developing information systems.

Overview

Information engineering methodology is an architectural approach to planning, analyzing, designing, and implementing applications within an enterprise. It aims to enable an enterprise to improve the management of its resources, including capital, people and information systems, to support the achievement of its business vision. It is defined as: *"An integrated and evolutionary set of tasks and techniques that enhance business communication throughout an enterprise enabling it to develop people, procedures and systems to achieve its vision"*.

Information engineering has many purposes, including organization planning, business re-engineering, application development, information systems planning and systems re-engineering.

History

Information engineering has a somewhat chequered history that follows two very distinct threads. It is said to have originated in Australia between 1976 and 1980, and appears first in the literature in 1981 in the Savant Institute publication 'Information engineering' by James Martin and Clive Finkelstein.

Information engineering first provided data analysis and database design techniques that could be used by database administrators (DBAs) and by systems analysts to develop database designs and systems based upon an understanding of the operational processing needs of organizations for the 1980s.

The Finkelstein thread evolved after 1980 into the data processing (DP)-driven variant of IE. From 1983 till 1986 IE evolved further into the business-driven variant of IE, which was intended to address a rapidly changing business environment. The then technical director, Charles M. Richter, from 1983 to 1987, played a significant role by revamping the IE methodology as well as designing the IE software product (user-data) which helped automate the IE methodology, opening the way to next generation Information Architecture.

The Martin thread was strategy-driven from the outset and from 1983 was focused on the possibility of automating the development process through the provision of techniques for business description that could be used to populate a data dictionary or encyclopedia that could in turn be used as source material for code generation. The Martin methodology provided a foundation for the CASE (computer-aided software engineering) tool industry. Martin himself had significant stakes in at least four CASE tool vendors - InTech (Excelerator), Higher Order Software, KnowledgeWare, originally Database Design Inc, (Information Engineering Workbench) and James Martin Associates, originally DMW and now Headstrong (the original designers of the Texas Instruments' Information Engineering Facility and the principal developers of the methodology).

At the end of the 1980s and early 1990s the Martin thread incorporated rapid application development (RAD) and business process reengineering (BPR) and soon after also entered the object oriented field.

Information engineering topics

IE variants

There are two variants of information engineering. These are called the DP-driven variant and the business-driven variant.

- DP-driven : The DP-driven variant of Information engineering was designed to enable IS Departments to develop information systems that satisfied the information needs of the 1980s - which was largely a DP-driven development environment. Most of the CASE tools available today support this DP-driven variant of IE.
- Business-driven: IE was extended into strategic business planning and developed the business-driven variant of information engineering. This variant was designed for rapid change in the client/server, object-oriented environment of the business-driven 1990s.

Business-driven IE is documented in the later books by Clive Finkelstein.

- Information Strategy Planning : The fundamental objective of Information Strategy Planning (ISP) is to develop a plan for implementing business systems to support business needs.
- Outline Business Area Analysis : The OBAA answers a range of questions related to implementation of a business area. Select tasks to include in a particular project that provide support for business decisions and objectives. Specific information needs and priorities for the business area are needed.
- Detailed Business Area Analysis : The purpose of a DBAA project is to provide detailed models as a solid basis for system design. The methodology helps find the right answers to the right questions. Applying the methodology is never an end in itself.
- Business System Design : The purpose of a Business System Design project is to specify all aspects of a system that are relevant to its users, in preparation for the technical design, construction, and installation of one or more closely related databases and systems. The key tasks are therefore structured to produce unambiguous consistent specifications, with the volume of detail necessary to make planning and technical design decisions.
- Technical Design : A Technical Design project prepares an implementation area for construction and installation. The key tasks are structured to produce a system and database that meet the user's acceptance criteria and are technically sound.
- Construction : The objective of the Construction stage is to produce a system, as defined in the technical specification, on time and within budget. The system should be of an acceptable quality, and contain all necessary operating and user procedures. The task is complete when the acceptance criteria for the business system are met.
- Transition : Transition is defined as the period during which newly developed procedures gradually replace or are interfaced with existing procedures. The execution of a Transition project obviously demands a thorough understanding of both the system to be installed and the systems to be replaced.

IE techniques

Some techniques that are used during an IE project are:

- Entity analysis : identifies all the things that the enterprise may want to hold data about. The analysis classifies all of the things into different entity types, revealing how they relate to each other. Which is being described in the entity model.
- Function analysis and process dependency : takes a function (a major business activity) of the enterprise and breaks it down into elementary business processes. From this, two diagrams are prepared: the process decomposition diagram, which shows the breakdown of a business function, and the process dependency diagram, which shows the interdependencies of business processes.
- Process logic analysis : describes the sequences of actions carried out by a business process and shows which data are used by each action.
- Entity type lifecycle analysis : describes the significant business changes to entities and confirm that processes have been modelled to effect these changes

- Matrix cross-checking : creates cross-references between data objects and processes to verify that they are necessary and complete.
- Normalization : provides a formal means of confirming the correctness of the entity model.
- Cluster analysis : helps define the scope of design areas for proposed business systems.
- Data flow and data analysis : makes a comparison possible between the business area models and the systems currently supporting this area, these current systems are analyzed using data flow and data analysis techniques.

Software tools

There are several tools supporting Information engineering

- Information engineering Facility (IEF) from Texas Instruments Software. This was subsequently sold to Sterling Software and then to Computer Associates. It still exists, in an evolved form within the Advantage suite. As of 2006 referred to as ALL:Fusion Gen, capable of generating J2EE and JAVA web applications in addition to legacy client/server and mainframe platforms.
- Information engineering Workbench (IEW) : Later renamed to Application Development Workbench (ADW) from KnowledgeWare. Knowledgeware was also acquired by Sterling Software. The product no longer exists.
- The business-driven variant of IE is supported by Visible Advantage, an Integrated CASE (I-CASE) tool and by Visible Advisor, a hypermedia Methodology product.
- Metastorm's ProVision product provides support for many types of modeling techniques using a repository based tool.
- Visio provides diagramming support to some of the Martin techniques.

Others included Bachman's Data Analyst, Excelerator and others.

Chapter-2

Geographic Information System

A **geographic information system (GIS)**, **geographical information system**, or **geospatial information system** is the system that captures, stores, analyzes, manages, and presents data with reference to geographic location data. In the simplest terms, GIS is the merging of cartography, statistical analysis, and database technology. GIS may be used in archaeology, geography, cartography, remote sensing, land surveying, public utility management, natural resource management, precision agriculture, photogrammetry, urban planning, emergency management, landscape architecture, navigation, aerial video, and localized search engines.

As GIS can be thought of as a system, it digitally creates and "manipulates" spatial areas that may be jurisdictional, purpose or application oriented for which a specific GIS is developed. Hence, a GIS developed for an application, jurisdiction, enterprise, or purpose may not be necessarily interoperable or compatible with a GIS that has been developed for some other application, jurisdiction, enterprise, or purpose. What goes beyond a GIS is a spatial data infrastructure (SDI), a concept that has no such restrictive boundaries.

Therefore, in a general sense, the term describes any information system that integrates, stores, edits, analyzes, shares, and displays geographic information for informing decision making. GIS applications are tools that allow users to create interactive queries (user-created searches), analyze spatial information, edit data, maps, and present the results of all these operations. Geographic information science is the science underlying the geographic concepts, applications and systems.

Applications

GIS technology can be used for: earth surface based scientific investigations; resource management, reference, and projections of a geospatial nature—both manmade and natural; asset management and location planning; archaeology; environmental impact study; infrastructure assessment and development; urban planning; cartography, for a thematic and/or time based purpose; criminology; geospatial intelligence; GIS data development geographic history; marketing; logistics; population and demographic studies; prospectivity mapping; location attributes applied statistical analysis; warfare

assessments; and other purposes. Examples of use are: GIS may allow emergency planners to easily calculate emergency response times and the movement of response resources (for logistics) in the case of a natural disaster; GIS might be used to find wetlands that need protection strategies regarding pollution; or GIS can be used by a company to site a new business location to take advantage of GIS data identified trends to respond to a previously under-served market. Most city and transportation systems planning offices have GIS sections.

History of development

In 1854, John Snow depicted a cholera outbreak in London using points to represent the locations of some individual cases, possibly the earliest use of the geographic method. His study of the distribution of cholera led to the source of the disease, a contaminated water pump (the Broad Street Pump, whose handle he had disconnected, thus terminating the outbreak) within the heart of the cholera outbreak.

E. W. Gilbert's version (1958) of John Snow's 1855 map of the Soho cholera outbreak showing the clusters of cholera cases in the London epidemic of 1854