

INDEX

RESEARCH PAPERS:

- 1) Integrating business process with technology based process, a case study from online railway reservation system-Information Integrity based Approach.
Prof. Pritee Hrishikesh Joshi,Muktangan English School & Jr. College,Pune

ARTICLES

- 1) Cloud Computing The Dawn of New Age In Computing
Porf. Gautam Bapat,MITSOM College,Pune
- 2) Soft Skills-Who needs them?
Porf. Ina Goplani,MITSOM College,Pune
- 3) Data Mining The Analysis Step of the knowledge discovery in databases process
Porf. Asha Nagoriya,MITSOM College,Pune

Research Paper on

**Integrating business process with technology based
process, a case study from online railway reservation system-
Information Integrity based Approach**

Submitted by

**Mrs. Pritee Hrishikesh Joshi
Muktangan English School & Jr. College,Pune**

Introduction

A business process or business method is a collection of related, structured activities or tasks that produce a specific service or product (serve a particular goal) for a particular customer or customers. A business process begins with a customer's need and ends with a customer's need fulfillment

Business Processes are designed to add value for the customer and should not include unnecessary activities. The outcome of a well designed business process is increased effectiveness (value for the customer) and increased efficiency (less costs for the company).

These processes are affected by changing requirements, new technologies, new market trends, and changing customer demands. The major reason of business failures is due to the **informational error**. Therefore it is necessary to identify all ambiguities and contradictions in the requirements and to come up with new business process.

The information used, generated and originated is having noise because that information is not having integrity. **Information Integrity Technology** is an upcoming technology of information science that focuses totally on trustworthiness of information and is one of several research areas Information Integrity is dependability and trustworthiness of information and is a key factor determining strategic business advantage. Its determinants are accuracy, consistency and reliability of information. The Information Integrity is defined as the flow of information from one process to another process should be correct, reliable, consistent known as Information Integrity. In each and every system there is a flow of information but whether it is accurate, whether all the environmental factors are considered for the case, all is to be checked in the scene that if any factor to be considered then whether it affects the system or no. If it affects then to make it more effective and efficient, we must remove the information gap or finding the error's within a process so to remove those gap's providing a new technology to overcome the gap's we prove a solution known as **Information Integrity Technology**.

I*I Technology minimizes information failures at minimum cost associated with business process and failures formed there. This study is to develop business processes or

rules using I * I to minimize product errors. It is common to categorize a system as *one that is system in being* (i.e., in operation) and *one that is to be* (i.e., design). System *in being* is the case of an existing system - system in operation, i.e., system for which external and internal customer requirements are validated and verified and no error is anticipated as system environment is taken to be fixed or static.

This technology will help the system should run in a effective and efficient way even if the environmental factors goes on changing. Many times it happens that the system is running smoothly but as time goes on, the environmental factors or the requirements goes on changing means the requirements are instant and local. To fulfill those requirements we have to make changes within a system so as to work efficiently and effectively. Instead of value adding one should do value creating activity which will gives you a chance/opportunity to create new process so as to work/process should be more effective and efficient. This new process is designed by integrating current business process with technology based process based on Information Integrity approach with the help of **Environment /Information Topology(ET/IT)** .Here in this research I am going to discuss case for **Online Railway Reservation System**. And also going to study how it can be integrated with new technology based processes to meet changing requirements.

Statement of problem

Online railway reservation has made booking railway tickets process very easy. A user can sit at any place and book his ticket through internet. He has to login on a website www.irctc.co.in . Once he logs in he has to fill required data and book tickets with the help of credit card. And the ticket is issued to the customer just on some clicks.

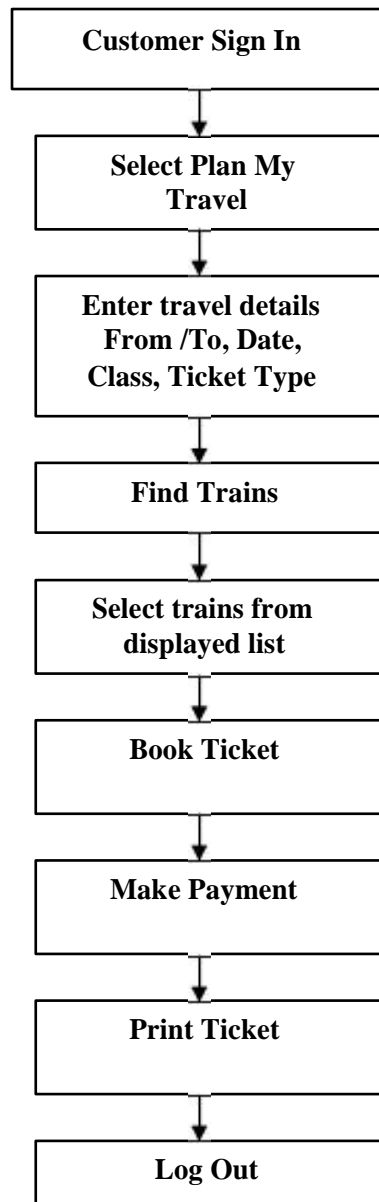


Fig.1.1 Diagrammatic representation of Online Reservation System

Many times it happens that the system is running smoothly but as time goes on, the environmental factors or the requirements goes on changing means the requirements are instant and local.

The problem statement is:” Information Integrity Technology system development for designing and developing integrity information system in the phase where requirements are instant and local.”

To fulfill those requirements we have to make changes within a system so as to work efficiently and effectively. I * I technology will help the system should run in a effective and efficient way even if the environmental factors goes on changing. Instead of value adding one should do value creating activity which will gives you a chance/opportunity to create new process so as to work/process should be more effective and efficient. **Whenever there is system performance failure it is because complete end to end set of interrelated activities that go to create value for customer have not been identified and accounted for design process.**

Thus a case of online railway reservation system is studied by implementing **Environment Topology** for different requirements of stakeholders. My research study also aims to **integrate current business process with technology based process** to solve the problems caused in current online reservation system. And to show how change in the business process will increase the integrity of the business and make it more beneficial.

Purpose of study

The purpose of research work is as follows:

1. The research work intends to study flaws in current online system due to static environment.
2. To identify different levels of users and their requirements.
3. To develop environmental topology for online booking system by identifying different stakeholders..
4. To come up with new business opportunity by integrating a business process with technology based process.
5. To define new processes to minimize the system failures.
6. Finally the research work intends to promote Information Integrity (I*I) Technology and to get the competitive advantage by reducing the risks and increasing the accuracy by developing new rules and processes

Research Methodology

Research is conducted for the purpose of discovering, interpreting, enhancing, developing standards to systematize measurements, and furthering advancement of knowledge. **Methodology** is the rationale behind collection of concepts, ideas, theories, and assumptions. Combined, a good research methodology is a technique of collecting data systematically.

A subsidiary of the Indian Railways, **Indian Railway Catering and Tourism Corporation (IRCTC)** handles the catering, tourism and online ticketing operations of the railways. IRCTC is famous for changing the face of railway ticketing in India. In 2003, with about six lakh registered users, the IRCTC website had become one of the largest and the fastest growing e-commerce websites in the Asia-Pacific region within a short span of its going online. **New IRCTC website** has drop down navigation, which provides proper information. User guides, Enquiries, General information, Agents, Special offers, IRCTC Zone are the navigation of IRCTC website. This navigation provides information about ticket booking, ticket canceling, browser setting, reservation inquiry, free tickets/gift, etc. The new interface is intended to increase the number of visitors who book rail tickets using the site. And hence the site is studied to find the problems encountered because of changing requirements and to increase its working efficiency. The information is gathered from different users. This **qualitative research** study is carried out using **a case study as a research strategy**. The study shows use of data generation methods like **interview, observation method, Internet Search, Literature Review, Reference Books, Research papers**

Case Study:

A cases study focuses on one instance of the 'thing' that is to be investigated; an organization, a department, an information system, a discussion forum, a system developer, a development project, a decision and so on. This one instance, or a case, is studied in depth, using a variety of data generation methods like observation, interview and documents.

There is no fixed criteria and way of conducting case studies, and it is done with a combination of other methods of qualitative research. This method studies in systematic way events, collection of data, analyses of information, and reporting the results. The end result is generally a clear perception and understanding of why and how things happen in the manner they do.

Data generation methods used in research study are as follows:

1 Participant Observation:

This method requires the researcher to become an active participant and to record his own experiences (subjective experiences), and make them available to a wide audience, for future studies and references.

Thus by participating in online booking activity some experiences were noted down for further study.

2 Direct Observations:

In direct observation method, there is no participatory involvement on the observer's part. One can record the happenings or observe from behind one-way mirrors. Thus booking processes were observed and noted down for further study of the research.

3 Interviews:

Interviewing involves direct interaction between the researcher and a respondent or group..A group of people including senior citizen, students, housewives, professionals were interviewed to collect data. This data is used to identify their different, new requirements.

4 Internet Search:

Mainly www.irctc.co.in website was studied to study current process and issues related with it. Also www.ciir.org.in site was studied to study in deep concept of Information Integrity and I*I technology.

5 Research Papers:

Different research papers related to I*I technology were studied from CIIR site.

Information Integrity Means

Information Integrity (I*I) is dependability and trustworthiness of information and controlling it is a key factor for determining strategic business advantage. Its attributes are accuracy, consistency and reliability of Information System (IS) and information there from [Rajaraman and Mandake, 1995].

Accuracy:

Accuracy (A) is defined as the degree of agreement between a particular value and an identified source. It can be assessed by identifying the relevant established source (standard) and by determining an acceptable tolerance. Specifically, the identified source provides the correct value – preferably the value corresponding to the optimum integrity.

Consistency:

Consistency (C) is defined as the degree to which multiple instances of a value satisfy a set of constraints. The multiple instances may exist across space (such as database and systems) or over time. Consistency is then with respect to a set of constraints and data/information is said to be Consistent with respect to a set of Constraints if it satisfies all Constraints of the data/information model.

Reliability:

Reliability (R) is complex attribute to define. Traditionally, it is a large concern in system development lifecycle model and refers to a wide range of issues relating to design of large systems (complex computerized information system) which are required to work in specified period of time.

From this point of view for an IS the definition of reliability given as “accuracy with which information obtained represents data item in whatever respect the information system processed it” can be seen to define the reliability requirement for the IS as a whole [Vijay Mandke].

Information Integrity is an interdisciplinary concept covering areas of Design Engineering, Systems engineering, Software Engineering, Information Systems, Decision Models, and Management and economic Sciences. Also it covers the area of Information management and economics. I*I is thus concerned with the correctness and exactness aspects of the information.

Application of Environment topology to current system

Environment Topology or Information Topology(ET/IT) is an abstraction from the environment discussed above. It is developed through an iterative and continuous process of information origination, evaluation and processing.

Attributes of Information Topology (ET/IT)

It is enriched with information envelope. Each layer has informational processes and decision mile-posts. This reduces the problem complexity while ensuring effectiveness and economic solutions. That is to say that how much of information should be originated in the topology. To this the answer is that corresponds to the information requirement of the dynamic decision stages.

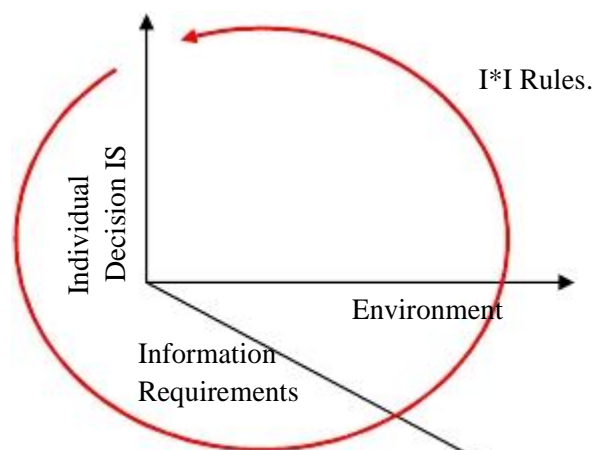


Fig.5.1 Attributes considered for ET/IT

The above diagram throws light on three attributes to build the ET/IT. They are environment, informational requirements and individual decision IS. Each is encapsulated by I*I Rules. The I*I Rules are in respect of following:

- Observation, verification, validation, recognition, forecasting, co-ordination, evaluating alternatives generated for their contributions to operable goals, recourse management.
- Requirements' Processing.

- Standard information (data element) networks.
- Re-evaluation, selecting flexible information decision for control implementation.
- Procedures for information networks for customized requirements.
- Procedures for flexible systems.

The informational requirements identified during the projects are as follows-

- Normal, standard requirement: Requirements are either given by the customer or assumed by the supplier, mostly latter. (Customer can be external as well as internal, internal customer coming into the picture when business process under go changes due shift in data driven technology application across the enterprise-wide supply chain.) Given this it is common for supplier to assume normal, standard requirements for a customer with assumed eligibility criteria. However, there are issues such as what-if eligibility not satisfied? Should that necessarily mean no consideration of that customer? For example, for product sale, supplier may assume that customer eligible who has money with him at that time. What-if the customer is not having money? What additional requirements supplier must anticipate, *ex-ante*, so as to not to loose that customer and not suffer business loss. In other words, it should be useful to define along with normal, standard customer eligibility requirements, requirement with marginal eligibility deficiency. There can be other requirements as well. And there are mentioned below.
- Requirements with deficiencies- major:
- Shifting requirements i.e. shifting in conceptual objects which can be
 - Informational objects, mainly to be found in service sector. Also, budgets at zero based etc.
 - Functional Objects – desired and maintainable.
 - Performance based objects.
- Requirements acceleration.
- Requirements delaying.
- Requirements coming with delays
- Transfer of requirements, horizontal, mobility of requirements
- Combination of requirements.

- Requirements with vertical mobility.
- Decline in requirements.
- Evolving requirements – changing priorities.
- Conflicting requirements, complex requirements.

Application of Environment topology to current system

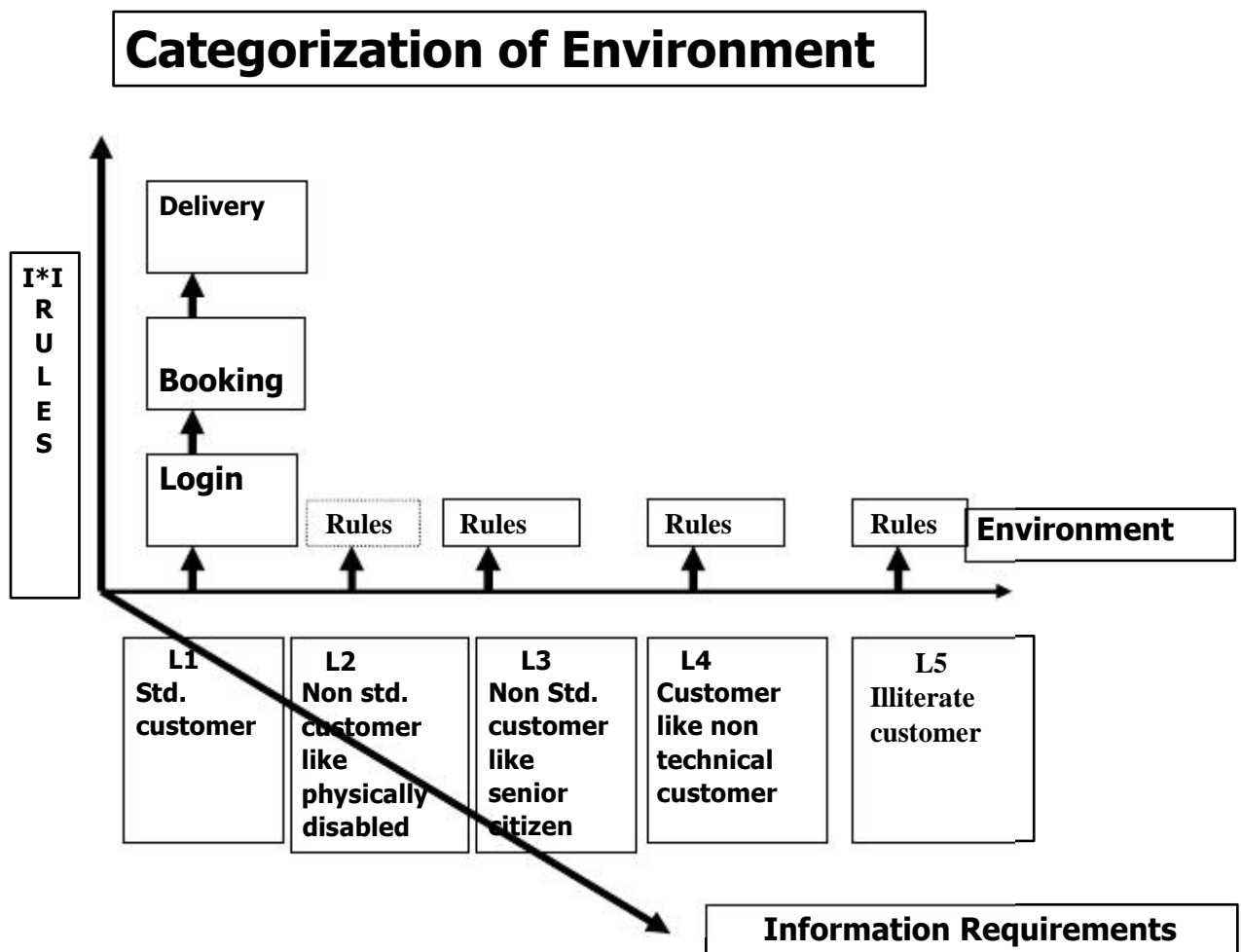


Fig. 5.2 Categorization of Environment

The above diagram shows standard and non standard customers. By considering these non standard customers new I*I rules are designed in such way that these non standard customers can also avail the facility of online reservation.

Applying **Categorization of Environment** to current online reservation system

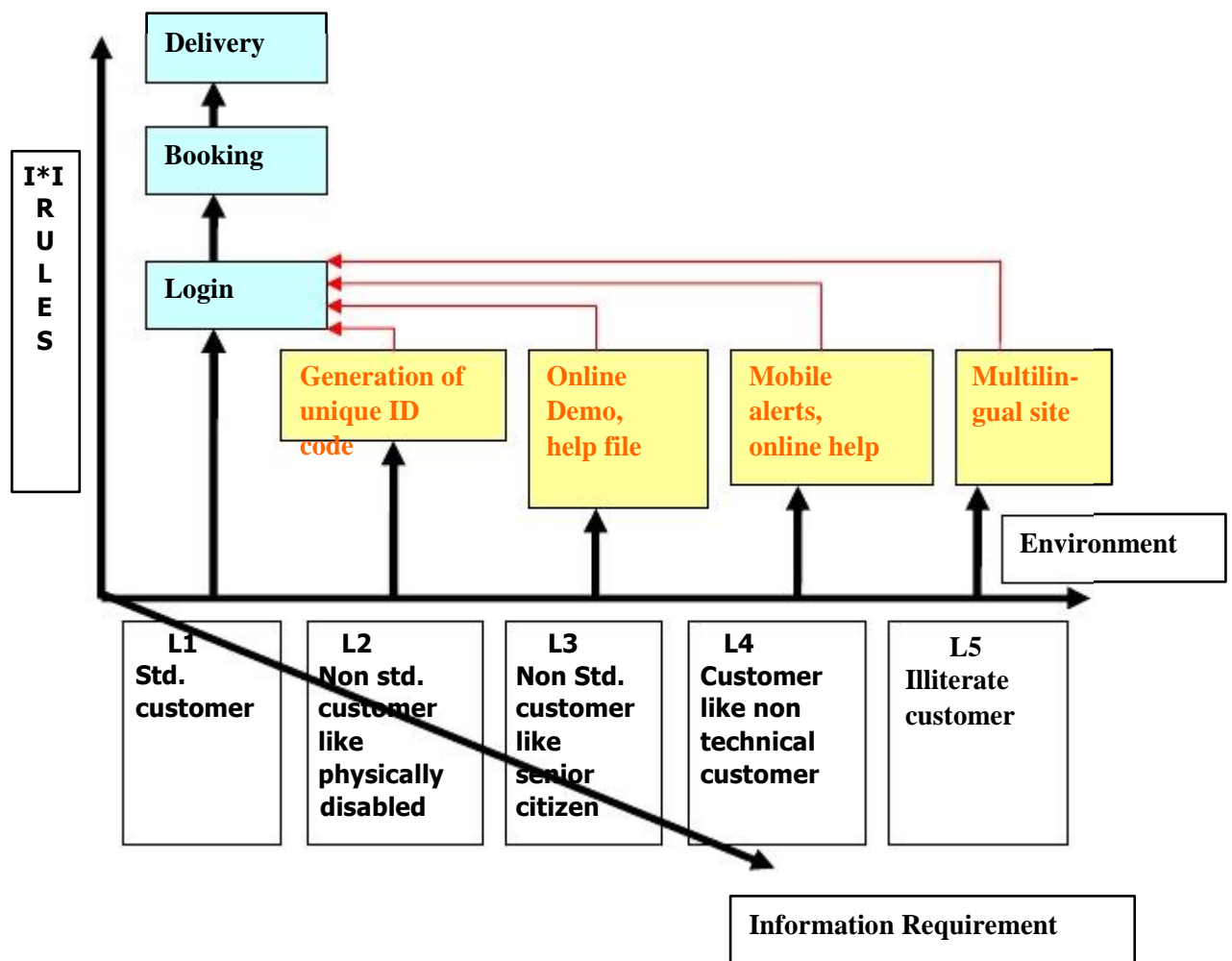


Fig. 5.3 Categorization of Environment applying to current reservation system

Different types of requirements are identified in Environment Topology .The new process should be designed in such way that even if new requirement added in the system, system runs smoothly. In this case even though there is non- standard customer occurs like senior citizen, physically disabled person, housewife, phobia tic person he can avail online

reservation facility. The process should be changed by integrating new business process. Hence I*I rules must be designed for each new requirement.

Level 2: Physically handicapped /disabled stakeholder

Rules:

1. Integrated voice system should be added so that he can avail facility of online booking.
2. Generation of unique code Id .With the help of which he an book ticket online.

Level 3: Senior citizen, housewife stakeholder

Rules:

1. Online demo, help file should made available on the site..
2. Payment can be made directly through savings account.

Level 4: Technology Phobia tic / non technical stakeholder

Rules:

1. Online demo, help file should made available on the site.
2. Site should be made more user friendly
2. These kind of stakeholder should be motivated to go with Mobile reservation system.

Level 5: Illiterate stakeholder

Rules:

- 1 Site should be made more user friendly.
- 2 Site should be made multilingual.
- 3 Mobile computing concepts should be promoted.

Summary and Conclusion

This research work mainly aims at developing I*I technology and Environment /Information Topology (ET/IT) for online reservation system. It begins with the study of current system. It identifies the anomalies in current system due to changing requirements.

This paper develops new business processes and I*I rules based on information integrity approach that will assure increase in the quality, accuracy and acceptability of the system by considering the different environments and different levels of requirements.

It suggests to develop I *I technology model for the current online reservation system. The diagram below shows an evolutionary scenario from a traditional IS to I*I System.

To achieve Information Integrity, systems should be designed for changing and complex environment and also for changing customer needs, which are local and global. The system aims at capturing, storing and processing different customer requirement with the help of the **Environmental Topology**.

Bibliography

Research Papers:

Mandke V. V.(2007) On Integrating Business Process with Technology Development – A Design Basis for Information Integrity Technology System, CIIR website.

Vijay V. Mandke CIIR Paper for Limited Circulation, New Delhi (June 2003) Making a Case for I*I in healthcare, CIIR website.

Vijay V. Mandke and Madhavan K. Nayar.Proceedings of information Technology for Business Management, 16th World computer Congress 2000(August 21-25), Beijing, China Information Integrity Imperative for Competitive Advantage in Business Environment characterized by Uncertainty, CIIR website .

Vijay V. Mandke (2006-07) Information Integrity System available from <http://www.ciir.org.in/COURSE MATERIAL/Lectures>.

Sakal Article dated on 25 September, 2009 “Innovation on Track “July 15, 2 0 0 6

Internet sites:

1 www.iretc.co.in

2 www.ciir.org.in

3 www.google.co.in

4 www.altavista.com

5 <http://www.slideshare.net/mssridhar/introduction-to-research-methodology-presentation>

6 <http://books.google.co.in/books>

7 <http://www.buzzle.com/articles/methods-of-qualitative-research.html>

8 <http://www.smartdraw.com/specials/flowchart.asp>

9 www.infogix.com

Reference Books:

Oates Broiny, 2006, *Research Information Systems and Computing*, New Delhi, SAGE

Cloud Computing

The Dawn of New Age In Computing

An introduction to cloud computing

By Prof. Gautam Bapat

Head- Department of Research & Innovation

This article is written for informative purpose. There are no new ideas or concepts introduced. It is written for the people who do not know cloud computing at all, but of course with little technical knowhow. This piece of text is just for the sake of information.

Cloud computing, in all of its forms, is transforming the computing landscape. It will change the way we deploy technology and how we think about the economics of computing. Cloud computing is more than a service sitting in some remote data centre. It's a set of approaches that can help organizations quickly, effectively add and subtract resources in almost real time. Unlike other approaches, the cloud is as much about the business model as it is about technology. Companies clearly understand that technology is at the heart of how they operate their businesses. Business executives have long been frustrated with the complexities of getting their computing needs met quickly and cost effectively. In a sense, cloud computing has started to become mainstream because these business executives have forced the issue into the forefront.

Cloud computing can be defined as the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices

as a utility (like the electricity grid) over a network most of the times the internet).

In a dynamic economic environment, a company's survival may depend on your ability to focus on core business and adapt quickly. Yesterday's profitable business model can't be counted on to translate into future growth and profits. As a business adapts to changing government and industry regulations, evaluates new business partnerships, and anticipates competitive threats, IT needs to help the business find new ways to respond.

Many players make up the world of cloud computing:

- ✓ The **vendors** providing applications and enabling technology, infrastructure, hardware, and integration
- ✓ The **partners** of these vendors that are creating cloud services offerings and providing support services to customers
- ✓ The **business leaders** themselves who are either using or evaluating various types of cloud computing offerings

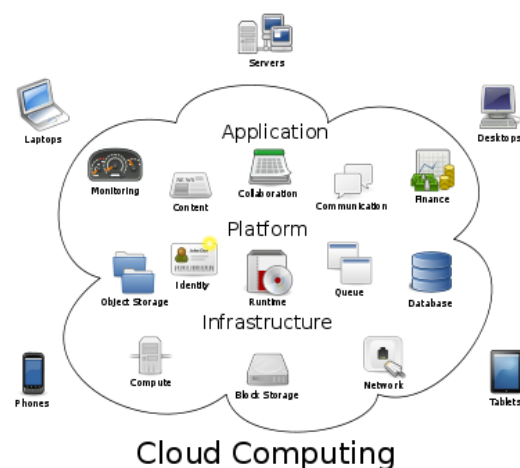


Image courtesy : wikipedia.org

Cloud computing is offered in different forms:

- ✓ **Public clouds**

✓ **Private clouds**

✓ **Hybrid clouds**, which combine both public and private

The world of the cloud has lots of participants:

✓ The **end user** doesn't really have to know anything about the underlying technology. In small businesses, for example, the cloud provider becomes the de facto data centre. In larger organizations, the IT organization oversees the inner workings of both internal resources and external cloud resources.

✓ **Business management** needs to take responsibility for overall governance of data or services living in a cloud. Cloud service providers must provide a predictable and guaranteed service level and security to all their constituents.

✓ The **cloud service provider** is responsible for IT assets and maintenance.

The cloud embodies the following four basic characteristics:

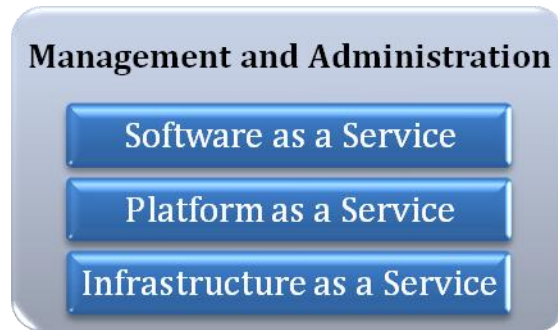
✓ Elasticity and the ability to scale up and down

✓ Self-service provisioning and automatic deprovisioning

✓ Application programming interfaces (APIs)

✓ Billing and metering of service usage in a pay-as-you-go model

The model of Cloud service delivery :



Let us know see, why businesses prefer Cloud. In other words why cloud is so economic :

✓ The standardized workloads can be executed on a highly integrated, massively replicable infrastructure stack. They don't have to support a wide array of workloads and a heterogeneous stack of hardware, middleware, OS, and so on.

✓ The computer hardware and network is highly streamlined and can be bought in bulk and configured to allow expansion. Often these companies require that hardware be engineered for their unique scaling requirements.

✓ All software can be stripped down so that only what is necessary is loaded.

✓ The service/software itself is written from scratch in a cloud-optimized way, tailored for efficiency at an instruction level.

✓ The provider may not offer or guarantee a specific service level.

✓ There is no need for virtualization technology to build virtual machines. The software can be engineered to the bare metal.

✓ The profile of the workload is measurable and predictable simply by numbers of users.

Many levels of security are required within a cloud environment:

✓ **Identity management:** For example, so that any application service or even hardware component can be authorized on a personal or group role basis.

✓ **Access control:** There also needs to be the right level of access control within the cloud environment to protect the security of resources.

✓ **Authorization and authentication:** There must be a mechanism so the right people can change applications and data.

You have to think about several issues before sending your organization into the cloud. There isn't just one approach. You might choose one or more of these approaches at different times for different reasons.

Consider a few simple examples:

✓ Your company is building a new application that will change the way you sell products online. You want to stress test this new application before releasing it to customers. Although you have a few extra resources inside your firewall, they aren't extensive enough to demonstrate if the new application will really scale. Using a cloud Infrastructure as a Service enables you to test the application effectively.

✓ Your company has run its own email internally for more than 20 years. It takes up a lot of space in the data center and requires a staff of ten people. Money is tight and the CIO must cut staff and capital expenses. The CIO finds a Software as a Service platform that can run the corporate email for a fraction of the cost of running email internally. Your company makes the move and the savings are dramatic.

✓ Your company is building a new but highly experimental application that might transform its business model. It might not be worth spending a lot of money on software and hardware upfront. In fact, if the project succeeds, the new application may be deployed in the cloud (and not within your company's own data center). Therefore, the company uses a Platform as a Service (PaaS) that includes its own well-designed and fully vetted development environment, new generation tools, and interfaces that allow it to connect to many different environments. No need to pretest all the components provided by the PaaS vendor—they're well designed and have been tested. The new application built on this platform is completed in record time and deployed to a test group of customers directly from the cloud service.

✓ Your company has started using a third-party SaaS solution for its customer-management application. It has successfully replaced the on premise customer-relationship management package that you've been running in the data centre for years. Now your company wonders what else could be moved out of the data centre into the cloud. How about the mainframe transaction processing system that handles all orders worth more than \$1 million? After some investigation, you realize that because the system is only used by a few individuals in the company and the information needs to be carefully governed, the cloud isn't a good choice.

✓ Your CIO has seen some new software that could solve a serious problem, but you aren't convinced that the solution is

right. Instead of buying a license, your company decides to use it as a service. After six months, it proves valuable. The software company offers you the opportunity to use the Software as a Service or on premise.

Deciding to investigate cloud computing is primarily a business decision. Therefore, start by evaluating these things:

- ✓ Your company's business strategy
- ✓ The role that your technology infrastructure plays in that strategy

For Developing a Road Map for cloud establishment ask these questions:

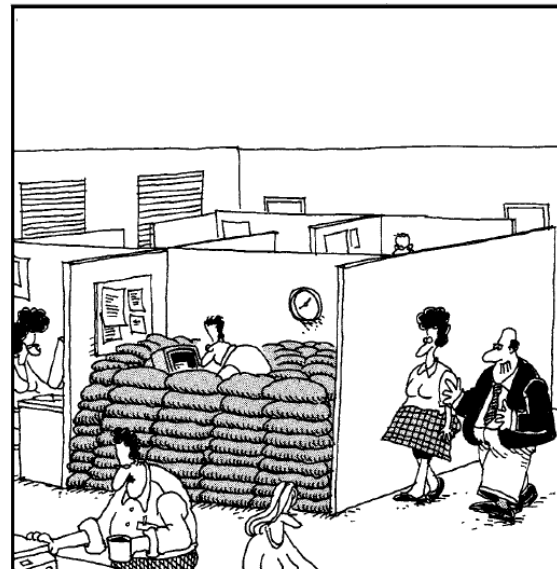
- ✓ What type of business am I in?
- ✓ Is my industry highly regulated?
- ✓ How do I benefit my customers?
- ✓ How much value am I getting from my existing data centre operations?
- ✓ What are my company's short-term and long-term goals?
- ✓ Am I under pressure to reduce capital expenses?
- ✓ Am I planning to develop new, untested technology-based offerings over the next 18–24 months?
- ✓ Am I looking to acquire complementary companies?
- ✓ How are my competitors dealing with their technology strategy?
- ✓ Are my competitors able to move more quickly than my company?
- ✓ What are my strategic advantages?

After you understand the company strategy, you can develop your own strategic plan. You must consider many things before developing a road map:

- ✓ The efficiency and effectiveness of your current data center
- ✓ Costs
- ✓ Risks
- ✓ Your organizational readiness

After you understand the issues and gaps, you can start designing your cloud plan — the road map that outlines the following:

- ✓ What are the services that you need to support your business growth?
- ✓ How you will roll them out?
- ✓ When you will roll them out (or in, as it were)?



"I assume everyone on your team is on board with the proposed changes to the system architecture."

Image courtesy : Rich Tennant

So, cloud computing is not only one more just new technology but a completely new vision to look at business.

- Prof. Gautam Bapat
Head- Department of Research & Innovation
MITSOM College, Pune.
gautambapat@yahoo.co.in

Soft Skills

Who needs them?

The concept of soft skills as a requirement for success in business is not a new one. Neither is it new to the IT and software sectors. The term 'soft skills' has been in common use in the business world since 1990. It is generally accepted that the term evolved from the word 'software' used to differentiate computer programs from the technical specifications in computer hardware. Yet, the word 'soft' associated with 'skill' has led to the misconception even today that these skills are easy to acquire.

Among the soft skills that were perceived to be required by HR heads in leading companies in the IT sector in 2002 were cross-cultural sensitivity, communication, customer interface skills such as listening, influencing and persuasive communication; presentation skills, meeting skills, and so on.

A huge competence gap was also being felt at the time at the project manager/leader level since people were good at doing, but not at getting things done. This brought into focus the need for leadership and team working and the need for soft skills training to develop employees as "people" in order to retain them into sharp focus. Why then is the lack of soft skills still very much a matter of concern in India? Why are soft skills so hard to find?

The answer lies broadly in these areas:

It is widely recognized that there is a great shortage of employable people in India at a time when we are seeing high growth rates in virtually every sector of the economy. There has been a proliferation of educational institutions to meet the demand for technical skills – however, many offer very poor education due to the lack of qualified faculty and suitable facilities. Neither is there room in the curriculum for training in soft skills, nor are there enough trainers available. Companies, particularly in the IT and software sectors have to devote enormous resources to training/retraining people in order to make them productive.

Unlike in the West, college graduates have virtually no exposure to the people skills required to be effective at work since they have little opportunity to work part time when they are students. Employees also fail to recognize that they need soft skills to work with others. The typical Indian professional believes that technical skills are the most valued skills in the workplace and that their performance evaluation is based primarily on technical output. This belief is reinforced by the fact that companies continue to hire on the basis of technical skills to meet their immediate needs.

Employees are often willing to devote their own time and money to upgrade their technical skills but are less inclined to spend their time and money on improving their soft skills

There is still considerable confusion about what soft skills are. For example, in our experience at Soft Skills India, many CEOs and HR managers still think of English language skills when they say that the communication skills of their employees need improvement. While some companies recognize that soft skills are acquired only after a lot of practice, and are not

just concepts that can be learned within a short period of time, corporate India still tends to expect quick fixes through short training programs.

What are soft skills and what soft skills are required by IT and software professionals?

According to IEEE, one of the world's largest professional associations, "Soft skills are the techniques you need to work with others. The things you weren't directly taught in engineering school. Skills like working on a team, networking with other engineers, public speaking, successfully and gracefully convincing others of your viewpoint all are considered to be soft skills." Google, Inc. sees soft skills as "Personal management skills such as attitudes and behaviors that drive one's potential for growth, and team work skills. "

"To a techie, anything outside of product and development skills could be considered soft skills", says Paula Moreira, writing on "Soft Skills for IT professionals". Among her unofficial list of skills required to get to the top, she lists teamwork, communication skills ranging from answering the phone to writing e-mails and proposals; presentation skills, selling skills, running meetings, leadership, problem solving and customer service.

A survey done by an independent research group found that the most important soft skill for IT staffers was perceived to be interpersonal skills by 37% of respondents, written or verbal communication (20%), ability to work under pressure (17%), overall business acumen (11%), professional demeanor (7%) and other skills by 8% of the respondents. A study on Information Technology Workforce Skills found that there was general agreement on 8 attributes and skills for employability. These are:

- ✓ Communication (oral and written) and ability to converse courteously

- ✓ Ability to work productively in teams and groups
- ✓ Customer service focus
- ✓ Ability to listen and comprehend
- ✓ Ability to be resourceful and constructive when solving problems
- ✓ Ability to analyze, prioritize work, evaluate and work with minimal supervision
- ✓ Ability to comprehend and express concepts in quantitative terms
- ✓ Ability to develop original solutions to novel problems quickly

Soft skills is a term used related to a person's EQ (emotional intelligence quotient) the cluster of personality traits, social graces, communication, language, personal habits, friendliness, and optimism that characterize relationships with other people. These skills are considered important by many recruiters as they believe that soft skills complement hard skills (part of a person's IQ), which are the occupational requirements of a job and many other activities.

Speaking on the importance of soft skills, a HR professional says a candidate with proper soft skills has an edge over others in this competitive world. "It is essential to gauge a person's soft skills before recruiting because behavior is often associated with one's thought process," he adds. Communication is a holistic subject and soft skills are an important subset. Without soft skills, there could be misunderstanding, particularly when people are from different backgrounds.

One of the most important aspects of soft skills in professional world is relationship management: the ability to inspire, influence, and develop others while managing conflict.

Some soft skills for professionals are

1. Interpersonal skills
2. Team spirit
3. Leadership skills
4. Social grace
5. Business etiquette
6. Negotiation skills
7. Behavioral traits such as attitude, motivation and time management
8. Customer service skills
9. Communication skills

Tips to enhance your soft skills are:

1) Awareness: You need to begin by making a list of skills that you lack by discussing with your close friends, co-workers etc. Awareness is the first step in building up your soft skills forte.

2) Interact: To build an understanding with your teammates, you should interact more with them. Also take up projects and get involved in some voluntary work.

3) Be optimistic: Exude that positivism and spread the infectious optimism.

4) Self-motivating: Set personal goals to reach new highs and keep targets for yourself.

5) Possess a forgiving and benevolent ego: Humble oneself, accept and learn sincerely from criticism.

6) Be a team player: Learn to trust, build comradeship, set common goals, provide support, friendship & get your hands dirty!

7) Communicate effectively : Be clear, tactful, and diplomatic and pace the conversation.

8) Think out of box: Be creative and hone your innovative and creative skills

9) Look at the Big Picture: Consider all aspects, potential opportunities, threats & contingencies

10) Learn how to multitask and prioritize : Practice empowerment and delegation, manage time well when multitasking and prioritizing.

Prof Ina Goplani

Assistant Professor
MITSOM College, Pune

1

Data Mining

(The Analysis Step of the knowledge discovery in databases process)

Data mining involves the use of

sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods consequently, data mining consists of more than collecting and managing data, it also includes analysis and prediction. Data mining can be performed on data represented in quantitative, textual, or multimedia forms. Data mining applications can use a variety of parameters to examine the data. They include association sequence or path analysis, classification, clustering, and forecasting.

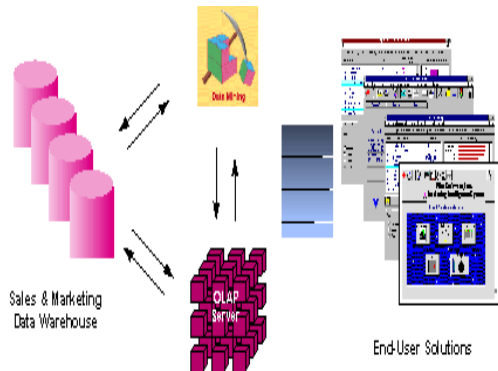
Data mining utilizes a discovery approach, in which algorithms can be used to examine several multidimensional data relationships simultaneously, identifying those that are unique or frequently represented. Two precursors are important for a successful data mining exercise; a clear formulation of the problem to be solved, and access to the relevant data.

Data mining involves six common classes of tasks:

1. **Anomaly detection** (Outlier/change/deviation detection) – The identification of unusual data records, that might be interesting or data errors and require further investigation.
2. **Association rule learning** (Dependency modeling) – Searches for relationships between Variables.
 - a. For example a supermarket might gather data on customer purchasing habits. Using association rule learning, the supermarket can determine which products are frequently bought together and use this information for marketing purposes. This is sometimes referred to as market basket analysis.
3. **Clustering** – is the task of discovering groups and structures in the data that are in some way or another "similar", without using known structures in the data.
4. **Classification** – is the task of generalizing known structure to apply to new data. For example, an email program might attempt to classify an email as legitimate or spam.
5. **Regression** – Attempts to find a function which models the data with the least error.
6. **Summarization** – providing a more compact representation of the data set, including visualization and report generation.

Architecture for Data Mining

Data mining must be fully integrated with a data warehouse as well as flexible interactive business analysis tools. Many data mining tools currently operate outside of the warehouse, requiring extra steps for extracting, importing, and analyzing the data. Furthermore, when new insights require operational implementation, integration with the



Models for Data mining

In the business environment, complex data mining projects may require the coordinate efforts of various experts, stakeholders, or departments throughout an entire organization. In the data mining literature, various "general frameworks" have been proposed to serve as blueprints for how to organize the process of gathering data, analyzing data, disseminating results, implementing results, and monitoring improvements.

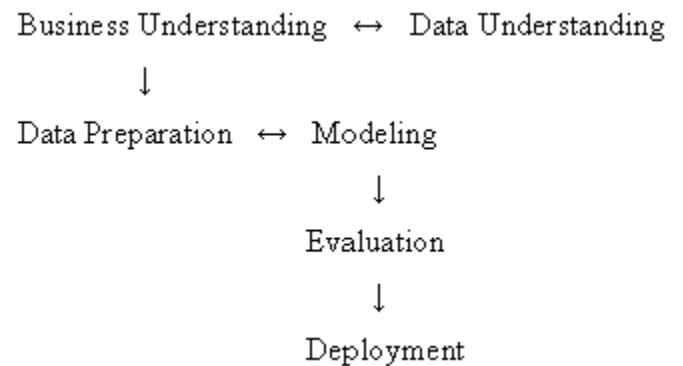
Another approach - the **Six Sigma** methodology - is a well-structured, data-driven methodology for eliminating defects, waste, or quality control problems of all kinds in manufacturing, service delivery, management, and other business activities. This model has recently become very popular (due to its successful implementations)

that grew up from the manufacturing, quality improvement, and process control traditions and is particularly well suited to production environments (including "production of services," i.e., service industries).

warehouse simplifies the application of results from data mining. The resulting analytic data warehouse can be applied to improve business processes throughout the organization, in areas such as promotional campaign management, fraud detection, new product rollout, and so on.

Following figure illustrates architecture for advanced analysis in a large data warehouse.

One such model, **CRISP (Cross-Industry Standard Process for data mining)** was proposed in the mid-1990s by a European consortium of companies to serve as a non-proprietary standard process model for data mining. This general approach postulates the following (perhaps not particularly controversial) general sequence of steps for data mining projects:



in various American industries, and it appears to gain favor worldwide. It postulated a sequence of, so-called, **DMAIC steps** -

Define → Measure → Analyze → Improve → Control

Another framework of this kind (actually somewhat similar to Six Sigma) is the approach proposed by

SAS Institute called SEMMA –

Sample → Explore → Modify → Model → Assess

This is focusing more on the technical activities typically involved in a data mining project.

All of these models are concerned with the process of how to integrate data mining methodology into an organization, how to "convert data into information," how to involve important stake-holders, and how to disseminate the information in a form that can easily be converted by stake-holders into resources for strategic decision making

Types of Data Mining

1. Supervised Classification (target):

- Logistic regression (discrete outcome)
- Multiple regression (continuous outcome)
- Decision trees (discrete outcome)
- Regression trees (continuous outcome)
- Neural Nets (discrete and continuous outcomes)




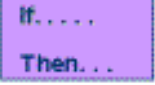
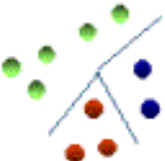
2. Unsupervised Classification (no target)

- Cluster analysis (K-Means, hierarchal, etc.)

- Self-Organized maps (SOMS).

What is data mining good for?

Data mining software allows users to analyze large databases to solve business decision problems. Data mining is, in some ways, an extension of statistics, with a few artificial intelligence and machine learning twists thrown in. Like statistics, data mining is not a business solution, it is just a technology. For example, consider a catalog retailer who needs to decide who should receive information about a new product. The information operated on by the data mining process is contained in a historical database of previous interactions with customers and the features associated with the customers, such as age, zip code, their responses. The data mining software would use this historical information to build a model of customer behavior that could be used to predict which customers would be likely to respond to the new product. By using this information a marketing manager can select only the customers who are most likely to respond. The operational business software can then feed the results of the decision to the appropriate touch point systems (call centers, direct mail, web servers, email systems, etc.) so that the right customers receive the right offers.

	Decision Trees
	Nearest Neighbor Classification
	Neural Networks
	Rule Induction
	K-means Clustering

Limitations of Data Mining

While data mining products can be very powerful tools, they are not self-sufficient applications. To be successful, data mining requires skilled technical and

Analytical specialists who can structure the analysis and interpret the output that is created. Consequently, the limitations of data mining are primarily data or personnel related, rather than technology-related. Although data mining can help reveal patterns and relationships, it does not tell the user the value or significance of these patterns.

Asha Nagoriya

Assistant Professor
MITSOM COLLEGE, Pune

