

Role of e-governance - Importance of e-governance

The role is to deliver effective administration transparency and public service. It gives varied department's info to the public and helps in decision making.

- It ensures citizen participation at all levels of governance. It leads to automated services so that all works of public welfare is available to all citizens.
- Each department and its action is closely monitored. Public can get their work smartly done and save their time. It provides better services to citizens and brings government close to public. Public can be intouch with the government agency. It cuts middle men and bribery if any from the picture.

### Challenges

1. Technical - There are technical problems in implementing e-governance transformation. Info. and communication technology infrastructure is the main challenge for electronic governance transformation. Architecture should be in place to provide a uniform set of guiding principles, models and standards. Governments must build an effective telecommunication infrastructure.
2. Privacy - Privacy and security of confidential reports is the main problem. There are concerns

about hacking defense info. and other data, tracking the website, mismanagement of private info. are big challenges.

3. Security - Security of info. and systems against disclosure or destruction of data is one of the main concern. It includes the security of documents and network security.

Maintenance and e-infrastructure should be protected in the form of firewalls and data accessibility should be limited. The security tech like encryption and digital signatures should be used to protect user id's, credit card no., passwords and other data.

4. Lack of training and qualified personnel - Lack of digital skills is one of the main challenges in developing countries. Appointing and training the personnel to acquire appropriate and technical and functional skills is very much essential for governance to be successful.

5. Digital availability - Accessing the internet might be a challenge. Since there is a lack of internet skills and internet availability, all citizens cannot access the internet. Lack of computer literacy is one of the main challenges.

Conclusion:- Digital transformation of e-governance has the potential to administrate systematically

It is an instrument to run a government for its citizen. The services can be provided to citizens on time and on demand.

### Future of E-commerce -

Cloud computing provides a new service consumption and delivery model inspired by Consumer Internet Services. Cloud computing drives down costs and accelerates cost reduction benefit. Cloud is making rapid inroads. E-Governance with cloud computing offers integration mang. with automated problem resolution, manage security end to end, and helps budget based on actual usage of data. At a global level, Cloud arch. can benefit government to reduce duplicate efforts and increase effective utilization of resources. This in turn helps the government going green, reducing pollution and effective waste mang. Enterprises and Small and Medium businesses are already reaping the benefits of cloud by using the pay-as-you-use service model, its massive scalability and ready availability. Since government requires a massive infrastructure it is imp. for government to use cloud computing on long term basis. Cloud arch. allow rapid deployment of turn key test environments with little or no customization. India's

National Informatics Centre (NIC) a division of the dept. of info. tech., has selected and deployed the open source eucalyptus software as the foundation for its cloud project, which calls for the execution of cloud-based e-governance projects on a broad scale. NIC is providing the network backbone and a wide range of ICT (Information & Communication Technologies) Services to government org. throughout India, including a nationwide communication network for decentralized planning, improvement in government services and wider transparency of national and local government institutions.

Eg of government investing in Cloud - European commission - Cloud computing facilitates public procurement among different states' member administrations and enables small and medium enterprises to gain access to public services

(2) Japan - A nation-wide "Kasumigaseki Cloud" is being developed to enable various Ministries to collaborate.

(3) Singapore - Cloud Computing is viewed as a major source of economic development.

Source: Cross-government call on cloud computing, the World Economic Forum. As majority of the commercial app. are now on cloud environment, so it can be potentially predicted that cloud

computing is going to be one of the preferred platform in both developed and developing countries.

The domains where the e-governance cloud computing services would be beneficial are -

#### 1. Unique Identification Authority of India (UIDAI) -

It is an integrated centralised project initiated by government of India to have all the details of the citizens of the country in one repository.

#### 2. Centralized Auditing -

Since cloud offers its applications to be centrally operated, auditing of the proofs and with less cost and expenditure could be done with less cost and more effectiveness.

#### 3. Management information system -

It offers integrated decision making platform for any Electronic Procurement, Invoice Processing and Stock.

#### 4. Deployment of citizen services -

Deployment of citizen's services can be implemented by adoption of cloud computing in less time and more effectively.

#### 5. Agriculture - Farmers forms the strength of any country. Agriculture is the domain where cloud computing can efficiently act as mediator or source of info. for the practices & research

that are being carried out in other countries by the farmers.

6. Education- Enabling cloud for education will provide us the <sup>best</sup> trainings and practices adopted by overseas educational hubs.

#### Interactive service model / (G2C2G) - ~~etc~~

Interactive service model is a consolidation of the other digital governance models and opens up possibilities of individuals in governance processes.

ICTs have the potential to bring every individual into a digital network and enable interactive flow of info. among them.

This potential of ICTs is fully leveraged in this model. As the participation is direct and not through representatives, it can bring greater objectivity and transparency in decision making processes, and give a greater feeling of involvement & empowerment, provided that individuals are willing to engage in the governance processes.

Under this model, the various services offered by the Government become directly available to its citizens in an interactive manner. It does so by opening up an interactive Government to Consumer to Government (G2C2G) channel in various aspects of governance such as election of government officials (e-ballots); decision to make on specific issues (eg.- health plans), delivery of

individualised government services, gauging public mood and opinions, targeting specific communities for specific governance advise or services bringing mass awareness.

### Applications -

1. To establish an interactive communication channels with key policy-makers and members of planning commission.
2. To conduct electronic ballots for the election of government officials and other office bearers.
3. To conduct public debates / opinion polls on issues of wider concern before formulation of policies and legislative frameworks.
4. Filing of grievances, feedback and reports by citizens with the concerned governmental body.
5. Establishing decentralized forms of governance.
6. Performing governance function online such as revenue collection, filing of taxes, governmental procurement, payment transfer, etc.

### E-Advocacy / Mobilization & Lobbying Model -

E-Advocacy / Mobilization & Lobbying model is one of the most frequently used Digital Governance model and has often come to the aid of the global civil society to impact on global decision making processes

The model is based on setting-up a planned,

directed flow of information to build strong virtual allies to complement actions in the real world. Virtual communities are formed which share similar values and concerns, & these communities in turn link up with or support real-life groups/activities for concerned action. The model builds the momentum of real-world processes by adding the opinions and concerns expressed by virtual communities.

The strength of this model is in its diversity of the virtual community, and the ideas, expertise and resources accumulated through this virtual form of networking. The model is able to mobilize and leverage human resources and info. beyond geographical, institutional and bureaucratic barriers, and use it for concerned action.

### Applications-

This model could be applied in the following possible ways-

1. Fostering public debates on issue of larger concerns, namely on the themes of upcoming conferences, treaties, etc
2. Formation of pressure groups on key issues to force decision-makers to take their concerns into cognisance.
3. Making available opinions of a suppressed group

who are not involved in the decision-making process into wider public domain

4. catalysing wider participation in decision-making processes.

5. Building up global expertise on a particular theme in absence of localised info. to aid decision-making.

### Maturity Model

The adoption and use of e-government services in developing countries is still a big issue.

A digital divide among the people, poorly offered e-government services and availability and access to the tech. by the people are some of the critical issues faced with e-government projects. Although occurrences of these issues in the developing countries are inevitable, if appropriate measures are taken when designing e-government projects, governments would be able to achieve wider participation from stakeholders in e-government activities. The design of e-government projects is driven by e-government maturity models, which are often called stage models. A maturity model is a conceptual framework that outlines how e-government projects should be assimilated in stages. Considering the enormity and complexity of e-government projects, it is often assimilated in stages.

Several efforts have been made to standardize e-government assimilation stages and to measure the level of maturity than an e-government project has achieved. This has resulted in the development of the several e-government maturity models. While existing e-government maturity models have been adopted to develop strategic plans to deploy e-government projects.

Maturity levels -

These 5 levels helps CIOs (Chief Info. improve the quality of digital government services for citizens. Choosing the direction of digital transformation, communicating that vision & justifying the necessary budget requests are the new challenges for modern CIOs. This applies equally to public sector CIOs. Overcoming these challenges requires a step by step strategy that is both affordable and sustainable.

1. Maturity Level 1 - Initial (E-Government) -

At this level, the focus is on moving services online for user convenience and cost savings, but data and its uses are extremely limited.

"If the organisational view is that a high percentage of online services or mobile access represents a modern digital government, then more

education & advocacy is needed to show what real digital government looks like and its benefits". To make the case for advancement, create case studies explaining how digital transformation will ease or remove high priority pain points for the organisation.

## 2. Maturity Level 2 - Developing (Open) -

E-government and open government programs often coexist with different leadership and priorities. Open government often takes the form of public-facing programs intended to promote transparency, citizen engagement and the data economy.

## 3. Maturity Level 3 - Defined (Data-Centric) -

On this level the focus shifts from simply listening to citizen or user needs to proactively exploring the new possibilities inherent in strategically collecting and leveraging data. The key performance indicators here are "how much of our data is open?" and "how many of our applications are built on open data?" It's tempting at this point to engage in vanity projects or skip ahead before the proper groundwork is laid; it's paramount to remain focused on designing and implementing data-centric strategies and processes.

## 4. Maturity Level 4 - Managed (Fully Digital) -

By this level, the organization, agency or department has fully committed to a data centric approach to improving government, and the preferred approach to innovation is based on open data principles. Data flows regularly across organization boundaries, leading to easier interactions and better services for constituents. It's possible at this stage to encounter privacy-related backlashes, as citizens can be uncomfortable with how their data is being collected and used.

Therefore, it is imp. to ensure that data is used within existing norms and regulations, and that this is clearly communicated.

##### 5. Level 5 - Optimizing (Smart) -

At this point, the process of digital innovation using open data is embedded deeply across the entire government, with buy in and leadership from the top tier of policymakers.

The innovation process is predictable and repeatable, even in the face of disruptions or sudden events of require rapid responses.

Types of government interaction in E-Governance-

1. G2C (Government to Citizen)

(G2C)

Government to citizen initiatives-

1. Computerization of land records. - In collaboration with NIC ensuring that land owners get computerised copies of ownership, crop and updated copies of Records of Rights (RORS) on demand.
2. Bhoomi Project - Online delivery of land records. Self sustainable e-governance project for the computerised delivery of 20 million rural land records to 6.7 million farmers through 177 Government-owned kiosks in the State of Karnataka
3. Gyandoot - It is an intranet based government to citizen delivery initiative. It was initiated in the Dhar district of MP in January 2000 with the twin objective of providing relevant info. to the rural population and acting as an interface b/w the district administration and the people.
4. Lokvani project in UP - Lokvani is a public-private partnership project at Sitapur district in UP which was initiated in November, 2004. Its objective is to provide a single window, self-sustainable e-Governance sol. with regard to handling of grievances, land record maintenance and providing a mixtue of essential services
5. Project FRIENDS in Kerala - FRIENDS (Fast, Reliable, Instant, Efficient Network for the Disbursement of Services) is a single window facility

providing citizens the means to pay taxes and other financial dues to the State Government. The services are provided through FRIENDS Janasevana Kendrums located in the district headquarters.

6. e-Mitra project in Rajasthan - e-Mitra is an integrated project to facilitate the urban and rural masses with maximum possible services related to different state government departments through Lokmitra - Janmitra centers kiosks.

## UNIT - 4

Data warehousing & data mining -  
Data warehousing is the process of constructing and using a data warehouse. A data warehouse is constructed by integrating data from multiple heterogeneous sources that support analytical reporting, structured &/or ad <sup>hoc</sup> queries, and decision making. Data warehousing involves data cleaning, data integration & data consolidation.

The term "data warehouse" was first coined by Bill Inmon in 1990. According to Inmon a data warehouse is a subject oriented, integrated, time-variant and non-volatile collection of data. This data helps analyst to take informed decisions in an organisation.

A data warehouse provides us generalised and consolidated data in multidimensional view. Along with generalised & consolidated view of data, a data warehouse also provides us online analytical processing tools. These tools help us in interactive and effective analysis of data in multidimensional space. This analysis results

A data warehouse works as a central repository place where info arrives from one or more data sources. Data flows into a data warehouse from the transactional system & other relational databases. Data maybe structured, semi-structured & unstructured data.

The data is processed, transformed, so that users & integrated

can access the processed data in the data warehouse through business intelligence tools, SQL clients & spreadsheets. A data warehouse merges info. coming from different sources into one comprehensive database.

By merging all of this info. in one place, an org. can analyze its customers more holistically. This helps to ensure that it has considered all the info. available.

Data mining is defined as extracting info. from huge sets of data. It is the procedure of mining knowledge from data. Data mining deals with the kind of patterns that can be mined. On the basis of the kind of data to be mined, there are two categories of functions involved in data mining.

1. Descriptive
2. Classification & prediction

Descriptive Function - It deals with the general properties of data in the database. Descriptive functions are:

1. Class / Concept description - Refers to the data to be associated with the classes or concepts. These descriptions can be derived by two ways:-
  - a. Data Characterization
  - b. Data Discrimination

2. Mining of frequent patterns - Frequent patterns are those patterns that occur frequently in transactional data. Kinds of frequent patterns:
- a. Frequent item set
  - b. Frequent subsequence
  - c. Frequent sub structure
3. Mining of Associations - Associations are used in retail sales to identify patterns that are frequently purchased together.
4. Mining of correlations - It is a kind of additional analysis performed to uncover interesting statistical correlations b/w associated attributes value pairs or b/w two item sets to analyze that if they have +ve, -ve or no effect on each other.
5. Mining of clusters - Cluster refers to a group of similar kind of objects. Cluster analysis refers to forming group of objects that are very similar to each other but are highly different from the objects in other clusters.

Classification and prediction - Classification is the process of finding a model that describes the data classes or concepts. The purpose is to able to use this model to predict the class of objects whose class label is unknown. Functions involved are:- classification, Prediction, outlier analysis and evolution analysis.

Types of data warehouse -

1. Enterprise data warehouse
2. Operational data store
3. Data mart

Enterprise data warehouse (EDW) -

Enterprise data warehouse is a centralised warehouse. It provides decision support service across the enterprise. It offers a unified approach for organizing & representing data. It also provide the ability to classify data according to the subject and give access according to those division.

Operational data store (ODS) -

It is nothing but data store required when neither datawarehouse nor online transactional processing (OLTP) system support org. reporting needs. In ODS, data warehouse is refreshed in real time. Hence, it is widely preferred for routine activities like storing records of the employees. This has a broad enterprise wide scope and used ~~as~~ for a routine business activity.

Data mart

It is a subset of a data warehouse. It specially designed for a particular line of business, such as sales, finance. In an independent data mart,

data can collect directly from sources  
Data warehouse and data marts are built on  
dimensional data modelling where fact table  
are

National warehouse census data -

A large no. of national data warehouses can be  
identified from the existing data resources  
from within the Central Government ministries.

1. The Registrar General and Census Commissioner of India decennially compiles info. of all individuals, villages, population groups, etc.
2. This info. is wide ranging such as the individual-slip, compilation of info. of individual households of which a database of 5% is maintained for analysis.
3. A data warehouse can be built from this database upon which OLAP techniques can be applied.
4. Data mining also can be performed for analysis and knowledge discovery.

Nation warehouse prices of essential commodities -

1. The ministry of food & civil supplies, government of India compiles:
  - a. Daily data (on weekly basis) for about.
  - b. 300 observation centers in the entire country.
  - c. On the prices of essential commodities such as rice,

edible oil, etc.

2. This data is compiled at the district level by the respective state government agencies and transmitted online to Delhi for aggregation and storage.
3. A data warehouse can be built for this data & OLAP techniques can be applied for its analysis.
4. A data mining & forecasting technique can be applied for advanced forecasting of the actual prices of these essential commodities.
5. The forecasting model can be strengthened for more accurate forecasting by taking into account the external factors such as rainfall, growth rate of population & inflation.

Data mining in the field of agriculture -

In agriculture sector where farmers and every businesses have to make innumerable decisions everyday and intricate complexities involves the various factors influencing them. An essential issue for agricultural planning intention is the accurate yield estimation for the numerous crops involved in the planning. Data mining techniques are necessary approach for accomplishing practical & effective sol. for this problem. Environmental cond<sup>n</sup> and variability in soil, input levels, combinations & commodity prices have made it all the more relevant for farmers to use info.

get and help to make critical farming decision. This focuses on the analysis of the agriculture data and finding optimal parameters to maximize the crop production using data mining techniques like PAM, CLARA, DBSCAN and multiple linear regression.

Mining the large amount of existing crop, soil and climatic data, and analysing new, non-experimental data optimizes the production and makes agriculture more resilient to climate change.

Today, India ranks second worldwide in the farm output. Agriculture is demographically the broadest economic sector and plays a significant role in the overall socio-economic fabric of India. Agriculture is a unique business crop production which is dependent on many climate and economy factors. Some of the factors on which agriculture is dependent are soil, climate, cultivation, irrigation, fertilizers, temp., rainfall, harvesting, pesticide weeds and other factors. Historical crop yield info. is all imp. for supply chain operation of companies engaged in industries. These industries use agriculture products as raw material, livestock, food, animal feed, chemical, poultry, fertilizers, pesticides and seeds. An accurate estimate of crop production and risk helps these companies in planning supply

change chain decision like production scheduling. Business such as seed, fertilizer, agrochemical and agricultural machinery industries plan production and marketing activities based on crop production estimates.

There are two factors which are helpful for the farmers and the government in decision making namely:-

- a. It helps farmers in providing the historical crop yield record with a forecast reducing the risk management.
- b. It helps the government in making crop insurance policies and policies for supply chain operation.

Data mining technique plays a vital role in the analysis of data. Data mining is the computing process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database system.

Unsupervised and supervised are 2 different types of learning methods in the data mining.

Data mining in rural development.-

The development of rural e-commerce and rural tourism has provided new ideal ideas for rural economic construction. The rapid expansion of rural e-commerce represented by Taobao Village

shows people the enormous potential of the rural e-commerce market. While e-commerce has changed the structure of rural economic development. With the rapid development of the economy and the vigorous promotion of the info age, products based on high-tech are constantly blended with traditional industries. While providing convenient services to people, it also greatly advances the process of informationization. The rapid economy development is also actively promoting the prosperity of the tourism industry. The heavy life pressure and the stereotyped daily routine of the city have prompted more and more people to join the rural tourism during weekend and holiday. Rural culture atmosphere and relaxed and comfortable environment also greatly enhance their quality of life.

The research data is mainly from research report data of China's tourism villages, the statistics of the National Bureau of Statistics of China, and the government's public data of some provinces. Acc. to the standards set by the Ali Institute, Taobao Village refers to a rural area with an administrative village as its unit, with more than 100 online shops or more than 10% of the active online shops occupying more than 10% of its administrative village households and the E-commerce transactions amounted to more than 100 million yuan.

Demonstration tourism villages are based on  
the government's list.

The time dimension of this study is from  
2014 to 2017.

## Data mining in health -

Electronic health records are dynamically turning out to be more popular among health establishment. With improved access to a considerable amount of patient data, healthcare firms are now in a position to maximize the performance and quality of their businesses with the help of data mining.

Since the 1990s, companies have utilized data mining for activities like credit scoring as well as fraud identification and a lot more. Nowadays several healthcare businesses across the globe are likewise starting to realize the potential advantages of medical data mining along with the predictive analytics.

The medical industry collects a dazzling array of data, most of which is electronic health records (EHRs) collected by HIPAA covered healthcare facilities. According to a survey published by PubMed, data mining is becoming increasingly popular in health care. The huge amount of data generated by healthcare EDI transactions cannot be processed and analyzed

using traditional methods bcz of the complexity and volume of the data.

Data mining provides the methodology & technology for healthcare organizations to:

- evaluate treatment effectiveness.
- save lives of patients using predictive medicine.
- manage healthcare at different levels.
- manage customer relationship, detect waste, fraud & abuse.

To shift through the collected medical data and to extract the useful knowledge hidden there, data mining is used as a part of the Knowledge Discovery in Databases (KDD) process.

Data mining involves the creation of association rules, the use of support and confidence criteria to locate the most imp. relationships within the data.

Other healthcare data mining parameters include:

- sequence or path analysis
- classification
- clustering, and
- forecasting

The healthcare industry possesses rich data sources, such as electronic medical records, administrative reports and other benchmarking findings.

Today, data mining in healthcare is used

mainly for predicting various diseases, assisting with diagnosis and advising doctors in making clinical decisions. But, the potential of data mining is much bigger - it can provide question-based answers, anomaly-based discoveries, provide more informed decisions, probability measures, predictive modeling, and decision support.

Using data mining, the healthcare industry can be very effective in such fields as:

- medical research
- pharmaceuticals
- medical devices
- genetics
- hospital management
- health care insurance, etc

Data mining in planning -

Recently, many companies have recognised data mining as an imp. technique that will have an impact on the performance in industries or companies.

However, there are many challenges to be faced when it involved industrial data such as ability to handle different types of data, graceful degradation of data mining algorithms, valuable data mining results, mining at different abstraction levels of data and different sources and protection of data security.

Environment in manufacturing fields are complex, dynamic, and have the stochastic systems. There are a no. of activities involves in managing and operating production system. Those activities include organizing work, selecting processes, arranging layouts, designing jobs, scheduling work and planning production.

Meanwhile, different stages of planning involve the operation such as process planning, capacity planning, aggregate planning, master scheduling, material requirements planning and order scheduling. These planning stages can be divided into three time frames; short, medium & long term. In business forecasting short-term usually refers to under 3 months, medium-term or intermediate, <sup>3 month to</sup> greater than 2 years. Long-range planning is generally done annually, focusing on range that greater than 1 yr. Process planning & strategic capacity <sup>planning</sup> are the major operations in long range term. Process planning converts the design into workable instructions for manufacture. It deals with determining the specific technologies and procedures required to produce a product or service. For the strategic capacity planning, it deals with <sup>1</sup> - term capabilities of the production system. It extends over a time horizon long enough to obtain those resources that affect product lead times, customer responsiveness, operating costs and a firm's ability to compete. So, for building the new facilities or acquiring new business, it suitable uses a capacity planning. For intermediate-range, there are aggregate planning, master scheduling and material

requirements planning. Intermediate - range planning usually covers a period from 3 to 18 months, with time increments that are weekly, monthly or sometimes quarterly. The output from the aggregate planning is the feasibility to hire or lay off workers, increase or reduce the work-week, add an extra shift, subcontract out work, use overtime or build up and deplete inventory levels. Usually aggregate planning's aim is to minimize total cost over the planning horizon include inventory investment, workforce levels and production rates. Master production schedule is a part of the material requirements plan that show the details of how many end items will be produced within specified periods of time. Material production schedule (MPS) is one of the imp. activity in manufacturing planning and control and works within the constraints of the production plan but produces more specific schedule. It breaks the aggregate production plan into specific product schedule, maintaining customer service levels and stabilizing production planning within a Manufacturing Resource Planning environment and a just-in-time base production systems.

Short - range planning covers a period from one day or less to six months, with the time

increment daily or weekly and bcz of that, operation planning activity in short - range is order scheduling. In individual machines, they are shop floor schedule to maintain and communication status info on shop orders and work center. Shop floor scheduling is the allocations of resources and the sequencing of tasks to produce goods and services, also to specify the time for each job starts and completes on each machines.

increment daily or weekly and bcz of that, operation planning activity in short range is order scheduling. In individual machines, there are shop floor schedule to maintain and communication status info. on shop orders and work center. Shop floor scheduling is the allocations of resources and the sequencing of tasks to produce goods and services, also to specify the time for each job starts and completes on each machine.

### Data mining in education -

One of the primary goals of any educational system is to equip students with the knowledge and skills needed to transition into successful careers within a specified period. How effectively global educational systems meet this goal is a major determinant of both economic and social progress. Some countries provide free education for all citizens from grade one through the university. Therefore, a large no. of students enter universities every year. For eg:- King Khalid University (KKU) accepted approximately 23000 students in 2013. It has become difficult to provide high quality teaching and guidance to such a large no. of students. Using datamining techniques to analyze students information can help identify positive

reasons for student failures. Data mining provides many techniques for data analysis. The large amount of data currently in student databases exceeds the human ability to analyze and extract the most useful info. without help from automated analysis techniques. Knowledge discovery is the process of non-trivial extraction of implicit, unknown and potentially useful info. from a large database. Data

mining has been used in knowledge Discovery<sup>(KD)</sup> to discover patterns with respect to users needs. The pattern definition is an expression in lang. that describes a subset of data.

Educational data mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings and using those methods to better understand students and the settings which they learn in. Different from data mining methods, EDM (Educational data mining) when used explicitly, accounts for the multi level hierarchy and lacks independent educational data. EDM methods comes from different literature sources including data mining, ML, psychometrics and other areas of computational modelling, statistics and info. visualisation. Work in EDM can be divided into 2 main

categories :- <sup>①</sup>Web mining & <sup>②</sup>statistics and visualisation. Another point of view proposed by Baker classified the work in EDM as follows :-

1. Prediction
  - a) Classification
  - b) Regression
  - c) Density estimation
2. Clustering
3. Relationship mining
  - a) Association rule mining
  - b) Correlation mining
  - c) Sequential pattern mining
  - d) Causal data mining
4. Discovery with models