# Information Sharing for e-Government

## Information Systems

Pablo Fillottrani  Elsa Estévez

Center for Electronic Governance
United Nations University - International Institute for Software Technology

July 2010





## Information Systems







## What is an Information System?

- an Information System is any combination of information technology and people's activities using that technology to support operations, management, and decision-making
- in a very broad sense, the term information system is frequently used to refer to the interaction between people, algorithmic processes, data and technology
- the term is used to refer not only to the information and communication technology (ICT) an organization uses, but also to the way in which people interact with this technology in support of business processes
- some make a clear distinction between information systems, ICT and business processes
  - an information system is typically seen as having an ICT component
  - information systems help to business processes





## Motivations

- business operation support processes business transaction of the organization transaction can be any activity of the organization
- decision making support assist lower management in problem solving and making decisions, using results of transaction processing and some other information also
- develop strategies to reach competitive advantages assist higher management to make long term decisions, handling unstructured or semi structured decisions. unstructured decision is one which there is no clear procedure for making it and if not all the factors to be considered can be readily identified in advance

## Information as a Business Resource

- when asked to identify key resurces in any business, businessmen will readily name people, equipment, money and buildings; they spend most time managing these resources
- information is often missing from the list, but today business cannot function without information
- moreover, information is difficult to assess, manage and measure although important management decision are based on it
- but information is a business resource that is used in every aspect of any business





## Information as a Business Resource

- information is routinely used to monitor the efficiency and effectiveness of the business, in the form of reports or inside information systems
- all business need to plan for their future and take high-level strategic decisions that rely on information about historic performance of the enterprise, projected future performance, custumers' needs and competitors' performance
- therefore, information is needed in every level of the enterprise and it is important to manage it in a consistent, accurate, timely and easily understood way
- infomation is data placed in context





## People

- the information systems department is typically responsible for maintaining IT infrastructure. It includes
  - programmers highly trained technical specialists
  - system analysts constitute the liasons between the IT group and the rest of the organization
  - information systems managers are leaders of information projects
  - chief information officer (CIO) is a senior managament position that oversees th use of IT in the organization





## Organizations and Information Systems

- information systems are built by managers to serve the interests of the organization
- the organization, must be aware of and open to the influences of information systems to benefit from new technologies
- factors that influence the interaction
  - organization's structure
  - standard business processes
  - surrounding environment (culture, politics)
  - management decisions

## Opportunities

- organization flattening large, bureaucratic organizations have downsized, reducing the number of levels in their hierarchies after information systems introduction, by broadening the distribution of information to empower low-level employess and increase management efficiency, and by allowing managers to receive much more accurate information on time, taking faster decision
- virtual organizations in which work is no longer tied to geographic location, linking suppliers, customers and sometimes even competitors
- increasing flexibility of organizations maximising the ability to sense and respond to changes, and taking advantage of new opportunities
- understanding organizational resistance to change introducing information systems require change in personal, individual routines can be painful are retraining and additional effort is often required





## Challenges, solutions

- Challenges
  - difficulties of sustaining competitive advantage competitive advantages do not last forever
  - difficulties of managing system-related change the natural inertia of organizations complicates the introduction of new tecnology
- Solutions
  - performing a strategic systems analysis identifying the types of systems that would provide a strategic advantage
  - managing strategic transitions changes in business goals, relationships with customers and suppliers, internal operations and business processes are necessary
  - promoting better management of organizatio's assets increasing revenue and reducing operating costs
  - creating work groups outside traditional places of work information systems lower agency costs, enabling the firm to manage more employees with fewer resources





## Data security

- data security is about protecting data against unauthorised users, maintaining privacy
- there should be an enterprise-wide data security policy in place, clear and concise
- policy is enforced by operating systems and DBMSs
- access control most users only need to access to a subset of the available data. They rely on authentication procedures such as logins and passwords
- groups of data access rights have to be defined
  - discretionary access control is where users who are granted access rights are allowed to propagate those rights to other users
  - mandatory access control is where access rights cannot be changed by users
- audit trails are included in most DBMS, recording what database objects were accessed by whom and when



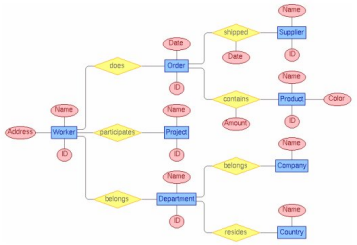


## Data integrity

- data integrity is about protecting the database against authorised users, enforcing constraints and consitency
- integrity constraints can be inherent to the underlying database model (for example, no component of the primary key is allowed to be null), or encoded in the logical schema

## Origins of data modeling (I)

- in the early days, focus of system development were business processes. Data was simple a byproduct of a process
- as technology evolved, methodologies were invented that evolved around processes (e.g. data-flow diagrams, structure charts)
- Chen's entity-relationship model (E/R) broke this pattern by introducing data as the center of business analysis and system design








## Origins of data modeling (II)

- at the same time, Codd's relational model develop the theory of relational database management system (RDBMS)
- the power of RDBMS together with the E/R modelling tool provided:
  - process-independence
  - business-focus data analysis
  - easy transition to logical model by normalization rules





## Roles of a Data Model

- understanding information requirements data model is used to document the information used by the business to highlight inconsistencies in existing systems, or to agree all the information requirements in systems to be developed. This model should be the result of pure analysis, untainted by design
- basis for physical database design once the conceptual data model is complete, it provides the start point for the desing of the physical database. The normalized data model could be easily translated into a relational database design. This process has two steps:
  1. first-cut database design the aim is to use the conceptual constructs to develop a relational design
  2. optimised database design in order to improve performance and other non-functional requirements, first-cut design needs to be enhaced by making use of built-in capabilities in RDBMS





## Corporate data modelling

- a corporate data model is similar to a project-level conceptual data modeling, but it has a much broader scope
- it involves all data used by the enterprise, there are no project data outside this model
- it will not necessarily be used as the basis for information system development. It expresses a view of data from the perspective of business
- it will ensure that data is commonly and unambiguously defined in all infomation systems, facilitating the sharing of data
- it provides a standard with which all systems must comply their interfaces with other information systems
- this approach means that database design migh be optimised, and not store exactly as in the corporate data model. Then extra processing to translation might be necessary
- it acts as a common database design, covering different areas

## Concepts

- corporate data model includes entities, attributes and relationships as any data model
- but some further concepts are necessary:
  - entity subtype is a generalization between entities, each instance of the subtype is also an instance of the supertype. For example, entity Cashier is subtype of entity Employee
  - mutually exclusive subtype relationships subtypes may be mutually exclusive in the sense that an instance of a subtype cannot be instance of other subtype. For example, Cashier and Administrative are subtypes of Employee, and no Cashier can be Administrative
  - complete subtype relationships all instances of supertype are instances of one subtype. For example, all Employess must be either Cashier or Administrative





## How to develop a data model

1. attribute-trawling involves studying all existing information systems, collecting data definitions (probably not documented), and sorting them out to obtain reusable definitions. Problems: there may be areas not supported by systems; there may be systems that do not actually meet users' expectations; it is unclear how definitions wil be analysed and compared; there may be too many to handle manually
2. joining project or area models involves independently modeling of data from the separate areas, and then amalgamating them. Problem: identifying common points where the models could join
3. top-down approach implies the development of a single conceptual data model (a "framewor model") covering core data requirements of the enterprise.Then it is used as the skeleton for separate project or area models, which are amalgamated later. Problem: how to build the starter model





## Corporate Data Model Principles

- develop the model "top-down" at least, in order to generate a "framework" data model, and then filling it to build area models
- give primacy to core business to correctly focus the process
- cover the whole enterprise this is to ensure no data requirements are missing. This is balanced with previous principle
- "future-proof" the model it must represent the true underlying nature of the information, and not how it is used at the time of the analysis
- develop cooperatively the data modelling team cannot work in isolation, consulting technicians and users from different areas
- gain consensus, not perfection there is a danger when a team seeks to develop the perfect model. The team should be prepared to publish and support the model in all business areas





## What is metadata?

- metadata is an integral part to information management because it provides contextual information for business data
- examples of contextual information are
  - the meaning and content of the data
  - policies and business rules that govern the data
  - technical attributes of the data
  - manipulation and usage of the data
  - lineage of the data

## Metadata initiatives

- no one will argue that data is critical to the business, yet data alone does not translate into business intelligence without metadata
- metadata provides the context for business data, facilitating technicians to manage systems and end users to use them, locating critical business data, relying on its data value, trusting on its accuracy
- there must be a management support to metadata





## Usage of metadata

- metadata should cover
  - what physical data is currently in our files and databases?
  - what does each data element mean in both business and technical language?
  - where is data located, and in how many different places?
  - how did it get into those files or databases?
  - how data can be accessed?
  - who own the data?
  - who is responsible for the content?
  - who updated it?
  - does each data element have a unique name?
  - are there privacy restrictions?





## Metadata initiative components

- you can start a metadata management initiative with the following components:
  - a database for metadata, or a metadata repository. Its tables represent metadata objects like Entity, Table, Attribute, Key. Its columns represent meta metadata like Name, Type, Definition, Length, Domain
  - a metadata administrator the person who creates and maintains the metadata repository, collecting, udpating, and integrating metadata from its various sources, and producing metadata reports
  - several policies assertions, or rules to be followed, set by the business to achieve or support a specific business goal. The business objective for managing metadata is to have control over business data like any other business asset
  - several procedures or practices that people perform to ensure they follow the policies





## Metadata classification

- metadata can be classified in
  1. business oriented business names, definitions and rules. Provided by end-users, business people and business analysts.
  2. technical table and column name, datatype, referential integrity rules. Provided by technicians during developing.
  3. process-related program names, transformation logic, or refresh schedule. Also provided by technicians during design and developing.
  4. usage includes by whom data is accessed, when, and for what purpose. Provided by technicians, not by end-users, by means of applications monitoring tools.

## Metadata repositories

- metadata repositories can be
  - off-the-shelf ready to use in a shorter time, but rarely satisfying all requirements
  - specifically built customized to business requirements, may be incrementally developed and needs full-time staff to develop and maintain
  - centralized eliminating metadata redundancies, ensuring consistency and consistency but requiring a flexible strategy to maintain
  - distributed simple metadata repositories stored together with data repositories, but lacking control and synchronism





## Data naming and conventions

- a data naming convention provide consistent, unique and meaningful names for all existing and new data elements. Conventions should be well known and easily followed
- abreviations and technical terms should be avoided
- problems:
  - conventions might be over-prescriptive
  - conventions may not deliver what is expected
- a thesaurus or controlled vocabulary is necessary





## Metadata Management

- archiving and purging define requirements and triggers that will archive or purge older or outdated metadata
- backup similar as a common database backup
- enhacements metadata expands and updates incrementally over time
- database tuning if access paths change or users grows, then databases may have to be tuned
- recovery from hardware failures and database crashes
- versioning it might be required to keep several versions of metadata on the repository to track history of changes