

# INFORMATION POLLUTION – THERMODYNAMIC INTERPRETATION OF LOSS OF INFORMATION INTEGRITY IN *IS* AND IN INFORMATION THEREFROM

(A Research and Knowledge Development Direction)

**Vijay V. Mandke**

Research Leader, Center For Information Integrity Research,  
Unitech Systems (India) Pvt. Ltd., New Delhi – 110 049, INDIA  
[vmandke@unitechsys.com](mailto:vmandke@unitechsys.com)

**Madhavan K. Nayar**

President, Unitech Systems, Inc., IL 60563, USA  
[mnayar@unitechsys.com](mailto:mnayar@unitechsys.com)

**Abstract:** Information is a higher order or derivative of matter and energy and depends on them for its existence. Informational errors lead to loss of Information Integrity. This results in inefficient and uneconomic material and energy processing. Implication of this is increase in the entropy of the system, which comprises both informational and physical work systems. Further, in the form of discharged system by-product wastes, it also contributes to the higher increase in the entropy of the system environment. This accelerates the time of *heat death* for the system. For the system designers and implementers this introduces the concern for the issue of increased environmental pollution due to the information pollution caused by informational errors in the open system that the information system (*IS*) is. What is significant is in the wake of the ever present system environmental factors of complexity, change, conversion, communication and corruption (5“C”s), this information pollution takes the form of a continual phenomenon. With reference to the research field of Environmental Pollution & Engineering, this modeling makes a case for a useful research direction in the area of Information Integrity Knowledge, Science, Technology and Industry Development for controlling information pollution

For a generic description, information pollution results from the loss of: information origination integrity, storage integrity, retrieval integrity and validation integrity. Also to be considered in the same vein is information pollution due to loss of: information processing integrity, communication and distribution integrity, *use* integrity, and information discard integrity or information storage for future *use* integrity. With reference to an information system (*IS*), information pollution is the issue caused by the loss of content integrity, process integrity, and system integrity. At a more functional level, informational errors in a design setting can be described as design errors, and so can be the informational errors in other settings covering the entire information system development and implementation life cycle (*ISDILC*) model. Then information pollution can be seen to result from loss of: design integrity, development integrity, deployment (testing) integrity, data and data product integrity, and detection (audit) integrity. This leads to the recognition of the reality of information pollution due to loss of system implementation integrity, operation integrity, maintenance integrity, etc. In summary, Information Integrity is a systems concept and recognition of eventuality of informational errors in different settings presents information pollution as a systems problem.

# 1. IS AND INFORMATION — A SOCIO-TECHNICAL VIEW

## 1.1 *IS and Real World, Information User, Informational Task of Decision Making, Information Purpose*

Information system (IS) exists in a real world. The real world consists of objects (some concrete, such as machines; and some abstract, such as accounts), humans, rules (established procedures, codes of practice, etc.), norms (often deeply engrained ways of designing, and implementing things and modes of thinking), and commands (e.g., computer programs, call of lists, etc.). the information user - internal or external to the *IS* - may be a human, a software, a machine or systems of theirs performing some task in relation to the real world. The task may be decision making, which may be simple or complex. The task may also be a mere message transmission as in a telephonic communication, where the issue of decision-making is to determine the objective of the sender. The objective of information is to use it to inform, evaluate, persuade, or organize other information. Creating new concepts, identifying and solving problems, decision-making, planning, initiating, controlling, and searching are just some of the purposes to which information is directed. The purpose of supplying information (not data) to machines is to provide instructions or to provide information for stored instructions to act upon.

## 1.2 *Information Sources of IS and Reality of IS as A Socio-technical System*

Most *IS* have these three major information sources: (i) the real world itself, which can be inspected; (ii) the designed *IS* – an artifact, which is intended to provide an exact (?) image of the real world; and (iii) an environment embedded *IS*. Specifically, the environment embedded *IS* is (a) concerned with data and information at meta level, (b) mostly operates at an informal level and sometimes substitutes for the designed system, and (c) often is used to provide qualitative and evaluative information about the real world. The extent to which information technology plays part is increasing rapidly but in the wake of analysis *here* it does not prevent *IS* being characterized as a socio-technical system delivering information for *use*. A socio-technical system is made up of relatively semi-autonomous and individualistic components - individuals, groups, enterprises – that can act for different or even cross-purposes. Similarly, then an *IS* consists of an enormous number of diverse and semiautonomous elements, loosely coupled in an intricate network of individuals and teams of people, hardware and software, procedures, regulations, communications, management policies and financial constraints that function in a variable and uncertain environment with diffused, decentralized management control.

# 2. IS – AN OPEN SYSTEM

## 2.1 *Closed System – A Thermodynamic Definition*

Thermodynamics defines a closed system as one with impermeable boundary and in which energy must remain constant. This is in keeping with the first law of thermodynamics, i. e., the principle of the conservation of energy, which states that the matter and energy can neither be created nor destroyed. Because this total energy must remain constant, a closed system can only *lose* its ability to do work. Inevitably, by the second law of thermodynamics, such systems must lapse into the state of totally unorganized, randomly distributed, inconvertible particles, which is the state of totally bound energy known as *heat death*.

## 2.2 *Open System – A Thermodynamic Definition*

How do then engineering systems sustain and some systems like organisms even increase their organization, continuing to do work for many years? The point is they operate as not closed systems *but*

as open systems. Thermodynamically an open system is defined as a system with continuous inputs and outputs of matter and energy. All the open systems, if they are to postpone for a time their inevitable *heat death*, must control the extraction and processing of matter, its internal distribution and storage, continuous conversion in to energy, and elimination as by-product wastes.

### ***2.3 Relationship between System, Information and Material & Energy processing in a Socio-technical System and significance of controlling information origination and processing***

Information is a higher order or derivative of matter and energy and depends on them for its existence. Because socio-technical system is loosely connected intricate network of semi-autonomous components and elements (see Sub-section (1.2)) and because social processing will depend on exchanges between them, it is reasonable to suggest that system processing in it will predominantly depend on information exchanged between its components and elements. Further, it is this information exchange *that* will then contribute to the significant flow of matter and energy in the socio-technical system. As a result, the socio-technical system, if it is to postpone for a time its inevitable *heat death*, (i) firstly must operate as an open system and (ii) secondly it must control the origination and processing of information between its components and sub-systems, its internal information distribution and storage, continuous conversion to energy, and elimination as by-product waste.

### ***2.4 IS with competitive advantage: Open System view of IS with requirement to originate and process information efficiently and economically***

In view of the socio-technical character of an information system (*IS*), the above then offers the thermodynamic basis for an *IS* with competitive advantage. But first what do we mean *here* by competitive advantage? Let us consider two information systems, namely, information system  $IS_1$  that sustains itself for time  $t_1$  and information system  $IS_2$ , which sustains itself for time  $t_2$ . Then  $IS_1$  is said to have a competitive advantage over  $IS_2$  if  $t_1 > t_2$ , i.e., if  $IS_1$  sustains itself for longer time and, preferably, even increases its information processing effectiveness and efficiency. In other words, for competitive advantage, an *IS* should avoid (postpone), to the maximum, time of its degenerating into the state of *heat death* wherein it (*IS*) has to function in an environment only full of inaccurate, inconsistent, unreliable and in situation even harmful data and information.

#### ***2.4.1 IS by-product wastes***

For this, the *IS* ought to continually originate, store, retrieve, evaluate, process, communicate, use, and subsequently discard or store for further *use* information efficiently and economically. Clearly, this view of an *IS* with competitive advantage is an open system view with the requirement of efficient and economic information processing. What this will enable the *IS* to control is its continuous conversion in to energy and elimination (degeneration) as by-product wastes. Some of the indicators of these wastes are: amount of computer paper consumed, size of staff performing repetitive informational work, difficulties in information distribution, negligible spread of information users making *use* of information, long time movement, presence of outdated unit record equipment, lack of documentation, existence of problem of information overload, difficulties in deciding, increased decision time cycles, frequent conflicts between results of two applications, lack of relationships between data in several applications, information failures that come with time delay, implications of latent errors made during *ISDL* cycle model, delivering of harmful information, etc.

### **3. BUSINESS PROCESS *IS* VIEW– INDIVIDUAL INFORMATION ORIGINATING AND PROCESSING SITUATION**

#### ***3.1 IS and Business Enterprise***

At a theoretical level, it is through markets that social processing advances in societies for promotion of common good of societal constituents. Business enterprises covering manufacturing, production and service industries across the spheres of human activity constitute all encompassing view of the markets. Understandably, business enterprises are socio-technical systems (see Sub-section (1.2)) and they perform in the real world, which is complex. It is to cope with this complexity that business enterprises develop systems.

In view of their socio-technical nature, business enterprises, too, demonstrate open system behavior. This increases importance of information in business. In economic terms, this leads to a larger and larger percentage of the total business resource being spent on information and information technologies. A business enterprise is an organization. System theory and economic theories suggest that complexity of an organization is ultimately limited by the amount of information that it can (economically) process, i.e., by the (costly) bandwidth of its communication channels. This resonates with the open system view of *IS* with competitive advantage (see Sub-section (2.4)).

Within the above framework, one then finds issues of enterprise (organization) management and of organizational strategies and decision-making intertwined with flow of information. For business activities/processes/procedures describing the enterprise, this suggests an information system (*IS*) view, designated as business process *IS* view.

#### ***3.2 Informational and Physical Work Systems***

Consider an entire supply chain from concept to delivery and follow up. This is a generic business process - a view applicable to manufacturing, production and service sectors. A business system is an orderly, harmonious group of interacting, interrelated, and interdependent procedural components. Further, it is an open system as it possesses a porous boundary with its environment, processes (i.e., imports as input and exports as output) information, and has an objective.

Accordingly, all business activities/processes/procedures are data processing procedures, that is, they all process data in some manner to deliver information for *use* in decision-making. With system integration, this makes for business activity *emphasizing* information and comprising informational and physical work systems. For competitive advantage, the requirement is to maximize informational work (IW) comprising the following activities: (a) generating from business process raw data/information in a complex and changing real world environment, which is characterized by uncertainty and hence errors, and (b) processing this information on current basis for undertaking the planning and evaluation of business design alternatives and selecting flexible information decision for control implementation at physical work system.

#### ***3.3 Multistage Decision Process based Model of Business Process *IS* View***

The research presentation entitled “Information Envelope and its Information Integrity Implications” at ICIQ 2001 studies modeling a generic business process as an integral to a closed loop information and control system, constituting a business process *IS* view. Most information processing involves some type of data conversion to information and is, therefore, closely related to a decision process with an objective.

As argued in Sub-section (1.1), even when the information is transmitted without changing form, as in a traditional communication system representation, the issue is to decide the purpose or objective of the transmission.

Traditionally, the decision process is viewed to comprise stages of forecasting, evaluation of alternatives and selection. Here information is considered basically as function of “source” (i.e. as “data”) and at most of “source” and “process”. However, the business process *IS* view is an open system; information being function of “source”, “process” and “recipient” (i.e. customer). For it a more workable model of a decision process spans multiple stages. These are: *obtaining* the ‘many factors’ & ‘multiple criteria’ characterizing business task (problem) complexity based on the long term goal set; *recognizing* the operable business goal from these multiple criteria; from the operable goal statement, *defining* the planning & design constraints and opportunity spaces; from the ‘many factor’ information variables characterizing problem complexity, *culling out* the useful (relevant) information variables; *recognizing* relationships (interdependencies) between the culled out information variables; *developing* the state transition models defining dynamic behavior of the culled out state (information) variables; and undertaking customized planning & design for *generating* the alternatives for evaluation and final selection of the flexible information decision for control implementation.

### ***3.4 Origination of Information and Alternatives Endogenous to Decision Situation – Structural Variant of Traditional Decision Situation***

What is significant is that this multistage *IS* view, among other stages, involves originating information in respect of the operable goal, generating alternatives, and selecting information in the form of the flexible information decision, all, “endogenous” to *the* decision situation. This is a structural variant from the traditional view of decision process, which (in the manner of a closed system) is concerned only with alternatives and information that are “already” generated, exogenous to the decision situation, and hence is a “collective” decision process. Thus, in multiple stage decision process based business process *IS* view, what we have before us, is an *individual information originating and processing situation*.

## **4. SYSTEM ENVIRONMENTAL FACTORS OF 5“C”**

As with traditional *IS* which is a “collective” decision process, this multiple stage decision process *IS* is also characterized by uncertainties due to system environmental factors of 5“C”, namely, complexity, change, communication, conversion, and corruption.

### ***4.1 Complexity***

It is to cope with the complexity of the real world environment that a modern business enterprise aims at increased bandwidth of its communication channels and processes increased information in the form of ‘many factors’ and ‘multiple criteria’ (see Sub-sections (3.1) and (3.3)). The more the variables and greater their interdependence, the greater that system’s complexity. Complexity places high demands on the information system’s (*IS*’s) capacity to set an operative goal, originate information, develop an information structure dynamics model to integrate findings, and design effective customized actions. This creates uncertainties and results in errors at various stages of the multistage decision process that the Business Process *IS* view is.

### ***4.2 Change***

As a socio-technical system, business service product is subject to change due extrinsic as well as intrinsic events such as societal pressures, legal and regulatory rules, component change, etc. Further, the business system must cope with very rapid advances in technology and practice. In this context, it may

also be mentioned that change in the global business system is accomplished laterally across several subsystems in which decision-making is distributed across many people and units. In such a diffused system, change can be a slow, often difficult process. There is yet another point of extreme significance - a system such as a global business operates in the real world, which is not passive but active. This invariably creates time pressures. All these change features leading to dynamics in components and elements of the business organization make it important to understand developmental tendencies therein. Further, it is not sufficient to observe and analyze (forecast) component and element information variables at any single moment of time but instead it must be determined as to, over a time, whereto the whole system is headed (trend or directional information). These change factor related demands on business process *IS* view, resulting in further requirements of information originating and processing, contribute to uncertainties and resulting errors in *IS*.

### ***4.3 Conversion, Communication, Corruption***

Then, as mentioned earlier, complex, modern business enterprise involves many human and machine interfaces-cum-interactions between vendors, designers, developers, decision makers, production, operation, maintenance, people and devices, technicians and equipments, and others. There are also requirements of manipulation (in a positive sense, that is). All these introduce system environmental factors of conversion (consolidation, decomposition or transformation of data), communication (movement of data/information within or across the enterprise) and corruption (poor motivation, desire for personal gain, carelessness, actions of people) in business *IS* environment. What is important for the investigation at hand is each encounter as this, indeed, each operation, technical maneuver, human-machine interface, or human and machine information processing creates uncertainty and presents an opportunity for error.

## **5. UNCERTAINTY IN BUSINESS PROCESS *IS* VIEW**

Specifically, acting externally and internally, 5“C”s introduce in *IS* uncertainties observed traditionally and beyond.

### ***5.1 Uncertainty at Physical Work System Level***

#### **5.1.1 At Plant Operations Stage and at Physical Variable Control Stage**

At physical operations stage and at physical variable control stage, uncertainties are introduced due to input noise, process parametric noise, and measurement noise.

#### **5.1.2 At All Control Levels**

At all control levels (i.e., physical variable, transaction processing and management decision controls, which in the wake of “application” emphasis are so impacted), uncertainties are due to information overload, lack of standardization, lack of relationship in data in several applications, errors in hardware, software errors, data entry errors, or accidental or intentional failures, etc.

#### **5.1.3 At Managerial Decision Level Controls and at Higher Level Controls**

Further, uncertainties are also introduced due to incomplete knowledge of system dynamics and due to judgmental errors both at managerial decision level controls comprising human-machine interfaces and at higher level controls like production control, which apart from human-machine interfaces, in all probability, may even include humans as part of the process to be controlled.

### ***5.2 Uncertainty at Informational Work System Level***

And then there are the multiple individual decision process stages, which are also impacted by 5“C”s. This has implications for the entire supply chain.

### **5.2.1 At Plant Operations Stage**

Briefly, the business process *IS* view is a complex *IS*. This complexity introduces at the operational level, hitherto unknown uncertainties in the form of complex error mechanisms coming from system development and implementation life cycle phases and that, too, coming with delay.

### **5.2.2 At Plant Operations Stage and at Controls Stages**

At the stages of operational level and at control levels, it introduces uncertainties that arise due to failures of embedded systems; and, at all business *IS* stages, uncertainties due to emphasis on system integration maximization. The later category of uncertainty is on account of resulting system interfaces exposing innermost system modules to uncertainties due to external system environmental factors and vice versa.

### **5.2.3 At Decision Process Stages Endogenous to Decision Situation – From Long Term Goal Set Obtaining Multiple Criteria to Selection of Flexible Information Decision**

Further, particularly at the decision process stages endogenous to decision situation, the complexity introduces uncertainties due to information processing under: multiple goals (implicit goals included); multiple factors & multiple criteria; a large number of interdependent information variables, varying with time and not completely and correctly observable; and system dynamics not well understood, which is because reality is not passive but – to some extent - active.

All these uncertainties result in information errors leading to loss of Information Integrity at each of the decision stages of the business process *IS* view and, therefore, in entire *IS* and in information therefrom.

## **6. BUSINESS PROCESS *IS* VIEW– CONTINUOUS INDIVIDUAL INFORMATION ORIGINATING AND PROCESSING SITUATION UNDER UNCERTAINTY**

For the purpose of present investigation, this draws attention to two relevant modeling aspects of the informational view of the business process.

### ***6.1 Multistage Decision Process based IS not a Static but Dynamic View***

Firstly, as observed in section (3.3), the multistage decision process based view of business process *IS* has stages of: from multiple criteria, *recognizing* operable goal; from operable goal statement, *defining* planning & design constrain and opportunity spaces; and from ‘many factor’ information variables characterizing problem complexity, *culling out* useful (relevant) information variables. What is required here is to *cull out* – not necessarily physically, but mathematically – and study facts (i.e., data and information variables) that are identified as relevant to the identified goal (Usefulness factor). As argued in Section (3.4), this requires treating business *IS* as an individual situation involving information origination.

An important question is: what if, in above *IS*, the “goal” leading to usefulness factor with reference to information originated, though given, *continuously* needs adjustment due to constantly changing environment? Such very well could be the situation in service sector – e.g., medical treatment of an adult patient. Or it could be the goal is not known or is out of date or is by itself complex. All these are conditions to be observed in the real world problem solving. Going further, particularly from a conservative point of analysis, consider a tightly coupled, high-reliability engineering system, which is large, semantically complex, time-pressured, and of high consequence. In the wake of unclear goal statement (implicit goals inclusive), this system description is also to be observed to run a risk, in the fashion of an open system, of taking a life of it’s own.

In such case tasks of culling out relevant facts (Usefulness factor with reference to data and information variables) and of defining their interrelationships under subsequent decision stages of the business process *IS* view cannot be treated as static ones. That is these tasks cannot be determined uniquely and exogenously as in case of closed systems. Rather, in the presence of 5”C”s, they acquire dynamic – open and *endogenous* to the decision making situation in that - character. This requires the data and information variables be *continuously* originated and processed. This reality leads to model information processing under the business process *IS* view as a *continuous individual information originating and processing situation in the presence of uncertainty*, so as to account for demands of continuously determined specific goal based individual situation in a complex and changing environment.

## **6.2 Business Process IS Errors are Informational Errors and Information Origination A Costly Activity**

Secondly, and perhaps most importantly, from Section (5) and specifically from Sub-section (5.2.3), it emerges that uncertainty due to ever changing system environmental factors of 5”C”s introduces errors at all decision stages of information origination and processing. In other words, due to 5”C”s, errors occur at stages of information origination, storage, retrieval, validation, processing, communication and distribution, *use*, and discard or storage for future *use*.

Traditionally, *IS* is designed with a static view of its environment and assuming that information is function only of “source” and at most of “process.” That is information processed is *not* understood to be function of its “recipient”, i.e., information user (customer) who may be a human or a software (includes instruction) or a machine. In management science it is useful to model information system (*IS*) as a network of horizontal structures of customer-supplier models. These customer-supplier models, which account for all internal and external customers, are considered at strategy, control and operations levels and in actual system may be combined in a variety of ways. They may be repeated, paralleled, and interrelated. Customer from one set of elements may become supplier to another set. Accordingly, these customer-supplier models and their networks represent information system, its sub-systems and their components, namely, hardware, software, procedures, regulations, communications, policies, financial constraints, etc.

The problem is, even with the reality of the ever present 5”C”s, these customer-supplier models at no stage anticipate errors due to ‘distortion’ in the once validated requirements of the internal and external *IS* customers. The assumption made is that information is correct, once validated. This makes their main concern – and therefore the main *IS* concern - to be *only* that information technology accesses, communicates, processes and distributes the *already* generated information exogenous to the *IS* decision situation. In other words, the *IS* model emphasis is not on “correctness” *but* rather on “exactness” aspect of customer information requirements. With information technology costs ever decreasing, the information processing for decision-making is, therefore, taken as a *costless* activity.

But the real world and its requirements are at striking variance with this assumption. All time implicit assumption of requirements (read information) as “correct” or “trustworthy”, once validated, leads to construction of data base management systems that remain concerned *only* with internal objects of *IS*. This is a closed system view of an *IS* and not acceptable. Information user is an open system and information processed is function of “information user”, also. This brings into play implications of ever present system environmental factors of 5”C”s on *IS*, its sub-systems, and their components. Such *IS* comprises individual decision process stages that are characterized by activities of information “origination”, which are *costly*. As a result, this implies the centrality of (a) informational errors (i.e., incorrect production of information) in a business situation setting (and not business errors as one tends to define) and (b) of their costs arising from uncertainty beset information origination and processing.

## 7. EFFICIENT AND ECONOMIC PROCESSING OF INFORMATION – INFORMATION AS BOTTLENECK AND INFORMATION INTEGRITY THE CRITICAL FACTOR

For competitive business advantage, as observed in Section (2.4), this *IS*, beset with uncertainty and errors, must process information efficiently and economically. This calls for controlling Information Integrity (I\*I) of *IS* and of information there from.

### 7.1 Bottleneck character of Information

To elaborate, businesses operate in market-based social systems, which from economic angle are governed by market price mechanism. For effective functioning, market price mechanism requires a common denominator to work with: money. Information about it is expressed in a form that makes it particularly *useful* and *usable* in the context of a market pricing mechanism. In other words, the success of information “I” with reference to money as a medium of exchange and unit of measurement (measure) *is* in the fact that it (*information “I”*) *minimizes transaction or comparing and selection costs*. This efficiency and economy in the market information gathering (originating) and processing systems that then provides the engine for both internal and external performance measuring systems based on concerned information “I” getting emphasized as the market grows in complexity. This explains the bottleneck character of the information used for the purpose.

### 7.2 Information Integrity a Critical Factor for Efficient and Economic Processing of Information – An Economic Angle

There is a little more to the analysis being pursued. Increased importance of monetary information ( $I_1$ ) is also accompanied by useful role this information plays in allocating resources at two levels. Specifically, efficient resource allocations in any product or factor market *require* competition. Competition requires, as necessary conditions at least two things: (a) information ( $I_2$ ) about market imbalances indicating business opportunities, and (b) information ( $I_3$ ) on working mechanisms (knowledge capital) for exploiting opportunities. Information (I) then is given by ( $I=I_1+I_2+I_3$ ). For efficient processing of information, there *has* to be a trade off between:

- Costs associated with originating and processing of information denoted by [ $COST_{ORIG\&PROC} \{I\}$ ], and
- Loss due to incorrect information, i.e., loss due to I\*I Risk [ $LOSS_{RISK} (I*I)$ ].

That is *that* optimum (ORIG&PROC) – denoted by (ORIG&PROC)\* - will prevail for which [ $COST_{ORIG\&PROC} (I)$ ] and [ $LOSS_{RISK} (I*I)$ ], i.e.,  $\{[COST_{ORIG\&PROC} (I)] + [LOSS_{RISK} (I*I)]\}$ , are minimum. Or, said in words, *IS* which, for a certain kind of information *origination*, storage, validation, processing, distribution and discard or storage for future use (*IDILC* model), is able to arrange them (costs) at the lower level *will* tend to prevail. From this it follows that, *to compete successfully*, (a) information “ $I_1$ ” – the aggregate information; (b) information “ $I_2$ ” on market opportunities, and (c) information “ $I_3$ ” constituting knowledge of working mechanisms for resource allocation must have *integrity*. This requires that each detail in each of these information statements (the aggregate or the measure, the opportunities, and the knowledge capital), and *not only* the bottom line statements, must have *integrity*; as it is only through ensuing of optimal integrity that it is possible to achieve efficient and economic processing of aggregate information in the business *IS* view described above.

This analysis has a far-reaching consequence. Firstly, it reinforces that information *origination* is a costly activity. Secondly, it shows that it is *only* by controlling I\*I that *IS* can be made to prevail. But, here, *IS* represents the business enterprise supply chain and, hence, what in fact prevails is the business enterprise supply chain. This indeed is the condition for the competitive survival of business enterprise system,

which is achieved through controlling I\*I. This establishes criticality of controlling I\*I for business competitive advantage.

### ***7.3 Minimization of Entropy in IS and Minimization of increase in Entropy in System Environment for Efficient and Economic Processing of Information – A Thermodynamic Angle***

There is another important angle to this searchlight. It is a thermodynamic angle. An *IS* processes much needed information for decision-making and ensures purposeful influence toward a continuously determined operable goal. Requirement in doing so, is to maximize the quantum and duration of the informational work (longevity of *IS*). Understandably then, the *IS* must do the following:

- (a) Minimize entropy in its systems so as to maximize organization in the system. Here ‘organization’ would mean order and end-directedness; and
- (b) Minimize increase in the entropy in the system environment, i.e., in the real world from which *IS* extracts and processes information. This is to minimize disorder in the system environment, which *IS* consumes for the purpose of sustaining and improving *organization* of the system.

Minimization of entropy in the system means minimization of entropy in the internal environment of the system. This perforce means minimization of missing information in the internal system environment. Similarly, minimization of increase in the entropy in the system environment would mean minimization of increase in the missing information in the system environment. *IS* consumes from system environment for information origination and processing for control implementation. Further, to increase *organization* in the system, the *IS* releases the energy through energy conversion and through elimination as *IS* by-product wastes (see Sub-section (2.4)). This leads to increase in entropy in the system environment.

### ***7.4 What must be controlled then? - Information Integrity***

Thermodynamic angle is the pervasive universal law of nature, which anything that is physical *must* obey. It observes that:

- (i) Missing information in internal system environment in respect of processes of information origination and processing of unstructured and aperiodic information and not only of processing structured and periodic information, and
  - (ii) Increase in missing information in the external environment of these processes,
- are* to be kept at minimum. If not, the former leads to higher entropy in the system and the latter to higher increase in the entropy of the system environment.

If above minimization is achieved at each of the decision process stages (see Sub-section (3.3)) and, therefore, for business process *IS* view, then the physical reality of *IS* can postpone the inevitable *heat death* for the time, which is the most fundamental design requirement governing physical work systems for the competitive advantage. For *IS*, this means requirements:

- (a) To maintain a desired (optimum) Information Integrity of *IS* and of information there from, and
- (b) To minimize increase in loss of Information Integrity in system environment.

And the economic angle - the market angle – so essential in sustaining the economic activities in societies suggests that to compete successfully *IS* is required to process information efficiently and economically, and for this to achieve information processed *must* be with Integrity.

We thus have two powerful scientific angles. One pertains *order* of nature (entropy), which results in disorder for information, and the other *order and organization* of information (integrity), which is disorder for nature. When put together they axiomatically propose that maintaining and improving Information Integrity is a necessary requirement to ensure sustained productive, efficient, and economic

functioning of the open system that the *IS* is; thereby establishing that for competitive advantage it is the Information Integrity that must be controlled.

## **8. INFORMATION POLLUTION – IMPLICATION OF LOSS OF INFORMATION INTEGRITY: A RESEARCH AND KNOWLEDGE DEVELOPMENT DIRECTION**

In the wake of informational errors leading to loss of Information integrity, Sub-section (2.4) recognizes higher increase in the entropy of the system environment through elimination as system by-product wastes (see Sub-section (2.4)). This is the issue of increased environmental pollution as a result of information pollution caused by informational errors in the open system that the *IS* is. In the wake of the ever present system environmental factors of 5“C”s, what is significant is that this information pollution takes the form of a continual phenomenon.

With reference to the research field of Environmental Pollution & Engineering, this modeling makes a case for a useful research direction of “Information Pollution”. For a generic description, information pollution can be seen to result from loss of: information origination integrity, storage integrity, retrieval integrity, validation integrity, processing integrity, communication and distribution integrity, *use* integrity, and of information discard or storage for the future *use* integrity. In a broad manner, for *IS*, it could be presented as an issue of information pollution from loss of content integrity, process integrity, and system integrity. At a more functional level (and to an extent in violation of Section (6.2)), if informational errors in a design setting are described as design errors and so are informational errors in other settings, then information pollution could be seen to result from loss of: design integrity, development integrity, deployment integrity, data and data product integrity, and detection (audit) integrity covering the entire *ISDLC* model. Similarly, one can also recognize the reality of information pollution due to loss of system implementation integrity, maintenance integrity, etc.

In summary, Information Integrity is a systems concept and recognition of eventuality of informational errors in different settings presents information pollution as a systems problem. As technology facilitates digital data as medium of information flow across the *SDILC* model and as informational work system characterized by system integration application is put to work, due to 5“C”s, every information movement in the *SDILC* information environment runs a risk of loss of integrity, thereby in the face of impending issue of information pollution harming the system capability to process information, material and energy efficiently and economically.

System information environment, infested with risk of integrity loss at every information movement, is like a polluted water lake or like an atmosphere over a non - environment - friendly industrial township. Every drop of water, which human drinks, or every air particle that a human breathes must be pure, or otherwise over a time period a life, if not the entire species, can be under danger. Similarly, every information movement must deliver information with integrity. Otherwise, given the business complexity and given difficulty to foresee the results of quite minor information pollution, over a period of time the business process *IS* is in danger of processing low integrity information decisions. This will put business system out of control leading to losing customers, if not an entire market, and therefore loss of competitive advantage.

Requirement therefore is:

- (a) To understand information processing models for following decision stages (Sub-section (3.3)):

- a. *Obtaining* ‘many factors’ & ‘multiple criteria’ characterizing business task (problem) complexity;
  - b. *Recognizing* (deciding) on an operable goal;
  - c. *Defining* planning & design constrain and opportunity spaces;
  - d. *Culling out* useful (relevant) information variables;
  - e. *Recognizing* relationships (interdependencies) between culled out information variables (environmental implications);
  - f. *Developing* state transition models defining dynamic behavior of culled out state (information) variables; and
  - g. Undertaking customized planning & design for *generating* alternatives for evaluation and final selection of flexible information decision for control implementation at physical work system.
- (b) To research uncertainties present at each of these stages in the form of issues of information distortion and noise (Section (5)),
- (c) To develop Theory of Information Error and its Analysis; and
- (d) To study Information Pollution (i.e., *IS* by-product wastes) due to loss in regard to “correctness” aspect of information due to:
- Information Distortion — Leading to loss of Accuracy (Non-conformance with established standard and acceptable tolerance), of Consistency (Non-conformance with established standards and acceptable tolerances in respect of spatial and temporal constraints), and of Reliability in the form of incorrectness due to incompleteness; and
  - Information Noise — Leading to loss of Reliability in the form of incompleteness due to inexactness of information. Specifically this will account for implications of incorrect system and component operations due to:
    - Mechanistic failure, service disruptions, failure of hardware, etc, and
    - Failure of system equipment, which is controlled directly by computer and, therefore, is dependent upon the correct functioning of the computer hardware and its software.

Specifically, these research initiatives should lead to techniques and technologies to ensure integrity of information at each information movement node (these cover nodes where data or information changes form as also nodes characterized by transmission without change of form where loss of integrity could be due to difficulty in deciding the purpose of transfer), as otherwise business runs the risk of working with low integrity (i.e. polluted) information environment. The universe of information movement nodes as defined here then offers a construct of the Information Integrity Space.

### ***8.1 Information Integrity Space - Business Opportunity to develop Information Integrity Technology for controlling Information Pollution***

Seen from another angle, the Information Integrity Space generates integrity improvement business opportunities consistent with integrity requirements of parts or entirety of business process *IS* view. Based on current perception of business requirements, in the first instance following Information Integrity Space dimensions are identified for developing integrity technologies:

- Prevention, monitoring, detection, verification, and correction of information errors.
- Security, audit, and control.
- Protection against corruption of information due accidental failures or deliberate fraud.
- Data scrubbing and cleansing in the creation of data warehouses.
- Design, development, operation, use, and maintenance of information systems.
- Conversion of existing systems due to mergers, acquisitions, and consolidations.

Modification of existing systems to accommodate changes such as Y2K, new legislation, or new technology

- Information Integrity requirements of specific industries such as banking, finance, telecommunications, engineering, transportation, defense, etc.
- Information Integrity requirements of various data forms such as voice data, video, etc.
- Information Integrity Technologies in the context of information explosion, application integration, and zero latency enterprises.

Every enterprise performs a variety of activities in respect of above and spends a significant amount of resources to ensure the integrity of information it delivers to its manufacturing/production units and processes, customers, partners, suppliers, and shareholders. Specifically, for this purpose, most enterprises have systems and procedures to verify the integrity of information they receive from external sources and of information they produce and use internally. These systems and procedures are usually custom-developed, and unique to each enterprise. Further, many of the Information Integrity related activities are performed in all parts of the enterprise and are, therefore, often duplicated.

In other words, Information Integrity solutions as above are being developed and applied in isolated, unique and ineffective ways within most enterprises, even though the underlying requirements are the same, and could be addressed more effectively and consistently through an enterprise-wide approach. Consequently, many enterprises are paying an inordinately high price for the level of Information Integrity they are able to achieve. For example, even in a country such as India, where computerization is still at a nascent stage, one estimate suggests that India spent around USD 412 millions on Y2K related businesses; although the positive side is that one third from it was earned by IT industry fixing the problem.

Studies indicate 75% businesses experience significant problems due to faulty data and that as many as 67% of businesses do not feel confident about the quality of their company's data. Though there are no readily available studies about the total cost of Information Integrity related activities within an enterprise, many executives and industry experts suggest that such costs could be in the range of 1- 5% of revenue. Yet another study estimates poor data quality costs business \$611 billion per year in the United States alone. Effective, enterprise-wide, and industry-wide approaches to Information Integrity using standards, best practices, and innovative Information Integrity Technologies while on the one hand can significantly reduce this cost, on the other hand they will open up the Information Integrity Space for new business opportunities in an unfolding Information Integrity Industry.

## ***8.2 A Case for Information Integrity Knowledge, Science, Technology and Industry Development for controlling Information Pollution***

Not since the industrial revolution has technological advancement held as much promise and uncertainty as today, the beginning of the 21<sup>st</sup> century. Promise is due to shift in economies from energy- based technologies to data - driven technologies, which are empowering business systems with a global market reach for competitive advantage based on seemingly endless access and processing of information. Uncertainty is because, due to the uncertainty in system environment, the information originated and processed is fraught with risks of errors in the business information system. This brings in the issue of information pollution and hence of loss of integrity of information system and of information processed by it which are key to the strategic business advantage in the presence of uncertainty.

Emanating from the reality of information pollution, Information Integrity emerges as a pervasive universal issue, which impacts business, government, and society in many ways in this electronic information age. Today, the knowledge and understanding about Information Integrity is rudimentary,

fragmented, and insufficient. On the one hand it has philosophical, sociological and economic aspects which are yet to be conceptualized, and on the other hand it has technical, technological and industry related requirements which only now have begun surfacing. Seen from this point, Information Integrity has potential for becoming a new discipline, a new science, even a new industry, very much like the environmental science and industry, which emerged as a result of society's concerns about the quality of air, water, and the earth.

The emergence of the new Information Integrity science, technology and industry will require the crafting and communication of clear, compelling, and consistent messages about the pervasive and key nature of Information Integrity, and the formation of a coalition of academia, thought leaders, professionals, practitioners, and organizations interested in Information Integrity.

## REFERENCES

- [1] Beniger J. R., *The Control Revolution – Technological and Economic Origins of the Information Society*, Harvard University Press, Cambridge, Massachusetts, USA, 1986.
- [2] Mandke V. V. and Nayar M. K., Information Integrity - A Structure for its Definition. *Proceedings of the 1997 MIT Conference on Information Quality* edited by Diane M. Strong and Beverly K. Kahn, Cambridge, Massachusetts, USA, October 25-26, 1997.
- [3] Mandke V. V. and Nayar M. K., Information Integrity Imperative for Competitive Advantage in Business Environment characterized by Uncertainty, *Proceedings of 'Information Technology for Business Management' Conference* edited by Prof. Renchu Gan, 16th World Computer Congress, August 21-25, 2000, Beijing, China, 115-125.
- [4] Mandke V. V., Nayar M. K. and Malik Kamna, Information Envelope and its Information Integrity Implications: For a complex, changing environment, modeling a generic business process as an integral to a closed loop information and control system characterized by uncertainty. *Proceedings of the 2001 MIT Conference on Information Quality* edited by Elizabeth M. Pierce and Raissa Katz-Haas, Cambridge, Massachusetts, USA, November 3-4, 2001.

0-0-0-0-0