IIRPB27F2KIND
13.1.07

INFORMATION INTEGRITY RESEARCH PROJECTS BANK

-Draft-
(Document not complete; Not for circulation)

by

Prof. Vijay V. Mandke
Research Leader
Center for Information Integrity Research
B-64 (FF), Gulmohar Park,
New Delhi-110049.
Email: vijaymandke@satyam.net.in

CENTER FOR INFORMATION INTEGRITY RESEARCH
New Delhi -110049
October 2001

-DRAFT-
(Document not complete; Not for circulation)

Information Integrity Research Bank Index

1. Studies in Socio, Economic and Technological implications of integrity (or lack of it) in IS.
2. Critical study of Data/Information Models.
3. Study of Information System (IS) Models
4. Study of Uncertainty in Business Environment
   (Note: This Project may be seen together with Project Sl. No.(18))
5. Study of Business System Environmental Factor of Complexity contributing to Uncertainty.
6. Study of Business System Environmental Factor of Change contributing to Uncertainty.
7. Study of Business System Environmental Factor of Communication contributing to Uncertainty
8. Study of Business System Environmental Factor of Conversion contributing to Uncertainty
9. Study of Business System Environmental Factor of Corruption contributing to Uncertainty
10. Study of Errors in Business IS View due to Uncertainty in Business System Environment
11. Study of enterprise- wide errors associated with generic organizational processes with particular reference to errors in an enterprise information system.
12. Industry specific application of study in Project at Sl. No. (11) above. Various industries that could be considered are : Manufacturing, Telecommunications, Banking, Aviation, Medicine, etc.
13. For different industry (application) sectors, study of errors stemming from well-established devices.
14. Study of theory of Error and Accident.
15. Adopt the Theory of Error and Accident (studied under Project at Sl. No. (14)) to study *Theory of Errors and Consequences of Loss of Integrity therefrom* in Computerized Information System (*IS*) constituting the business process/activity and in Information processed by the *IS*.
16. Based on the Project at Sl. No. (15), for different application sectors, namely, manufacturing, banking, telecommunications, medical sector, etc., develop understanding of errors and consequences of loss of information integrity therefrom in *IS* constituting the respective business process/activity.
17. Study of Latent Errors in Large, Complex and Dynamic Application Domains.
18. "All Business Processes whatever else they do, make automobiles, sell real estate, run hotels, or whatever, they all process information and in that they all are information systems(*IS*)." Explain.
    (Note: This Project may be seen together with Project Sl. No. (4).)

19. On the nature of 'Business *IS* View'
20. Develop a process model for Decision-Making in a work environment that is a combination of extreme dynamism, intense time pressure, high complexity, perpetual uncertainty, and palpable risk.
21. For human-machine systems, study the methodology (a) to detect whether or not an error leads to incorrect action (misadministration) and what, if any, type of misadministration occurs, and (b) to develop approaches to eliminate possibilities of error occurrence.
22. Application of study in Project at Sl. No. (21) above to study errors in human-machine systems under Computerized Information Systems (See this Project along with Project at Sl. No. (11)).
23. Industry specific applications of study in Project at Sl. No. (22) above (See this Project along with Projects at Sl. Nos. (11) and (12)) .
24. Application of study in Project at Sl. No. (21) above to study errors in human-machine systems at different phases of Software Development and Implementation Life Cycle.
25. Modeling Business Process as integral to an information and control system for the Environment characterized by Uncertainty –A Controls interpretation.
26. Discuss the assertion "Information Integrity is a pervasive phenomenon."
27. Develop a view of Information Pollution.
28. What are the various types of Information Pollution?
29. Discuss various Information Pollution sources in the context of an organization.
30. What can be considered as the Symptoms of Information Pollution in an organization? What are the methodologies for evaluating Information Pollution?
31. Business process *IS* view (see Figure (18.3)) comprises a series and parallel network of core *IS* model given in Figure (3.13). Specifically, each of the stages of this core *IS* model carry out information movement in the form of transformation and transfer of data and information, so as to facilitate the business *IS* view to deliver information in the form of information decision for control implementation. Further every data transformation as also data transfer involves a decision. Thus one workable reality model of informational movement (constituting the informational work) which is at the bottom of a business process *IS* view and constitutes its atomic description is in terms of a *decision making process model* (see Figure (19.1)). With data driven technologies keyed to the flow of information across the supply chain and on the Net, the Framework for Project at Sl. No. (-) observes that these decision making processes representing the information flow are characterized by dynamic decision making situations, putting every stage of the *IS* development Life Cycle and their users (external as also internal) at the sharp end of operation characterized by dynamic and complex problem solving situations, and, therefore, are prone to errors that accompany complex problem solving. Study the nature of these dynamic, complex problem solving situations and of errors therein with particular reference to issues of user behavior and psychological factors influencing them (Note: This Project Framework may be read along with the Framework for Project at Sl. No.(17)).
32. Within the Framework of the Project at Sl. No. (31), identify the primary characteristics of the Complex Problem Solving (this Project should be seen along with the Projects at Sl. Nos.(4) – (10), (17) and (20)).

33. As observed in the Frameworks for Projects at Sl. Nos. (-,-,32), system complexity, non- observability of system states (also termed as intransparence), internal dynamics (change), and incomplete or incorrect understanding of the system reality model, are basic to all dynamic decision making situations which are the hallmark of demands of operating at the sharp end of operation, so critical and universal in the business process *IS* view in the wake of incoming of commerce tied to data driven technologies and the Net. This calls for planned approach to solving complex problems. Develop a workable schema for such entire problem solving process so as to ensure correct information decision for appropriate action (Note: This Project Framework may be read along with the Framework for Project at Sl. No. 20).
34. Business process *IS* view is a system. A system is defined by its input, process, output, system environment, and system objective or goal. As discussed in the Framework for Project at Sl. No. (--), it is the property of end-directedness, i. e., objective or goal that distinguishes an open system from a closed system, and it is the open system that is characterized by the requirement of information processing between its sub-systems and components. In other words, when solving a dynamic decision making problem, one is actually solving a complex *IS* problem which is a micro-ecology of a decision making process model having a property of end-directedness, i. e., system objective or goal. As a result, steps of identifying the requirements of goal setting and of setting goal acquire very important position in developing a workable mechanism for complex problem solving. Discuss the issue of requirements and setting of "system objective" within this framework of organization for complex problem solving.
35. Develop a reality model of how problem solvers in real situations deal with goals. That is to say investigate/discuss if real life problem solvers like planners and implementers as they work on situations needing problem solving recognize requirements of goal setting and apply
goal setting methodologies as described in Framework for Project at Sl. No. (34).
36. A system means more than simply acknowledging the existence of many variables. It means recognizing the different ways the variables can affect each other and themselves, i. e., the interrelationships between the variables. Discuss different categories in which these interrelationships can be grouped.
37. **Assignment**: Based on Framework for Project at Sl. No. (36), study a real world Situation from your own environment - professional, home, economic, social, etc. - so as to bring out the complexity present in such situations and in the information models representing them.
38. Discuss the reasons for errors in developing information models for real world, Complex situations.
39. Discuss the limitations of "reductive hypothesis" as a methodology for information model building while solving complex, real world problems.
40. **An Assignment**: As discussed in the Framework for Project at Sl. No. (-), *Complexity* is one of the system environmental factors that introduces errors in the business leading to loss of INTEGRITY in *IS* and in information therefrom. Specifically, *Complexity* stands for the existence of many interdependent variables. Study if a "COMPLEXITY QUOTIENT" can be developed for a system.
41. All enterprises/organizations-man made systems, business firms, cities, or economic

systems- are liable to be diverted from their chosen paths of growth or stable operation by the shocks which fall on them from the outside world, or by pressures generated within the system. The 'shock' may be an opportunity, such as a firm attempting to react to an enlarged market, or it may be unpleasant, as when changes in international market prices affect the balance of trade. The shocks may be unexpected, foreseen or self-inflicted, but the essential problem for decision makers is to decide on the information decision so as to control the enterprise or organization so as to take advantage of favorable opportunities while defending it (the system) against unpleasant upsets (noises). All these are dynamic systems behavioral examples which often also display elements of systems delays.

Study systems engineering techniques for modeling dynamical systems as described here. Further study how these techniques can be applied for studying *IS* modeled inductively.

42. **Assignment**: The network age is upon us. Computers are no longer used stand-alone. An office workstation is usually connected to remote printers, file servers, database servers and world-wide electronic mail facilities. A point of sale terminal will connect to in-store stock control systems, banking systems for credit validation and possibly even a remote consumer analysis organization. A travel agent can access the on-line reservation systems of dozens of airlines, hotels or car rental companies. Even the simple home computer may have modern access to electronic mail, bulletin boards or remote computing facilities. Proceeding further, a typical factory or industrial plant will now have dozens of computers connected by networks cooperating to control car assembly plants, oil refineries, steel mills, food manufacturing plants, or power generation.

    Study strategic, control and operational information requirements for managing a network or distributed system as above. Also discuss information architecture when management is implemented as a distributed system.
43. **Assignment**: Each information system has predefined procedures. Each procedure provides an accurate and timely course of action. Procedures perform desired tasks of recording, organizing, and classifying data; identifying alternatives; and maintaining adequate records necessary to achieve specific objectives. Describe activities, procedures, and tasks corresponding to an Accounting System of an enterprise.
44. **Assignment**: An enterprise/industry has an open system for every business activity. Complex business systems are generally divided in to sub-systems. A sub-system is a smaller functional unit, serving the overall configuration in order to ensure its success. For example, registration is a subsystem of a college's student service system.

    Develop activities identified in the course of Project at Sl. No. (43) as open systems.

    Note: Open systems are interdependent with their environment.
45. **Assignment**: Discuss dimensions of Information Requirements. Elaborate by giving examples.
46. Problem solvers live and operate in a four-dimensional system. In addition to the

three dimensions of space, this system includes the fourth dimension of *time*, which moves in one direction, and that direction is toward the future. Discuss the difficulties problem solvers have in studying time sequences and recognizing temporal patterns.

47. Towards the goal of complex problem solving, Framework for Project at Sl. No. 33 has discussed steps in the organization of complex information decision and action. These steps are: (a)Formulation of Goals (For requirements of this step and for what difficulties problem solvers have in carrying out this step, see Frameworks for Projects at Sl. Nos. (34-35)), (b)Formulation of and Gathering of Information (For requirements of this step and for what difficulties problem solvers have in carrying out this step, see Frameworks for Projects at Sl. Nos. (36-39), (c)Prediction and Extrapolation (For requirements of this step and for what difficulties problem solvers have in carrying out this step, see Framework for Project at Sl. Nos. 46), (d)Planning of Actions, Decision Making, Execution of Actions, and (f)Review of Effects of Actions and Revision of Strategy.

    Within above framework, for complex problem solving, discuss the methodology for: (d)Planning.
48. Based on the Frameworks for Projects at Sl. Nos. (-47), develop an overview of inadequacies of human thought in dealing with complex systems.
49. Study the ways to respond to complex problem solving situations?
50. **Assignment**: Framework for Project at Sl. No. (46 …. Check the reference number!) has described the difficulties in developing time sequences, i. e., the system dynamics, which in turn lead to errors in information decision in complex system. With reference to this discussion, through example(s) show how errors in information forecasting take place if numbers are solely interpreted on the basis of their size. Further explain why in order to understand what numbers mean, it is necessary to take into account the process that produced them. Also explain what difficulties humans have in carrying out this requirement?
51. Computerized Information Systems (*CIS*) constitute the technology for organization of information decision for complex systems. Discuss errors in *CIS*.

    (Note: This project may be developed with particular reference to small and large engineering systems, though the investigation will also be applicable to soft systems.)
52. **Assignment**: Based on discussion on faults resulting in errors in *CIS* (see Framework for Project at Sl. No. (51)), investigate how errors are created, i. e., how errors take place?
53. **Assignment**: Traditionally, classification of faults in *CIS* is discussed at the “logic” level. Extend this approach to study classification of undesirable events in *CIS* to other universes; namely, “physical” and “information” level.
54. **Assignment**: Fault-tolerance or fault-tolerant computing, or more generally, the ability to produce correct computational results even in the presence of faults, errors and unexpected situations has been an important concern since the beginning of the digital computer age.

    Study the development of fault-tolerant computing.

55. **Assignment**: From the Framework for Project at Sl. No. (54), it can be seen that the literature on "Reliability of Computer Systems" dwells on the issue of fault-tolerant techniques in the quest of reliable, i. e., failure-free or failure-secure or failure-tolerant systems. Within this framework, survey the literature to understand approaches pursued for increased reliability or dependability of computer systems.
56. As elaborately discussed in Frameworks for different Projects, Computerized Information Systems (*CIS*) contain errors. This results in loss of integrity in information systems and in information therefrom. For a *CIS* model as in Figure (3.13), study errors in an operating information system and in information therefrom.
57. Based on the Framework for Project at Sl. No. (56), study loss of integrity in information system and in information therefrom due to errors in operating *CIS*.
58. Based on the Frameworks for Projects at Sl. Nos. (56-57), develop Information Integrity Attributes.
59. Based on the Frameworks for Projects at Sl. Nos. (56-58), discuss the inadequacy of application controls to ensure Information Integrity.
60. **Assignment**: Based on the Frameworks for Projects at Sl. Nos. (56-59), discuss what strategy is then required for Information Integrity Improvement in *CIS*.
61. Develop a Information Integrity Research Survey Report.
62. Develop the concept of Information Integrity System.
63. **Assignment**: Based on the perception of information requirements as given in the Framework for Project at Sl. No. (62), define the *IS* errors due to difficulties in obtaining correct and complete set of information requirements.
64. **Assignment**: Define *IS* Errors in data/information model < Entity, Attribute, Value>. Specifically, give a checklist for defining Entity, Attribute and Relationship whereby presence of error could be signified by the virtue of checklist not being met.
65. Study errors at various stages in *IS*; namely, Data Origin, Data Processing and Information Use Stages.
66. Identify *IS* errors due to Design Errors (Design reviews), Development Errors (Testing), Deployment errors (Security & Control), Data Errors, and Detection Errors (Audit):
67. Analyze *IS* Errors.
68. **Assignment**: *IS* errors lead to loss of integrity in *IS* and in information therefrom. This calls for information integrity improvement technology. Specifically, **(a)**Fault/Error avoidance techniques aiming at preventing faults/errors from entering the system during the design stage, **(b)**Fault/Error removal techniques which aim at finding faults/errors within a system before it enters service and which are implemented at the stage of hardware and software testing, **(c)**Fault/Error detection techniques used during service to detect faults/errors within the operational system so that errors and their implications (adverse effects/failures) may be minimized, **(d)**Fault/Error tolerance system designed to allow the system to operate correctly in the presence of faults/errors, and **(e)**on-line Fault/Error detection/estimation/prediction
and control or correction (i.e., on-line feedback control mechanisms characterized by on-line learning and control/correction) mechanisms implemented during operations stage (i.e., within the operational system), are all examples of integrity technologies.

Survey various application areas of Information Integrity Systems (I*I S) using I*I Technologies as above.

69. Develop I*I Requirements.
70. Describe Information Use hazard Analysis Techniques.
71. Develop an Information Error and Effects Analysis Table.
72. HAZOP for study of errors in IS and in Information.
73. Discuss Hazard analysis within the *IS* development life cycle.
74. Study of Hardware Faults.
75. Study of Hardware Design Faults.
76. **Assignment**: On some aspects of Specification faults.
77. Study of Software Faults.
78. Discuss fault coverage approach as a measure to assess success of fault avoidance/removal/detection/correction/tolerance (that is I*I) technologies.
79. On developing the Information Integrity Space (I*IS).
80. Economic impact of Information Integrity.
81. Methodology for quantifying economic impact of Information Integrity
82. Develop *IS* Models for Integrity Analysis.
83. **Assignment**: Study the Framework for Achieving Information Integrity (I*I).
84. **Assignment**: Develop I*I Architecture.
85. Describe the subject area of Integrity Engineering.
86. Assignment: Define the term Design Integrity.
87. **Assignmen**t: Define the term Implementation Integrity.
88. **Assignments**: For different application areas, study following topics:
    (Note: Each topic could give rise to different assignment.)

    - IS Models
    - Information Flow Models
    - Linear Flow Models
    - Developing quantifiers for Integrity attributes
    - Methodologies for establishing industry specific standards for attributes, integrity profile and cumulative integrity indices
    - Developing Integrity Analysis Methodologies
    - Establishment of user data rules list for error detection
    - Software development in respect of:
    - Program for integrity Attribute indices, Integrity Profile and Cumulative Integrity indices
        - Sampling program, Sampled Records Audit Trail Program
    - Programs for error detection/estimation/prediction
    - Development of Error detection Data Base
    - Presentation of reports on: Errors/ causes identified: their patterns and trends
    - Identification of Integrity improvement opportunities
    - Deciding and implementing Integrity Improvement Action
89. **Assignments**: Study Application Area and User environment Specific I*I Product Development so as to ensure:

- Standard data rule list
- Standard sampling software
- Standard sampled records Audit Trail software
- Standard sample file and error File
- Standard Statistical runs
- Standard reporting facility
- Standard filters, estimators, predictors
- Standard support data base to preserve integrity analysis results
- Standard mechanisms for deciding and implementing I*IA (Information Integrity Improvement Action)
- Standard for documentation and training

90. **Assignments**: Study User and Application specific I*I Technology Development, Testing and Evaluation.
91. **Assignments**: Conceive and Develop Research Projects to explore Information Integrity Architecture for long term development and implementation initiatives along the lines indicated below:

    - Structures of business processes
    - Structures of information Systems for business processes
    - Information Flow models for Information systems
    - Structural requirements
    - Hardware requirements
    - Software requirements
    - Information Integrity Architecture to meet integrity requirements
    - Protocols
    - Standards
    - Functional locations

92. Study of Enterprise-wide requirements of Information Integrity for managing Risks in Banking with special reference to Investment and Effects Portfolios (doctoral research topic as suggested by Professor Saha, Dean, National Institute of Bank Management, Pune).
93. Develop Information Integrity Attributes for:
    - Data (Data origin stage)
    - Process
    - Output, i.e., Information (Information for Use)
    - Network
    - People
    - System

    (Note: Should information use stage which could be a decision support stage not be added to above stages?)
94. Application of Systems Engineering Techniques (specifically of Modern Digital Control Systems Theory) to Fault-Tolerant System Design (1): Study method for detecting failures if feedback sensors in an operating control system. It may be noted that a failure in a sensor which provides a feedback signal in a control system can

cause serious deterioration in its performance. Early detection of the incipient failures will help in replacing the faulty sensor before serious trouble develops.

95. Application of Systems Engineering Techniques(specifically of Modern Digital Control Systems Theory) to Fault-Tolerant System Design (2): Study the design of failure sensitive detectors , or detection filters, which have the capability for real-time failure detection and identification in sensors and actuators. The proposed solution to the failure detection and identification problem could be based upon the reference model approach used in observers and Kalman filters.
96. Application of Systems Engineering Techniques (specifically of Modern Digital Control Systems Theory) to Fault-Tolerant System Design (3): Develop a method for failure diagnosis using Kalman filters. Specifically, the raw measurements from a dynamic plant may be processed using a Kalman filter to generate residuals. Failure detection and isolation may then be done by conducting Generalized Likelihood Ratio (GLR) tests on these residuals. Finally, the approach adapted may be based on the assumption that abrupt system changes occur infrequently i.e. the system models considered are a faithful representation except for a sporadic system failure.
97. Application of Systems Engineering Techniques (specifically of Modern Digital Control Systems Theory) to Fault-Tolerant System Design (4): Study fault diagnosis in dynamical systems using parameter identification methods.
98. **Assignment**: Planning for Distributed Data (1): Study properties inherent in certain data which lead naturally to distribution.
99. **Assignment**: Planning for Distributed Data (2): Study properties inherent in certain data which lead naturally to centralization.
100. **Assignment**: Planning for Distributed Data (3): Study reasons for multiple Distributed copies of data.
101. **Assignmen**t: Planning for Distributed Data (4): Study problems with distributed data.
102. **Assignment:** Planning for Distributed Data (5): Study forms in which distributed data can exist.
103. **Assignment:** From the six types of data distribution given in the Framework for Project at Sl. No. (101), in the first three data types, namely, *replicated*, *subset* and *reorganized* data, the same data may exist on two or more machines. In that case an important design question is: “Are the multiple copies of data synchronized?” In other words, when the value of the attribute is changed in one copy, is it immediately changed in the other copies?

    Study synchronous and non-synchronous data types.

104. **Assignment:** Study observed average failure rates for typical electronic components, for instruments, and for larger and more complex mechanical, electrical and electronic items of equipment.
105. **Assignment**: Study how factors of (a) Quality, (b) Temperature, (c) Environment, (d) Stress, (e) Complexity, and (f) Maturity of the manufacturing process influence failure rates of components and elements.
106. **Assignment**: Study failure rate models for electronic components.

107. **Assignment**: Study calculation of equipment failure rate from component failure rate data.
108. **Assignment**: Study software reliability models.
109. **Assignment**: Study methods for improving software reliability.
110. **Assignment**: Study factors influencing Human reliability.
However, it is far more difficult to predict software and human reliability.
111. **Assignment**: Study the literature for software reliability models for (a) Development Phase, (b) Validation Phase, (c) Operational Phase, and (d) Maintenance Phase.
112. On the problem of Development Lifecycle Model for Information Integrity System.
113. On the problem of fault-tolerant control system design
114. **Assignment**: Survey Dependability measures.
115. For a core *IS* model as in Figure (---), study data exposure groups and data exposure control points as discussed by Fisher.
116. With the objective of analyzing errors in *IS* and their integrity implications on information, study the totality of the Information 'envelop' (i. e., the total Information base) that can be considered to describe information processed under an open, complex system.
117. Within the Framework for Project at Sl. No. (116), develop a systems representation of Information envelop and its processing for a complex and changing business environment, and delineate emerging Information Integrity requirements.
118. On Information Integrity Imperative – Furthering Lessons of Maxwell's Demon
119. **Assignment**: Based on the Framework for Project at Sl. No. (118), develop the proposition that: Information Integrity Technology constitutes a control technology for an *IS* to optimize extraction, processing, distribution for use and storage of information with the objective of maximizing the operational life span of *IS*.
120. **Assignment**: "Information is a deceptive and ambiguous concept." Comment.
121. Study methodology to determine value of Information.
122. **Assignment**: In the Framework for Project at SL. No. (121), the value of information in decision making has been studied. Project at SL. No. (120), in its introductory paragraphs, has also talked about mathematical definition of information when information is considered at transmission/ communication level. Study the methodology developed under Information Theory or Communication Theory for calculating quantity of information in transmission.
123. **Assignment**: Study the concept of Utility of Information as discussed in the Literature.
124. Based on the investigations presented through the Frameworks for Projects at SL. Nos. (116-117), develop Information Development Life Cycle (IDLC) for a problem solving situation in a complex and changing environment characterized by uncertainty.
125. Based on the investigation under Project at SL. No. (124), study Overall Information Integrity Development Lifecycle Model.
126. Construct an example to illustrate I*I view for a business process in a complex,

changing environment characterized by uncertainty.