WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

The Information System: Data or Intelligence?

Otto H. Poensgen
and
Zenon S. Zannetos

404-69

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139
The Information System: Data or Intelligence?

Otto H. Poensgen
and
Zenon S. Zannetos

404-69

July 1969
Abstract

The extensive use of computers for management purposes has raised the issue of the role of management information systems. Unfortunately the rush to exploit the advantages of the computer, i.e. speed, consistency and "patience", introduced a bias toward massive and "brute force" approaches for data generation. This paper argues that the proper role of information systems is to provide managers with intelligent information rather than raw data. It then proceeds to describe the stages through which the information system can advance to acquire the necessary intelligence so as to aid managers in their planning and control activities.
The Information System: Data or Intelligence?
Otto H. Poensgen and Zenon S. Zannetos

I. Introduction

We postulate that the purpose of an information system is to provide the decision makers with intelligence regarding:

1. The choice of objectives
2. The determination of critical decisions and operations
3. The efficient allocation of resources, (human, including the timing of their deployment)
4. Organization of the above steps.

In the above steps we also include the necessary modification in the light of post decision intelligence.

To serve these purposes especially resource allocation and control, a management system needs capabilities for:

1. Purposive search
2. Data gathering and storage
3. Data processing and transformation
4. Inference

The last two steps convert data into intelligence. It is the purpose of this paper to:

* The authors are respectively, Professor of Business Administration at the Universitaet des Saarlandes, Saarbruecken, Germany, and Professor of Management at the Sloan School of Management, Massachusetts Institute of Technology.

1 Intelligence as used in this article is to mean data selected and structured such as to be relevant in a given context for a decision. We recognize that intelligence so defined is the output of a system (human or otherwise) with a capacity to apprehend facts and propositions and their relations and to reason about them. (Webster's Seventh New Collegiate Dictionary, p. 440).
1. Link these steps in a conceptual framework

2. Indicate what to expect of an intelligent information system

3. Point out step by step where progress can or should be made.

II. The Information System - A Conceptual Scheme

The information system may be viewed as a part of the control system, which deals with the necessary signals, which join together the various parts of the control process i.e. the reception and measurement of messages, transformations, comparison with other messages or other pre-stored information (incl. standards), further transformation to information, and, finally, plans - including those to modify the system itself. The control system contains in addition to the above, the effector that makes physical adjustments on the basis of the signals.

Exhibit 1 portrays the information system: its elements and their connection. The components of the system are data, information, plans, and the central activity, which is termed intelligence. Data, information, plans, reported results are both food for and output of intelligence. The environment of the system as far as it impinges upon the system consists of the firm's activities, the firm's management, and the firm's environment. Let us now go into the details.

Intelligence has two functions. First, and this is termed intelligence of level 0 to 1, it works on data to generate plans for the activities that implement the purpose the information system exists for, i.e. procurement and transformation, of physical, financial, and other resources

1 "Data" is a term that will not be defined here but is assumed to be in agreement with its Latin meaning.
THE LOOP Still "OPEN"  
(Control System)

Wherever flows or activities are not fixed by procedure the loop is closed by managerial activity.

Broken lines indicate a flow not part of the system or a variable.

1) Necessary to tailor response time to exigencies of the problem - e.g. fast translation of results into a change of the for the activities or to implement management by exception.

2) Including the alternative of doing nothing (i.e. no change).

3) Not to clutter the exhibit any further the impact of the environment, the activities, the information system upon the manager has not been shown.

Environment of the firm

DATA BASE

Data are classified for 'tag' as indicated by meta-intelligence.

Environmental influences upon the outcome

Plans and other intermediate or other potentially useful information are added to the data base.

INTELLIGENCE

Selection of data for relevance for problem at hand.

Transformation of data into (i) a plan or (ii) a result.

Screening for elimination from the data base

0 < level ≤ 1

META-INTELLIGENCE

Improvement of intelligence activities (in view of forecasts, plans, results, changes in the environment).

1 < level ≤ 2

Improvement of the above process (i.e. of 'simple' meta-intelligence).

2 < level < ∞

MANAGER

Persons outside the system, taking, altering or abolishing plans.

Information on the results is added to the data base.

Information on the results is added to the data base.
of the firm and distribution of products or the services in the pursuit of the firm's objective(s); second, it works on the data to improve the intelligence process itself, this we may call meta-intelligence ($1 < \text{level} < 2$). The latter includes improvement of meta-intelligence itself. It seems fruitless i.e. devoid of operational usefulness at this stage to go beyond this and speak of meta-meta-intelligence, etc., ($2 < \text{level} < \cdot$). At both levels, that is "intelligence" in the more narrow sense and "meta-intelligence," the functions are absorption of data and selection for relevance, generation of requests for relevant data, and transformation and transmission of the output. The output may be called information, namely, data perceived as relevant or potentially relevant\(^1\) in the given context. The information may be an intermediate output (e.g. a forecast) that is added to the data base, and at the same time processed further until a "plan" has been generated, which is a special category of information. The plans themselves will, therefore, become part of the data base. They again may be plans pertaining to the above-mentioned activities, or they may be plans for improvement of the information system itself. A final function of intelligence is the discarding of data, when information becomes obsolete or storage space is required for more important information or for transformation functions.\(^2\)

---

\(^1\)For a more precise definition of the term see any text on information theory, e.g. Raisbeck (14) Ch. I, or Ashby (3), p. 177-179.

\(^2\)Even if storage capacity were unlimited it must be kept in mind that the larger the stored data base, the longer the time required to retrieve data unless there are search procedures and a classification scheme for the data that takes into consideration the probability that a datum will be needed. Even so, in many cases it will be more economical to throw away a piece of information than to downgrade and reclassify it.
The (non-intelligence) activities as controlled by the plans and affected by the firm's environment will generate results which then are added to the data base. The data base thus consists of:

a) items generated by or collected from the environment, including the activities, (results) upon the system's request - as a standing order or ad hoc.
b) intermediate output of the intelligence activity,
c) plans
d) items the system is made aware of without its explicit request.

To implement d), of course, there must be receptors and storage space. An unexpected report about a competitor's activities may serve as an example. If there is no room for a vital piece of information (or for any other reason) management may change, make, or abolish plans, thus superseding the system. The importance and frequency of such action is an indication of the level in the hierarchy of the system.

The cycle length between data - intelligence - plans - results - data will of course be different for different data, different plans, different results. It is not even fixed in any given instance. There may be no feedback unless results surpass or fall short of a predetermined value. We might term this system control by exception. Or, activities of meta-intelligence may lead to resetting cycle time.

The above description holds for any subactivity or the firm as a whole. It is a matter of meta-intelligence to determine how close an integration of the various subactivities should be sought for.

All of the above does not describe what should be but what is. In a primitive system (level below 1) much of the data exist only in the memory of the human actors, and the actors have much discretion in their behavior, or to put it less complimentary: different human beings would perceive
information differently, make different plans and arrive at different conclusions on the basis of differences in character, intuition, education, experience, and access to data. The entire modification of the system (meta-intelligence) might be outside the system and within the human domain. The more advanced the system the less interference there is with plans directly, with intelligence, and meta-intelligence. Of course, some interaction will be there in any case: without purposeful human interaction the system would run for ever or deteriorate regardless of what happens. In this context it is appropriate to clarify the meaning of the word "level". "Level" is in general a hierarchical ranking. "Level 1" implies a deterministic application of rules with no modification of the information system occurring through the system itself (e.g. no resetting of standards). Level 2 includes those activities which affect the perception of the status of the system but it does not include modification of the rules which govern its behavior. Thus the system may be capable on the basis of feedback received to choose among prestored alternative decision rules, and reset the standards which in turn will affect further action. Finally "level 3" is concerned with the rules which govern the overall behavior of the system. It should be obvious that the higher the level at which an inanimate system operates the more intelligent it is, the more "objective and verifiable"("scientific") it appears, and to the extent that it has institutionalized the learning experience of the organization it is not, in the short run, affected by the predilection, training and experience of any particular individual.
III. Stages in the Process of Acquiring Intelligence

This section of this paper attempts to render more concrete the preceding section's conceptional scheme by breaking down the continuum of system sophistication that exists in reality, into distinct stages to facilitate the exposition. The classification may also line out the direction progress has taken or may be taking, though no claim of immutability is made for the sequence. The description of the higher stages provides the authors with the opportunity to voice their own opinions as to where progress is promising of desirable.

1. Storage of Raw Data without Classification

Since the amount of information that potentially could be stored is infinite for any realistic scheme, rules for inclusion or exclusion of data must exist. An example of a primitive system might be a street vendor who, say, throws all his receipts of cash from his suppliers in a box without further ado and keeps no other record. The only purpose might be to be able to prove that a payment was made.

2. Development of a System for Classification and Aggregation of Data

This we may call elementary organizing process. The system is still very unsophisticated, but it operates under simple rules of classification which facilitate storage and retrieval. Thus, the user can associate on the basis of such classification and can recall stored data other than through retrieval and processing of the entire file.

3. Automatic Comparisons of Data

Using the system classification schemes, the system juxtaposes data units output, for example by reporting this year's sales to date with last year's sales to date, or cost incurred in Factory 1 and cost incurred in Factory 2, cost element by cost element, or

---

1 This section is partly based on material included in Reference (18).
even estimated cost (if it were a formal input to the system) and actual cost. This characteristic enables the system to extract differences which are then used for problem identification by the manager or a subordinate assigned to this task.

4. Simple Manipulations of the Data

Cost per unit might be calculated, or the difference between actual and estimated cost in absolute amounts or percent. (Standard costs provide an example.) The system now has the capacity to store simple cause- and effect relationships for measuring deviations. Thus, the system has acquired the capability of incorporating models of behavior. On the basis of what is stored the inputs are associated with (a) the model in which the inputs appear, and (b) the chosen level of the model, which assigns numerical values to variables in the model. The system, however, has no capability to challenge the models. The latter are still exogenous to the system. It is useful though to have such systems capability for budgeting purposes (e.g., flexible budget procedures), and for testing the quantitative consequences of "alternative" levels of input factors. We must also note that for the first time the output of the system can be classified as information, rather than data.

5. Institutionalization of Management by Exception

At a fifth, conceptually different level, the kind of output of the system is dependent on the magnitude of the stored data (either directly or via transformations that are performed automatically).

Such a system might point out a variance only if it reaches a certain absolute level or a certain percentage standard cost. The system reacts on the basis of prestored cues. It may even assess the probabilistic level of significance of deviations on the basis of prestored models. This is a step towards institutionalizing
management by exception. Such progress is easy to implement with modern computing machinery, since it has the logical capability to choose between branches on the bases of the magnitude or sign of predetermined variables within the system. Electronic data processing equipment should enable us to make progress even beyond the above level, for example: a more complete information system need contain:

A (1) not only the estimated inflows and outflows
(2) including the bases for estimation
(3) and the actual outcome
(4) with an analysis of the differences and their causes

B (5) but also the principal opportunities rejected
(6) with the methods of estimation applied in their case
(7) and the cause of their rejection
(8) and, finally, an ex-post-facto analysis of what would have happened had the rejected alternative been chosen
(9) compared to what happened (3) to the alternative chosen.

The reason is this: it is entirely possible that anticipations were realized and again occurred in agreement either absolutely or relatively to the expectations, but that alternatives contemplated and rejected at the time would have led to better (absolute) results. The post facto analysis hopefully can lead to a better choice of alternatives via a better prediction of their consequences. The opposite case is also thinkable: the alternative chosen may have led to absolute losses or losses in relation to expectations, but the rejected alternatives would have led to even greater ones. Under these circumstances the actual losses (absolute and relative) need not be interpreted as indicating
false prediction, analysis, or execution. The best result possible may have been reached, if all possible alternatives were indeed analyzed at the time of the original decision.

We also need a post hoc comparison not only of actual vs. estimated results of the chosen alternatives and a comparison with those the rejected alternatives might have yielded but also an ex-post analysis whether all alternatives (relevant) were considered or whether ex post better ones are now recognized to have been available. The question must be raised periodically whether the process of searching for alternatives stands in need of improvement, but this leads us beyond the level under consideration.

6. Examination and Change of the Models Themselves

A system which possesses the sophistication required by levels 4 and 5 depends on exogenously derived rules. It is their injection from outside into the system which enables it to transform (manipulate) data and report them. These rules reflect the system designer's (or his customer's or superior's) expectations of the future, his assessment of the possible, and his conception of causal relationships between input and output— in short the designer's 'model' of the firm or a segment thereof.

To analyze what would have happened, had one of the rejected alternatives been accepted, surely is no more difficult than the analysis of the potential consequences for the alternative before rejection, since all that information and more is available (e.g. actual prices, material supply, and the like). The difficulty would rather seem to be in a bias towards the rejected alternatives—small difficulties and even major ones like breakdowns, quite likely would have been overlooked, while, of course, they do show up in the actual figures for the alternative chosen. To avoid this, a party (or a part of a computer's memory and instruction set) not knowing about the actual events should estimate the actual results with no more information than is available for the rejected set. For repetitive decisions this may be discontinued after a bias has been established with a sample of actual outcomes on the one hand, and outcomes for the same alternative estimated ex post with incomplete information. This may lead towards a better feeling of what classes of events are predictable and which are not.
his model of a process or a project. More advanced systems are capable of updating, revising, refining, challenging, or even plainly refuting these models. Such models, we might say, have 'intelligence' of a level higher than 1.1

At this level there seems to be limitless room for improvement. In the last part we will try to be specific and indicate what are in our opinion, some of the more promising avenues, areas and techniques for advances. Some preliminary remarks will provide the setting.

(a) No model or other statement ever describes a part of reality completely. Some aspects, details, surrounding circumstances will always be missing. Any description of "model", to use the in-word, will obviously abstract reality.2

(b) No statement about reality3 can ever be made with certainty. The point has been argued forcefully by A. J. Ayer.4

(c) From (a) and (b) it follows immediately that there is no such thing as decision making on the basis of complete information. There are only degrees of incompleteness.

---

1 This is simply a convenient label for a system with model changing capability. The definition has the advantage we hope, of not clashing with the meaning the word has for the average reader.

2 For a cogent presentation of this assertion, see W. Ross Ashby (3) p. 109 f.

3 Other than about one's own perceptions.

4 L. A. Ayer (5), p. 36-68.
(d) Two or more data may be manipulated in a number of ways that is infinite for all practical purposes. For example, a cost element at standard and at actual may simply be juxtaposed or the difference may be calculated, or the differences may be put in per cent of either of them or the average etc. This rather trivial remark leads us to the next point:

(e) A good information system possesses the usual desiderata such as speediness, low cost, absence of ambiguity. It also "milks" the available data well, overlooking no essential aspect contained in it by transforming it and reporting it in a form that is most helpful for the issue at hand, (the plan). This is clearly an "intelligence" activity, especially if at this stage the method itself of manipulating and reporting is modified on the basis of criteria and ways that are prestored i.e. injected by the human being.

The reporting may be done in several ways simultaneously if there are several issues. Users are thus steered away from a misleading monistic figure to serve all purposes. Finally, a good information system has the reported information point beyond itself by indicating where additional information or increased precision would be most valuable.

7. Highly Advanced Stages

Still more advanced stages bring the functional specification itself under scrutiny, the system being the more sophisticated the wider the range of functional forms is that it challenges, updates, and develops, and the more general the criteria for
the system's progress are. A sophisticated system thus provides a continuous progress toward determination of the applicability of alternative functional relationships and at an even higher level derives abstract relationships of more general applicability.

IV. An Example

Let us start elucidation of some of the tasks that we see for the system at the higher stages by discussing an example. A relatively unsophisticated system, like the standard system normally found in use, may have a numerical standard, say maintenance dollars, set by a human being. Standard and human being are part of the system. The system calculates variances, ratios, in our example dollars of maintenance per machine hour, and on the basis of some decision rules, that form part of the system, prints them out or Withholds them. The human being, often a cost accountant, decides whether and where the standards themselves require updating.

In a slightly more sophisticated system, as for example in flexible budgeting, the cost accountant endowed with the glorified title of "systems analyst" might simply specify a relationship

\[ C = F + V Q \]

\( C \) = Total Maintenance Cost
\( F \) = Fixed Cost for maintenance
\( V \) = Variable Cost per machine hour
\( Q \) = Number of Machine Hours of Operation

but leave it to the computer or its minions to determine the numerical value of \( F \) and \( V \) by some technique such as linear regression, or a simple visual fitting of a line through past data, or by setting so-called scientific standards.
Of course, if the relationship is not linear, but if there are curved segments (where overtime occurs, where non-preferred equipment is used, etc.) the usual flexible budgeting system will not point this out as such, let alone correct itself. It may grind out many variances which are labeled significant and it is left to the systems analyst to interpret them correctly by changing the relationship.

That is, the system either labels a discrepancy between planned and actual result the fault of execution or it leaves this question undecided and just locates the discrepancy more or less precisely. The determination of the course is left to someone outside the system (stage 5).

If it is felt that the mathematical representation is not correct but the main causes of maintenance are, the functional form may be changed. To stay with our example, a quadratic term may be added.\(^1\) Suppose now, that the unexplained discrepancies continue casting doubt on the cause-effect relationships. The system then should be capable to turn from a search in depth (functional forms) to a search in breadth (possibly at random) by looking for other variables associated with maintenance costs. This avenue increases system sophistication.

The above methods may not only be inefficient but the expected search might be long and the variable the system comes up with, may be very enlightening. It might suggest labor hours or amount of material processed, both of which are closely related to machine hours but provide no further explanation of our discrepancy. A higher level system might now change the functional form of specification. Upon finding a high collinearity between machine hours on the one hand and labor hours and amount of material processed, the system cannot discover this unless exponential forms are contained in the memory. The memory need not be that of a computer. A file or the knowledge of the human executant may serve the same purpose.

---

\(^1\) If the true relationship, say, is exponential (as in interest calculations)
processed on the other, it might investigate whether a shift in the proportions of machine hours to labor hours is associated with a deviation of the maintenance hours from their standard. A correlation might reveal that jobs demanding a high labor input per hour strain the machinery—possibly because it is no longer run in the automatic mode, but in the manual one with its possibilities of human mistakes.

Further progress can be made by structuring the search. The system designer might have instructed the system to try in such a case to correlate with the discrepancy, overtime, new hiring in the last month, time worked on the night shift specifying the order. The computer proceeds until it has found a plausible cause or until the set of variables it is instructed to investigate is exhausted. Thus, a message to this effect might be the only output. Such a system would tackle the same problem the same way every time, it would learn nothing.

A still more progressive system would record the variable(s) that provided an explanation. In our example this might be the time elapsed since the last overhaul. When the same problem arises again, the system might start the search for an explanation with the variable that proved most helpful the last time (or most frequently in the past) and take up the human choice next. Since the search procedure depends on prior solutions of the same problem we may say that the system learns.

It is well to point out that the system cannot infer from the data everything a human being can. In our example, it might be that whenever very hard steels are milled there is an increase in maintenance shortly
thereafter. For the system to come up with that answer it would have been necessary to have hardness of steels processed recorded with the various jobs, and, furthermore, to have hardness of steels included in the set of variables to be examined.

Whatever the level of system sophistication, it is indispensable that the system has some capability to determine when the search is to be stopped and appeal to its human master once it encounters data or constellations of data that it finds inexplicable and "beyond its pale". This should not be a "go - no go" type of decision but should be much more differentiated.

In the simplest of cases it states that a particular, well defined piece of data is missing - it might just be late this month or quarter. In other cases the system might indicate that data on a particular area are missing, say demand or expected level of economic activity in an industry, without specifying which type of data best would close that gap. In still more serious cases it might declare that it is stymied, lining out briefly the question and the steps taken so far and the degree of confidence in them.

We would expect progress to be faster in the direction of "in-depth analysis" than in the direction of "in-breadth analysis" because the computer cannot (as yet) recognize patterns involving many variables in the way of human beings. On the other hand the computer has the advantage of speed and consistency, in the sequential analysis required for an "in-depth" search.
V. Concluding Remarks

We find it difficult if not impossible to distinguish clearly between "goal setting", "decision making", "(operational) planning", "control", and "execution". The differences seemed to be in the connotation; "decision" carries with it a determination for action, for getting results (although we speak of tentative decisions just as we do of tentative plans or goals). "Goals" or "objectives" seem to leave the strategy of details of implementation open, whereas "planning" (especially of the operational type) seems to promise just that. Whether a set of figures constitutes a goal or an action plan thus depends on how specific this set is felt to be. If the person looking at the set of figures feels it is as specific as he will make it, he may term it a plan! The recipients of the plan or a part of them may term them goals, if in their mind and from their managerial level it is not evident how they are to be implemented, and they feel that these details have to be worked out ("planned") first.

"Operational planning" very often is meant to be the original determination of steps deemed necessary or desirable to define and reach a certain goal, or several goals, whereas "control" includes modification of plans if required by subsequent events in order to reach the goal--or even adjustment of the goal. The term also carries overtones of activities that belong to the province of behavioral science--such as motivation and performance evaluation of people--principally of those executing the plans. If a modification is major, it is occasionally called "replanning". Generally, however, we maintain "control vs. planning" is not a task-inherent
characteristic. To decide whether a particular job is to be classified as a controlling rather than a planning activity we should know the details of the task environment, including the organizational hierarchy. Both planning and control functions are inherent in all activities of an organization and apply to all levels of management. Let us clarify this by an example. Suppose that the headquarters of a conglomerate company decide in the light of recent relative profit performance to cut the investment budget of division A and increase that of B. This is a modification which may be termed an act of control of headquarters. At one level lower--divisional headquarters--this means a major change for them requiring replanning. Thus, they might decide not to build a new plant but to expand an old one. Still one level lower, in the engineering department, it means a new set of plans apparently uncorrelated to the previous one, must now be worked out. The engineer might not even know that the old set is scrapped, for him it is a totally new problem. He is planning now.

This suggests a further question: Is it sensible to distinguish between "planning" and "execution"? In executing the plans or responding to a "controlling" action of his superior, a subordinate may make plans on how to modify the actions of one or more of his own men by first reflecting (hopefully) and then communicating with him. (This process may of course be interactive). Then he observes the results of the communication process, reflects again and communicates again to modify (control) the

---

1 This is somewhat in line with a view expressed by one of the authors, that the traditional separation of the corporate (or even divisional) functions in staff and line is largely out of date for an enterprise that has an integrated approach to any given product or project including purchasing, personnel procurement, production, etc. See Poensgen (12) p. 391-392; see also Fish (10) for a different approach to the problem.
process. This may repeat itself at lower levels right down to the worker. We must notice that even at the worker level it is common that he himself does not perform the physical change of a location or form but simply initiates the action, e.g. by pushing a button, communicates his wishes to and observes a machine. As Richard A. Beaumont, (6) p. 4 aptly remarks "We are moving from a society of doers to a society of observers".

To sum up, what is needed is an information system capable of applying rules of correspondence to data stored on a reasonably large number of classes of variables and also capable of searching for the most appropriate rules to find associations within and between them guided by its own (as well as the planner's) experiences and, finally, capable of modifying and elaborating these rules on the basis of the same experience. Such a system may be described more or less synonymously as an associative one, or a system of self-organizing files, or one of pattern recognition capability or a system capable of inference and learning from experience.

This brings us to the end of our exposition. We have left out at least two aspects that will vitally affect the specific form a comprehensive information system will take. First we have said nothing concerning the goals of the organization served by the information system. The reason for the omission is simple; "No one really knows what individuals or any organization wants, or what they should want".¹ It also seems to us that if the goals are specified clearly enough to permit measurement of progress towards them, and of sufficient duration to make sequential adaptation of efforts towards these goals meaningful, the considerations of this paper are

¹American Accounting Association (1) p. 69
not greatly affected by them, since we consider methods that promote efficient ways of reaching a goal, and apply to a wide range of goals or objectives.

Second, our paper has not touched upon the issues of motivation, performance evaluation, and organization structure. It is evident that these issues will have a substantial influence upon the design of information systems and will no doubt force modification of methods that would be applied under a more mechanical view of feedback and feedforward. In order to determine the exact nature of these impacts we must do both more thinking and also carry out extensive empirical investigations on the organizational aspects of information system design especially in the light of modern technology.
BIBLIOGRAPHY


2. ANTHONY, Robert N., Planning and Control Systems - A Framework for Analysis, Boston: Division of Sponsored Research, Graduate School of Business Administration, Harvard University, 1965.


