Assessing the Value of IS

A Systematic Approach Toward Assessing the Value of an Information System*

By: Niv Ahituv

Abstract

A multiattribute utility approach is adopted to assess the value of an information system. Various economic analyses of the value of information are reviewed and the conceptual problems regarding the definition of this value and some measurement difficulties are discussed. A list of possible utility attributes is proposed for a reporting system value assessment, and for each attribute a measure and a utility function is suggested. Some techniques which constitute a joint utility function are presented, accompanied by two examples. A real case of minicomputer selection is given in order to illustrate the structured approach.

Keywords: Information evaluation, system selection, information economics

ACM Categories: 2.44, 2.49

Introduction

The various approaches toward information system evaluation can be roughly classified into two different categories.

1. Pragmatic assessment such as cost/benefit analysis. A review of this approach was published recently by King and Schrems [25].

2. Theoretical evaluation based on Decision Theory. The general foundations were developed among others by Ackoff [1], Marschak [30], and McGuire and Radner [33]. Some of them have been applied in accounting research (Feltham, [15]; Mock, [36]; Feltham, [16]; Demski, [11]; Butterworth, [8]).

A third approach can be placed in between the above two. This approach attempts to formulate a utility function for certain information problems and then to find the system which gives an optimal solution for the function. Grochow [19] and Dexter [12] proposed such an approach toward timesharing systems evaluation. Grochow's work is a rare case where empirical data as well as a theoretical model were used. Ahituv [2] uses a similar approach to evaluate data entry and validation systems.

This article should be classified in the third category. It will attempt to list possible attributes of the user utility function associated with reporting systems. It will also attempt to assess the nature of the utility gained by the user from each individual attribute, and then to outline how those fragments of the utility function can be combined.

The next section discusses conceptual problems with regard to the definition of the value of information. The third section discusses some measurement problems. The fourth section lists possible attributes of user utility gained by using reporting systems. The fifth section outlines some techniques of assessing and optimizing user's utility function, or at least reaching a satisficing solution. Some concluding remarks are indicated in the last section.

The term "satisficing," coined by Simon [44], is used here to distinguish between an optimal solution and a solution which might not be the optimal but at least meets some mandatory constraints.

*MIS Quarterly/December 1980 61
Conceptual Problems

Any fundamental discussion with regard to the value of information raises at least three major questions.

1. Whose value are we talking about? Is it an individual, a team, an organization, or any group of individuals?
2. What type of value are we talking about? Is it the value perceived by the user; is it the marginal improvement of the user performance revealed after receiving the information; or is it the normative value analytically computed?
3. Who is performing the evaluation and when? Is it the decision maker or user served by the system who evaluates it either continuously or ex-post; or is it an external or objective evaluator who recognizes all the parameters and performs an ex-ante analysis?

With regard to the first question, usually there are no severe theoretical obstacles as long as the utility of an individual is considered (see for example Radner [39]). When an information system should serve a group of individuals, we have to face the “Impossibility Theorem” [4] before allowing any further discussion. According to the theorem, it is not possible to assign a common set of preferences to a group of individuals, unless the assignment is “imposed or dictatorial” [4, p.59]. Usually, the problems evolving from this theorem are overcome by assuming that common preferences can be assigned to the group, as done in Team Theory [31, 40]. Similarly, this article will assume that a common set of preferences exists, e.g., a utility function will be developed for the user of an information system, where the notion “user” does not necessarily imply an individual.

With regard to the second question, three different definitions of the value of information are found in the literature. The following example illustrates them.

Suppose a gambler in a horse race is able to purchase information about the results of a future race prior to its commencement. Suppose the gambler believes that the proposed information is true, complete and private, i.e., it will not be disclosed to the public, and is probably willing to pay for that information. The amount the gambler is willing to pay reflects the subjective evaluation of the worth of the information, namely the perceived value of the information. Suppose later, after the information had been acquired, some income was gained from gambling. The difference between that income and the income that could have been gained without the information is the virtual betterment of the gambler’s performance, namely the revealed (realistic) value of the information. By using some simple tools of Probability Theory, we can also compute the expected payoff of what should have been gained if the gambler used the information optimally—assuming the gambler wished to maximize the expected payoff. The difference between this payoff and the expected payoff when the gambler has no prior information is the normative value of the information.

Thus, there are three different values of information; see related discussion in Mock [36]. The perceived value is usually discussed in papers based on empirical research. Gallagher [17], and Munro and Davis [37] measured this value by using a semantic differential scale. A similar approach was taken by King and Rodriguez [27]. Neumann and Segev [38] investigated the perceived value of a banking information system among branch managers. Ronen and Falk [41] measured the “monetary equivalent” of the perceived value by “selling” information to the participants of their experiment; see also Zmud [51].

The revealed value of information has been measured in some experimental research. A series of such experiments were carried out at the University of Minnesota [13]. Hedberg [21] and Edstrom [14] investigated the additional value gained by using an online reporting system in comparison to a batch processing system. Mock [35] compared the revealed value to the normative value and found them to be different. In a way, his results conform with Simon’s ideas about decision makers being satisficers rather than optimizers [44].

The normative value is intensively discussed in Information Economics. Feltham [16], Tapiero [50], and Stohr [48] calculated this value for inventory control information systems. Butter-
worth [8] presented its application to accounting systems.

There is not an ultimate answer to the question of which value should be considered. The perceived value is available only if experiments and interviews are undertaken, and it is useful for economic considerations only if it is expressed in monetary figures, which is not frequently done. The revealed value probably best reflects the realistic impact of the information. However, when information for middle or top level decision making is investigated, it is very difficult to estimate that value. This is because decisions in these levels are unstructured, their implications are likely intangible and range over a long time, and therefore the revealed value of particular information is hard to isolate. Both the perceived and revealed values require some post hoc analysis. Further analysis can be done by a normative method.

This article adopts an approach based upon the Multiatribute Utility Theory [24]. In a way it is a combination of the above “values.” The determination of the utility attributes and the characterization of the user preferences [51] can be based on the perceived or revealed value. Further analysis can be done by a normative method.

Referring to the third question of the above, the identity of the evaluator has some implications on the approach to the evaluation itself [16]. If the evaluator is an operator within the system (i.e., the decision maker according to Churchman [9]), he or she is not supposed to have a comprehensive knowledge about all the states and the components. The evaluation is performed on a day to day basis, resulting in gradual improvements to the existing system. If the evaluator is external to the system, some sort of a “know-all” system designer [9], the evaluation gets the overall perspective required for designing a new system.

This article takes the point of view of the outsider, the objective designer who wishes to better the utility of the user. Consequently, the approach would be to devise tools for pre-implementation, or ex-ante, rather than post implementation, or ex-post, valuation. However, since we would like to incorporate user’s views and expectations into the analysis, we have to obtain some data which are normally available only after implementation. This type of data is obtained through experiments [13, 35], questionnaires [37], and interviews in empirical research related to perceived and realistic values. These methods may be useful for the approach discussed here. With those remarks in mind, we can proceed now to discuss some technical problems.

**Measurement Problems**

It is almost generally agreed that the utility gained by using information is not a unidimensional function [26, 46, 51]. McGuire expresses it bluntly by saying that “any search for a one-dimensional measure of ‘informativeness’ is a vain one” [32, p. 109]. Thus, a multiatribute utility function should always be selected for information system evaluation. This approach involves some technical problems.

First, what are the relevant attributes? Terms such as timeliness, accuracy, reliability, etc., tend to play an important role in the evaluation process [23, pp. 82-83]. However, it is not certain that all of them are always relevant. In other words, it is likely that the dimensionality can be reduced by assigning only the relevant attributes to each case investigated. For example, Neumann and Segev [38] found a spill over effect of the content attribute over several other attributes, i.e., the evaluation of content causes evaluations of other characteristics to be similar.

Second, how would each attribute be defined and measured? The answer is relatively simple when timeliness is considered, but it might be nontrivial with regard to some other attributes.

Third, how would the effect of each attribute on the utility function be assessed? For example, it is relatively easy to declare that users prefer short response times to long ones, but is it a linear or a threshold function?

Finally, what is the joint utility function constituted from all the relevant attributes? In order to

---

The term “utility function” refers to a cardinal function reflecting preferences.
formulate the joint utility function one has to assess the tradeoffs among the various attributes, and those are not so easily detected. Moreover, sometimes only an ordinal, or value, function is feasible, and sometimes even achieving that is too ambitious.

The next section develops a list of possible utility attributes. Each attribute will be developed through the following steps:

1. definition of the concept behind the attribute,
2. definition of a possible measure for the attribute, and
3. characterizing the nature of the [marginal] utility function related to that attribute.

Before proceeding it is important to note that this article presents an approach toward information systems evaluation. It does not pretend to propose an ultimate method, nor to provide the readers with an exhaustive list of attributes. It only claims that a similar approach can be adopted in many cases of information evaluation.

### Possible Utility Attributes for a Reporting System

In order to simplify the presentation, we confine ourselves only to a reporting system. We assume that at a given point of time, \( t_0 \), there exists a database\( ^3 \) containing data already validated and updated. The only feasible operations on that database are selective retrieval, manipulation (aggregation, sort, computation, etc.), and displaying. We exclude any reference to the preparation of the data, namely, the capturing, validating, and updating of the data files, \( i.e., \) we assume that the user cannot affect the creation of the database.

First, three major categories of attributes are proposed:\( ^4 \):

1. timeliness,
2. contents, and
3. format.

The cost attribute is not introduced here; however, it is discussed in the subsequent sections. For each of the three categories, we will try now to determine various attributes.

### Timeliness

One has to differentiate between online and batch processing reporting systems before evaluating the time factor. Online systems are usually designed to provide fast answers to queries. Therefore, in such cases, the response time should be considered. Batch processing systems are usually designed to provide information in a fixed frequency thus, in addition to response time, the frequency should be considered. Two attributes, therefore, emerge.

**Response Time**

This is defined as the difference between the time the information is received and the time it was requested (\( i.e., t - t_0 \)). Assume \( t_0 = 0 \) and get \( t \) as a measure.

**Assumption.** The utility function is a nonincreasing function of \( t \) \( [43, 49] \). The nature of the function might be continuous, or discrete (threshold), \( etc. \)\( ^5 \).

### Frequency

It is important to provide data within the framework of a given frequency in cases of periodical reporting systems. For example, the share prices in the stock exchange are highly valuable as long as the trade continues. The usefulness of the information decreases after the end of the business hours, and then starts to increase again before the beginning of the new business day, and so on.

In periodical systems, the variable remains the same as above (\( i.e., t \)), yet the utility function is different.

**Assumption.** The utility function is a decreasing cyclic function of \( t \) (such as \( u_t(t) = \sin(a_t)(t + b_t) \)) where \( a_t \) and \( b_t \) are positive constants).

\( ^3 \)The term "database" is used here to refer to any collection of data. It is purely conceptual and does not refer to any specific hardware/software technology.

\( ^4 \)See related, though not similar, discussion in Neumann and Segev \( [38] \), and Zmud \( [51] \).

\( ^5 \)All the attributes are taken as being deterministic for the reasons of simplicity.
Assessing the Value of IS

The last function can be exemplified by taking $a_1 = \pi/2$ and $b_1 = 1$. Some typical results are exhibited in Table 1.

### Table 1. An Example of a Cyclic Decreasing Function

<table>
<thead>
<tr>
<th>t</th>
<th>$u_t(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0.20</td>
</tr>
</tbody>
</table>

It can be noticed that the figures reflect periodical changes in the utility as well as a decay over time. Of course, this is only an example and in a real case one has to assess and calibrate a representative function.

There may be some other time related attributes. For example, Keen and Scott Morton mention the currency of data as one of them [23, p.84]. However, due to the initial assumption that the user does not affect data generation timeliness attributes are not discussed any further here.

### Contents

The content characteristics of a report are somewhat vague [38]. They may include various aspects such as accuracy, relevance, fineness (aggregation level), or order of magnitude (units, thousands, millions), etc. This article proposes two attributes which are content related— the similarity and aggregation level.

### Similarity

The notion of similarity [2], which will be explained later, is introduced in order to represent the relevance and the accuracy of the reported data. We will introduce it from an abstract point of view and then discuss its applicability.

Let D be a finite set of data actually received by the decision maker. Suppose that after a decision has been made, and the real state of the world is already known, the decision maker can determine what should have been the optimal set of data, ex-post. In other words, the decision maker, restricted to the data available in the database, designates a subset of the database that, if had been displayed, would have led to the best (subjective) decision. Denote this subset by $D^*$.

Let $m(A)$ be an operator which provides the number of members in a set, and assume that each data item in $D^*$ has the same importance with regard to the decision problem.

**Denote.**

a. $m_1 = m(D \cap D^*)$, the number of members that are common in D and $D^*$,

b. $m_2 = m(D^* - D \cap D^*)$, the number of members of $D^*$ that were not provided to the user, and

c. $m_3 = m(D - D \cap D^*)$, the number of members that were provided but not desired.

A comparison between $m_i$, $i = 1, 2, 3$, and some common notions of statistics will be presented later. Figure 1 helps to picture $m_1$, $m_2$, and $m_3$.

Now define $s$ as follows:

$$s = \frac{m_1}{m_1 + m_2} - \frac{m_3}{m_1 + m_3}$$

Let us examine some properties of $s$:

a. If $D^* = D$, i.e., the data provided are exactly the data desired, then $m_2 = m_3 = 0$ hence $s = m_1/m_1 = 1$.

b. If $D^* \cap D = 0$, i.e., the report does not provide any desired datum, then $m_1 = 0$ hence $s = -m_3/m_3 = -1$.

c. If $D \subseteq D^*$, i.e., the data provided are desired but still there are missing data, then $m_3 = 0$, hence $s = m_1/(m_1 + m_2)$, namely $0 < s < 1$.

d. If $D \supset D^*$, i.e., the data desired are provided but there are also undesired data, then $m_2 = 0$, hence $s = 1 - m_3/(m_1 + m_3)$, namely $0 > s > 1$.

Generally speaking, $s$ is minimal, equals $-1$, when there is no similarity at all between the

---

*Assuming that neither D nor $D^*$ are null sets. Also assuming that missing data and unrequired data have the same value. The latter assumption can be relaxed by assigning different weights to the components of (1).
desired and the provided data. \( s \) is maximal, equals 1, when the data are perfect. \( s \) increases when the similarity increases. Thus, \( s \) is suggested as a theoretical measure of similarity whenever each datum carries the same value for the user. If various subsets of the data are associated with different values for the user, each subset \( j \) can be corresponded to its own \( s_j \) and then a "weighted average" of the \( s_j \)'s can be computed.

Now the assessment of \( s \) should be discussed. Referring to statistical errors of the first and second type, one can interpret the formula for \( s \) (1) as follows:

a. "An error of the first type" would be the probability \( p_1 \) that a datum is desired, i.e., should belong to \( D^* \), but it has been omitted from the report.  

\[
p_1 = 1 - \frac{m(D \cap D^*)}{m(D^*)} = 1 - \frac{m_1}{(m_1 + m_2)} = \frac{m_2}{(m_1 + m_2)}
\]

b. "An error of the second type" would be the probability \( p_2 \) that a datum is not desired, i.e., should not belong to \( D \), but it has been provided.

\[
p_2 = \frac{m(D \cap D^*)}{m(D)} = \frac{m_3}{(m_1 + m_3)}
\]

Assuming that the probability does not vary for each datum.

This is because either the datum is not relevant or because it is incorrect or inaccurate.

Equation (1) can be presented as \( s = 1 - p_1 - p_2 \), namely, a function decreasing when the probabilities of error increase.

These probabilities can be forecasted by using some past data and/or some similar cases. Mace, Crowe, and Jones [29], while discussing data validation systems, suggest that either moving average or exponential smoothing seem appropriate techniques to assess such probabilities. Cushing [10] and Bodnar [7] also discuss some implementation aspects of reliability assessment, i.e., probabilities of errors, with regard to accounting control systems. There is no unique method of assessment, but all these articles regard it as feasible.

Whenever such probabilities are not available or cannot be estimated, a surrogate variable may be considered. For example, one may find it useful to examine the adaptability of the report to the requirements of the user as stated in the initial design of the system. That might help to estimate the relevance factor. The correctness of the data can be estimated by samples [29]. Once the similarity measure, \( s \), is determined, its utility function should be assessed.

**Assumption.** The utility function is a non decreasing function of \( s \). The use of the similarity measure is exemplified in the fifth section.
The degree to which data should be aggregated is a frequent issue in Accounting, MIS, and Econometrics. Ijiri [22, appendix B] gave a mathematical interpretation for the aggregation process. Lev suggested to adopt the entropy function for aggregation measurement [28]. Generally, it is agreed that some amount of information is lost by its aggregation. Whether this loss has any qualitative significance is still a debatable issue [18, 20, 41].

A decision maker faced with many figures, potentially has a better opportunity to "know" what is going on. On the other hand, the human perception is limited [34] so one might overlook some important facts due to information overload. Sometimes the domain of possible actions or decisions to be taken does not require a full detailed report, namely, aggregated data will lead to the same decisions as more detailed data [16, p.31].

The usefulness of the entropy function [28] for measuring the value of aggregation was found to be limited only to cases where the user utility function is of a logarithmic nature [5, 41]. Nevertheless, this function can still serve as a measure for the quantity, rather than quality, of aggregation, as will be proposed here.

Let us limit our concern only to numerical non-negative numbers. Assume that a list of data, Do, contains all the raw data that can be displayed to the user. Denote Do = {d1, ..., dn} and normalize the list by computing

\[ d_i^* = \frac{d_i}{\sum_{i=1}^{n} d_i} \]

The entropy of the raw data would be:

\[ h_0 = -\sum_{i=1}^{n} d_i^* \log d_i^* \]

Suppose the members of Do are aggregated into a new list, D, whose entropy is h. The absolute loss of entropy, h0 - h, does not have a high significance since it is highly dependent upon the number of items in Do and D. A relative measure would serve us better.

Following the same line of thought as with the similarity measure, suppose post hoc, the user can determine the optimal report required for his decision making process. Denote the entropy of this report by h*.

**Assumption.** A possible utility function for the aggregation attribute has a single peak at h*.

For example:

\[ U_h(h) = -a_3[(h-h^*)/(h_0-h^*)]^2 \]

(assume h_0 > h^*, a_3 > 0)

(The coefficient a_3 serves for calibration purposes). Denote H=(h-h^*)/(h_0-h^*) and note that 0 < H ≤ 1 whenever the aggregation level is less than desired; H = 0 when the aggregation level is optimal; H < 0 whenever aggregation is overdone.

The proposed measure ignores the meaning of the figures, the title of the items, which is a common disadvantage of all the attempts to deal with entropy. However, if the aggregation is initially restricted only to "reasonable" aggregations from an accounting point of view, that disadvantage may partly be overcome.9

The problem remaining is the assessment of the "optimal entropy," h*. Actually, it is a problem of an ex-ante assessment of a value which can be investigated ex-post. Ex-ante, it should be estimated by using questionnaires, interviews [24], and experiments [13, 41]. For example, the designer can prepare some alternative pilot reports, each of which varies in aggregation level; users will be requested to opine on each alternative. In fact, if the decisions involved are fairly structured, users' preferences can be experimentally measured [41].

Entropy is not the ultimate measure for data aggregation. Many articles inquire into aggregation problems without the use of entropy (see [6, 47]). The following example illustrates an aggregation problem in which entropy has not been applied. Although the example is simplistic, we believe that it may clarify the idea of this attribute.

9Shannon and Weaver [42], when they first introduced the entropy concept to information theory, warned their successors from entropy "abuse": "The word information, in this theory, is used in a special sense that must not be confused with its ordinary usage. In particular, information must not be confused with meaning" [p. 8].
Example. Suppose a top manager supervising several departments is periodically informed by receiving N figures (for example, there are r departments and each has z budgetary items hence N = rz). The manager makes judgments and decisions by scanning the figures and carrying out comparisons between various figures appearing in the output. It is assumed that the importance of each department and each figure is the same.

Let N be sufficiently large to make it quite difficult to analyze all the figures and come to definite conclusions in a reasonable time. The maximum number of possible comparisons between every pair of figures is proportional to N², (N(N−1)), and it is too large to be handled by the manager. On the other hand, providing the manager with only one figure which is the total of the entire list of N figures is meaningless. Thus, the first characteristic of the manager's utility is determined by the bounded capacity to perceive and analyze data [34, 44].

The manager would like to detect exceptions in the departmental activities. If two figures are aggregated together and there has been at least one exception regarding one of the original figures, it is not certain that the manager will detect the exception because two counter direction exceptions might compensate each other. Even if it is discovered, the manager will not be able to promptly determine its location. Therefore, the second characteristic of the utility is the ability of control.

The above verbal description is formulated in a quantitative manner.

Denote. N — the number of the original raw data,

n — the number of figures in the report provided to the manager, 1 ≤ n ≤ N, and

k = N − n.

This problem can be examined from two points of view—pessimistic and optimistic. The pessimistic view assumes that the loss of information is maximized. For example, if $d_1, d_2, d_3,$ and $d_4$ are the original data, then in a report containing two figures they will appear in the form of $(d_1 + d_2)$ and $(d_3 + d_4)$, or similarly, so that no original figure is available. The optimistic view assumes that the loss of information is minimal; e.g., a two figure report produced from $d_1, d_2, d_3,$ and $d_4$ will exhibit one original figure, say $d_1,$ and aggregate the rest. Certainly, in reality the manager knows which numbers have been combined and which remained separate. However, since it is assumed here that the figures do not differ in importance, and aggregation is a must, the upper and lower bounds are set for the evaluation.

From the pessimistic point of view, Table 2 describes the maximum number of items for which the control is lost (denote $\bar{M}(k)$) if n figures are presented to the manager.

<table>
<thead>
<tr>
<th>n</th>
<th>$\bar{M}(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0</td>
</tr>
<tr>
<td>N−1</td>
<td>2</td>
</tr>
<tr>
<td>N−2</td>
<td>4</td>
</tr>
<tr>
<td>⋮</td>
<td>⋮</td>
</tr>
<tr>
<td>N−k</td>
<td>2k</td>
</tr>
</tbody>
</table>

The preceding table is valid as long as $k ≤ N/2$. When $k > N/2$, all the items are not controlled from the pessimistic point of view.

It is obvious here that $\bar{M}(k) = 2k = 2(N−n)$, $k ≤ N/2$. Suppose the probability that an exception occurs is the same for each item, then the conditional probability to detect it, given it has occurred, is:

$q_1 = 1 - \bar{M}(k)/N = 2n/N−1$, $n ≥ N/2$

Obviously, $q_1 = 1$ when $n = N$, and $q_1 = 0$ when $n = N/2$, or less.

$q_1$ would be a measure for the quality of control in a pessimistic evaluation.

From an optimistic point of view, the minimum number of items that are not controlled if n figures are presented to the manager should be determined from an optimistic point of view. Denote this number by $M(k)$. Table 3 describes its nature.

For $1 ≤ k ≤ N−1$ we get

$M(k) = k + 1 = N−n + 1$. 
Assessing the Value of IS

Table 3. Optimistic Estimate of the Number of Data Items Garbled in Aggregation

<table>
<thead>
<tr>
<th>n</th>
<th>M(k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0</td>
</tr>
<tr>
<td>N - 1</td>
<td>2</td>
</tr>
<tr>
<td>N - 2</td>
<td>3</td>
</tr>
<tr>
<td>N - 3</td>
<td>4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>N - k</td>
<td>k + 1</td>
</tr>
</tbody>
</table>

Taking the above assumptions regarding the probability of exception the conditional probability of detection is derived:

\[ q_2 = 1 - \frac{M(k)}{N} = \frac{(n-1)}{N} \text{, for } 1 \leq n \leq N - 1. \]

For example, if \( n = 2 \), then \( q_2 = 1/N \), meaning that \( N - 1 \) figures have been aggregated and only one original figure remained unchanged.

Regarding the detection probability as a representative of the quality we may conclude that the quality is limited between \( 2n/N - 1 \) and \( (n-1)/N \), where \( N/2 < n < N - 1 \). The quality is limited between zero and \( (n-1)/N \) where \( 1 < n < N/2 \).

Since the pessimistic and optimistic boundaries of the quality function are both linear in \( n \), we may say that the quality function, \( q(n) \), is a linear function of \( n \), and write it as \( q(n) = a_1 n + a_2 \), where \( a_1 \) and \( a_2 \) can be estimated as shown above.

Regarding the human capacity to perceive and analyze data [34], it seems that any utility function representing that characteristic should have a single maximum value in the range between \( n = 1 \) and \( n = N \). Recall that the number of possible comparisons is proportional to \( n^2 \). A quadratic function might serve though not necessarily, for that purpose: \( w(n) = -b_1 n^2 + b_2 n + b_3 \). If \( n_0 \) is known to be the optimal number of \( n \) from cognitive psychology viewpoint, then the ratio, \( b_2/2b_1 = n_0 \), may serve as a constraint for the equation of \( w(n) \).

Finally, in order to constitute a joint utility function, one has to calibrate \( q(n) \) and \( w(n) \) in order to obtain monetary values and to inquire about the tradeoffs among them. The cost of producing and reporting \( n \) items out of \( N \) has to be assessed as a function of \( n \). We may assume that the cost is linear to \( n \), since the data are more aggregated, less space is required for their storage, fewer retrieval operations are performed, yet more calculations are executed, and hence, arrive at a quadratic single attribute utility function of \( n \) with few constraints, whose solution should lead to an optimal \( n \).

The preceding example is certainly simplistic, yet it exhibits an approach to handle aggregation. We return now to the general discussion on the attributes of content.

The content attributes that have been presented do not pretend to be exhaustive, nor mutually independent. Their tradeoffs and overlapping domains should be investigated while the joint utility function is assessed, as discussed in the next major section. They were brought in order to demonstrate an approach toward content evaluation.

Format

Many possible attributes exist with regard to the format of a report. Among others, one may consider the following:

**Medium**—the medium by which the report is provided, such as a printed report, visual display, plotter or graphs, or microfilm, etc. [13, 51].

**Ordering**—the way the data is arranged in the report, such as the order of the columns or horizontal ordering, or the sequencing of details and totals, or vertical ordering, etc.

**Graphic Design**—the graphic setting of the report, such as colors, letter sets and font, etc. [13].

The main problem regarding the evaluation of the format related attributes is that quantitative variables cannot be obtained, hence the use of an analytical approach is almost impossible. Two methods can be used for the assessment of the value as perceived by the user. First, since in reality the number of alternatives here is very small, the user may be asked to rank order
Assessing the Value of IS

Second, one may use some of those attributes, which are no more than technological characteristics, to set constraints for the other categories of attributes. For example, the usage of a line printer sets a constraint on response time. The vertical order of the printed lines might set constraints on feasible aggregations, if sorting is not desired and the files are sequential.

The above discussion presents a somewhat arbitrary list of attributes related to the assessment of the value of a reporting system. Its main purpose is to demonstrate an approach, rather than to set an ultimate attribute list.

The Joint Utility Function

The ideal situation, from the evaluator point of view, would be when every relevant attribute is known and measurable; the utility function related to each individual attribute is clearly defined; and the tradeoffs among the various attributes are available and provide a clear mathematical formulation of the joint utility function, which also includes the cost function. In this case, an adequate optimization technique will lead to an optimal solution for a reporting system. An ideal situation like that is illustrated in the following example.

Example [2]. Suppose a user wishes to install a semiheuristic inquiry system. The system should provide information about individual customers, but the identification of the customer is not always unique. Sometimes only the name of the customer is available and sometimes the name is not even correctly spelled. There are various software packages performing such types of retrieval. The packages differ in their reliability or similarity measure, response time, and cost.

Assume the user utility function regarding the similarity measure, $s$, is found to be $U_s(s) = -a_2 \exp(-b_2s)$, where $a_2, b_2 > 0$ and $a_2$ serves as a calibration coefficient which transforms the value of $U_s$ into monetary figures. $s$ itself is based on the probabilities of errors of the first and second type described above, e.g., $s = 1 - p_1 - p_2$ where $p_1$ is the probability that an existing desired record is not retrieved, and $p_2$ is the probability that the record retrieved is not the one desired.

The user utility function regarding the response time $t$ is $U_t(t) = a_1 \exp(-b_1t)$, where $a_1, b_1 > 0$, and $a_1$ calibrates the function to provide monetary values.

Now suppose the cost is increasing linearly with regard to the increasing of $s$ and the decreasing of $t$, namely:

$$c = -a_3t + b_3s + \text{constant} \quad (a_3, b_3 > 0)$$

If the joint utility function is found to be additive with regard to $U_t$, $U_s$, and $c$, and if some coefficients can be calibrated, a utility function of the type is achieved:

$$u = a_1 \exp(-b_1t) - a_2 \exp(-b_2s) - (-a_3t + b_3s) + \text{constant}$$

The solution to this function under some constraints for $t$ and $s$, at least $t > 0$ and $-1 \leq s \leq 1$, might give some insight to the optimal system for the organization involved. After the theoretical equation (2) has been solved, one may look into the actual values of $t$ and $s$ characterizing each of the proposed systems, and select a system whose attributes are the closest to the values obtained in the optimal solution.

The prospects of facing an ideal case in reality are poor. More often “quick and dirty” approaches are needed which would be less sophisticated but sufficiently convincing, and above all, make sense. The following discusses and demonstrates such methods.

Suppose that constructing a joint utility function is not feasible. Moreover, even the marginal utility function for each individual attribute is not available. The user can only rank order the possible values of attributes, assuming that they are measurable, according to individual preferences. In that case, one of the following evaluation methods is recommended.

Lexicographical ordering

This approach is useful if the value attributes can be rank ordered according to their importance, and tradeoffs are not accepted among them, i.e., the user is not willing to trade a possible decrease
of a certain attribute for an increase in another [24, pp. 77-79]. For example, suppose the response time, \( t \), is the most important attribute. Hence, if \( t_1 > t_2 \) then system 2 is selected regardless of the value of any other attribute, assuming that the user prefers a shorter response time. If \( t_1 = t_2 \) then the next attribute is examined and so on. This method should be used if there is a clear dominance of importance among the various attributes. It can be refined by bounding attributes to be between some given constraints. However, cases in which such dominance exists are quite rare.

**The efficient frontier**

**Definition.** System \( j \) dominates a system \( k \) if at least one of its attribute's value is better than the corresponding value of system \( k \) and the other attribute's values of \( j \) are not worse than the corresponding values of \( k \) [24, pp. 69-77]. The set of all the systems that are not dominated by any other system, cost is considered as well, will be called the "efficient frontier." Each proposed system which does not belong to the efficient frontier can be eliminated. However, the problem remaining is how to select a system out of all members of the efficient frontier. Keeney and Raiffa [24] suggest some heuristic ways such as setting constraints to all the attributes less one, and selecting the best system with regard to the remaining attribute, as long as the constraints are satisfied.

A further refinement suggested by them would be to weight each attribute and select the system with the best "weighted average." Virtually, this is the well known method of weighting [43] which assumes the existence of a linear utility function and linear tradeoffs; it is therefore often criticized.

Some of the above ideas have been used in order to select a minicomputer for a small business. As happens very often when theoretical concepts are applied to a real case, the evaluator has to tolerate some compromises. For example, in this case the price of the equipment considered ranged around \$40,000, hence the selection process could not be too costly in comparison to the above figure. This prohibited the use of sophisticated analysis of a utility function, rather than linear weighting. The analysis also had to comply with a certain degree of understanding mastered by the business management. Still, the case shows that the aforementioned ideas are applicable and can be used to solve real problems.

**Case [3].** A regional distributor, or franchise, of an international company that produces cameras, films, x-ray machines, developing chemicals, etc., had some severe problems in managing and controlling the developing laboratory. Too many customer orders were lost, delayed, mishandled, or mistakenly mailed. Sudden shortages in developing chemicals caused delays in lab operations. The management had a vague feeling that computerized data processing would improve its operations and its planning. A feasibility study was completed, recommending the installation of a minicomputer to handle the lab data processing and the bookkeeping of the entire firm.

The feasibility report, with a cover letter, was sent to various manufacturers of minicomputers. A mandatory requirement was that a turnkey system should be proposed, including: hardware, software, operating procedures, and assistance in implementation. The possibility of smooth future upgrading was highly stressed. Six proposals were received, and they were analyzed and compared as shown in Table 4.

It is important to note that there was not any attempt to assign absolute "grades" to the proposals' attributes. The entire evaluation was based on relative comparisons among the various proposals.

In the next step, the number of appearances of "S," "A," and "I" were counted for each proposal, resulting in the following as shown in Table 5.

The "linearity" approach was then used to rank order the various proposals, by associating three points for "S," two points for "A," and one point for "I." Another rank ordering was performed for the purchase price. Both orderings are presented in Figure 2.

In other words, the price ordering from the bottom up was D, F, E, B, A, and C. The quality ordering from the bottom up was F, D and E, C, A, and B.

The evaluation rule was applied and proposals D and B were found to dominate all others; all other proposals are inferior either to B or to D. Thus, it
Table 4. Relative Ranking of Proposals’ Attributes*

<table>
<thead>
<tr>
<th>Attribute</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>hardware</td>
<td>S</td>
<td>S</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>I</td>
</tr>
<tr>
<td>upgrading ceiling</td>
<td>S</td>
<td>S</td>
<td>A</td>
<td>A</td>
<td>S</td>
<td>I</td>
</tr>
<tr>
<td>modularity of upgrading</td>
<td>S</td>
<td>S</td>
<td>I</td>
<td>A</td>
<td>A</td>
<td>I</td>
</tr>
<tr>
<td>software</td>
<td>A</td>
<td>S</td>
<td>S</td>
<td>A</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>similar experience of manufacturer</td>
<td>A</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>backup</td>
<td>I</td>
<td>I</td>
<td>A</td>
<td>I</td>
<td>S</td>
<td>S</td>
</tr>
</tbody>
</table>

*S - superior; A - average; I - inferior

Table 5. Totals of Relative Grades

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Number of “S”</th>
<th>Number of “A”</th>
<th>Number of “I”</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 2. Efficient Frontier
was recommended that proposals D and B be investigated further. At this point some members of management and user departments, i.e., lab and bookkeeping, were taken to live demonstrations of systems B and D and began negotiations with manufacturers B and D. The price difference between B and D was around $20,000, $25,000 versus $45,000, and what was actually being investigated was whether system B was worth the difference. Eventually, proposal B was accepted. The system is now implemented and the user is quite satisfied, mainly because management understood and was involved in the selection process, though its knowledge in DP was limited.

Conclusions

The issue of information systems evaluation suffers from scarce and scattered theoretical background as well as severe technical problems of measurement. However, it seems that if the measurement problems are at least partly overcome, there are sufficient methods of evaluation.

This article attempts to contribute to the solution of both problems. It condenses and classifies some theoretical background for information systems evaluation. It also suggests a systematic approach toward applying this theory. Some examples were given to demonstrate that approach.

The discussion here could not be exhaustive. Hopefully, it will lead to further research in that area.

References

[16] Feltham, G.A. Information Evaluation,


[34] Miller, G.A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Review, Volume 63, Number 2, March 1956, pp. 81-97.


Assessing the Value of IS


About the Author

Niv Ahituv is an Assistant Professor at the Faculty of Commerce and Business Administration, at The University of British Columbia. He formerly lectured at The University of Calgary and at Tel-Aviv University, and managed the Data Processing department at the Bank of Israel. He holds a B.Sc. in Mathematics, an M.B.A., and M.Sc., and a Ph.D. in Information Systems. His articles have appeared in Computers and Operation Research, The Computer Journal, Information and Management, and others. His main areas of interest are Economics of Computers, Information Economics, and Information Systems Management and Development.