# Using Value-at-Risk for IS/IT Project and Portfolio Appraisal and Risk Management

**Stefan Koch**
**Department of Information Business, Vienna University of Economics and BA, Austria**
stefan.koch@wu-wien.ac.at

**Abstract:** This paper makes the case for adopting a risk measure from the finance sector for IS/IT project and portfolio evaluation. The proposed value-at-risk approach constitutes a well-tested approach in high-risk environments, especially banking, and reports the expected maximum loss (or worst loss) over a target horizon within a given confidence interval. Value-at-risk is computed using either an analytical, parametric approach, or resorting to simulation, either based on historical samples or Monte Carlo methods. The main advantages of using value-at-risk measures are that they are methodologically consistent with modern IS/IT evaluation approaches like real options, that they offer possibilities for management and assessment of IS/IT project portfolios, and that the results are easy to interpret.

**Keywords:** IT investment, risk management, value-at-risk, project portfolio

## 1. Introduction

In the last years, the evaluation of IS/IT projects has been the centre of much debate. One of the reasons for this debate has been the e-Commerce and Internet hype on the trading floors, and the fact that the respective bubble later on exploded. Naturally, investments into new technology, especially IS/IT, and respective start-ups need to be carefully analysed, especially in this new environment. Associated with this trend, risk management, either within an organisation performing one or multiple IS/IT projects or for an investment in several start-ups has become a center of attention (Remenyi 1999, Benaroch 2002, DeMarco and Lister 2003).

Regarding valuation of IS/IT projects, the real options approach (Trigeorgis 1998) gained prominence in MIS literature (Santos 1991, Benaroch and Kauffman 1999, Taudes 1998). This approach is based on option theory from finance, and tries to incorporate the management's flexibility into decision making. Especially several possible options like abandonment, or expansion (growth) options offered by pilot projects are of interest in IS/IT projects. In the literature, several applications for real options have been described, including software growth options used in evaluating software platform decisions (Taudes 1998, Taudes et al. 2000), or investment timing in the development of point-of-sale (POS) debit services (Benaroch and Kauffman 1999, Benaroch and Kauffman 2000). In the last years, the focus has shifted from evaluating one (or more) known options embedded in an IS/IT project towards active management and planning of options in IT investments for controlling risks (Benaroch 2002).

Extending the approach of using analogies with finance, this paper argues for adopting a value-at-risk approach in evaluating IS/IT projects and for risk management. The main advantages of using value-at-risk measures are that they are methodologically consistent with modern IS/IT evaluation approaches like real options, constitute a tested and used approach in high-risk environments, especially banking, that they offer possibilities for management and assessment of IS/IT project portfolios including existing dependencies, and that the results are easy to interpret.

The structure of this paper is as follows: First, an introduction to value-at-risk will be given, highlighting both shortly its history in the finance sector and the main points of the computation itself. Then, the application for evaluating a single IS/IT project will be discussed, afterwards detailing the use for IS/IT project portfolio risk management. In both cases, small illustrative examples are given and discussed.

## 2. Introduction to value-at-risk

### 2.1 History and applications

The history of value-at-risk is deeply interwoven with the finance sector and especially banking. In the strive for financial stability, a first landmark decision was the 1988 Basle accord by the central banks from the G-10 countries, which defined a minimum standard of capital requirements for commercial banks, using a percentage of risk-weighted assets (Basel Committee on Banking Supervision 1988). As this first approach has faced criticism, including that neither portfolio risk, nor netting, nor market risk have been accounted for, modifications have become necessary. In 1993, one of the most important documents, the Group of Thirty's report on derivatives was published, explicitly endorsing value-at-risk for



Reference this paper as:
Koch S (2006) "Using Value-at-Risk for IS/IT Project and Portfolio Appraisal and Risk Management" *The Electronic Journal Information Systems Evaluation* Volume 9 Issue 1, pp 1-6, available online at www.ejise.com

measuring market risk (Group of Thirty 1993). This concept was then popularized by the RiskMetrics system originally developed by J.P. Morgan (Morgan Guaranty Trust Company 1994). The Basle accord, after an amendment for market risk in 1996 (Basel Committee on Banking Supervision 1996), in its latest version from 2001 now also 'strongly recommends' that banks disclose their value-at-risk. The U.S. Securities and Exchange Commission (SEC) also now requires all large U.S. publicly traded corporations to report quantitative data on market risk in their report to the SEC, listing value-at-risk as one of three possible methods for doing so (SEC 1997, Jorion 2001, Jorion 2002). Recently, it has been empirically shown that value-at-risk disclosures of banks are significantly related to future market risk (Jorion 2002).

In the last years, applications of value-at-risk measures have started to begin in areas other than finance, including inventory management (Luciano et al. 2003), the purchasing process (Sanders and Manfredo 2002) or even real estate investement (Kevenides 2002).

## 2.2 Computing value-at-risk

While several definitions for value-at-risk can be formulated, it basically indicates the greatest potential loss of a position or a portfolio, which can be verified with a certain probability, in a defined time horizon (Tardivo 2002, Best 1998). Or, as Jorion puts it, value-at-risk summarizes the expected maximum loss (or worst loss) over a target horizon within a given confidence interval (Jorion 2001). These definitions already hint at several important characteristics of value-at-risk: It can be computed both for a single position or for a diversified portfolio, and it has some discretionary power, in that both the holding period (time horizon, target horizon) and the confidence interval need to be defined by the user. The holding period should be set with the type of portfolio considered taken into account, setting a horizon corresponding to the period necessary for orderly liquidation (Jorion 2001). For example, a bank computing their value-at-risk for a portfolio of highly liquid currencies might even use one day as holding period. The confidence interval chosen should necessarily either reflect regulatory imperatives, risk attitude, or depend on characteristics of the underlying distribution. Having set both holding period and confidence interval, value-at-risk is computed by estimating the probability distribution of gains and losses of the considered position or portfolio over the time horizon, and then finding the point at which the probability of incurring greater losses corresponds to the set confidence interval (in fact to one minus the confidence interval). Therefore, value-at-risk reports one, easy to interpret figure: The loss of money that is not exceeded at the probability of the confidence interval over the defined time horizon. In the most general form, value-at-risk can therefore be derived from the distribution of the future portfolio value $f(w)$, finding for a given confidence interval $c$ the worst possible realization $W^*$ such that:

$$c = \int_{W^*}^{\infty} f_\tau(w)\, dw. \quad (1)$$

The value-at-risk can be either reported relative to the mean (the expected portfolio value) or as absolute loss relative to zero.

For computing value-at-risk in practice, three approaches are proposed, each with specific strengths and weaknesses. These are the parametric, or analytical or variance-covariance approach, historical and Monte Carlo simulation (Pearson and Smithson 2002, Stambaugh 1996). Sometimes the latter two are grouped together under the name of simulation or full valuation methods (Best 1998, Jorion 2001, Tardivo 2002). While historical simulation necessitates large historical samples (and attendant assumption of stable volatilities), Monte Carlo simulation naturally can become complex and costly in computer resource for large real-world portfolios (although in the last years several ways were proposed to increase the speed of Monte Carlo simulations (Pearson and Smithson 2002)). Most often used, due to being the first version having been developed, ease of implementation and conceptual fit with modern portfolio theory, is the parametric approach.

The main hypothesis behind the parametric approach is that the future portfolio values (and hence returns) follow a parametric distribution, the most common assumption is that they follow a normal distribution. Therefore, value-at-risk can be derived directly from portfolio standard deviation $\sigma$ (using a multiplicative factor $\alpha$ dependent on the chosen confidence level). For a single position with initial investment $W_0$, the value-at-risk below the mean then is given by:

$$VAR(mean) = W_0 \alpha \sigma \sqrt{\Delta t}. \quad (2)$$

For a portfolio of assets, as the return of each single asset is assumed to be normally distributed, the portfolio return as a linear combination of normal variables is necessarily normally distributed as well. Due to the diversifying effects of a portfolio, the value-at-risk

 

of a portfolio is not the sum of the value-at-risks of all single positions, but needs to incorporate the respective covariance matrix. The delta-normal method defines relations between financial positions and underlying, primitive risk factors which again are normally distributed. For an instrument whose value depends on a single underlying risk factor $S$, first the portfolio value at the initial point is computed, together with the first partial derivative $\Delta_0$ with respect to the underlying risk factor *S*, the sensitivity of value to changes in the risk factor at the current position, termed delta for derivatives. The potential loss in value $dV$ is then computed as

$$dV = \Delta_0 \times dS, \qquad (3)$$

using the potential change $dS$ in the underlying risk factor. If the distribution is normal, the value-at-risk can be derived from the product of the exposure and the value-at-risk of the underlying variable:

$$VAR = |\Delta_0| \times VAR_S = |\Delta_0| \times (\alpha \sigma S_0). \qquad (4)$$

For a portfolio, the delta-normal method uses a set of primitive risk factors, onto which the positions are mapped using the respective delta-positions denoted by vector $\chi$, and the covariance matrix $\Sigma$ between risk factors over the target horizon together with the specified confidence level to compute the portfolio value-at-risk:

$$VAR = \alpha \sqrt{\chi' \Sigma \chi}. \qquad (5)$$

Especially with derivatives like options, due to their non-linear nature, including the second derivative using delta-gamma approximation is recommended to increase the fit.

For a more thorough treatment of value-at-risk than is possible here, the works of Jorion (2001), Best (1998), Pearson (2002) and Allen, Boudoukh and Saunders (2004) are useful starting points.

## 3. Value-at-risk for IS/IT project evaluation

The first and most important question is whether the value-at-risk can in general be determined for an IS/IT project. Following the most generic terms and definition of value-at-risk, it can naturally be derived. Every IS/IT project has a certain amount of uncertainty, therefore a probability distribution of gains and losses over a set time horizon exists. Necessarily, any arbitrary confidence level can thus be set, and the cutoff point in the probability distribution specifying the loss not exceeded with corresponding probability can be determined.

Before specific problems of computation, uses and advantages and disadvantages are addressed, specification of both confidence level and time horizon in the context of IS/IT projects need to be discussed. While the confidence level can be determined quite analogous with classic value-at-risk, e.g. using 95% or 99%, but keeping in mind possible characteristics of the underlying distribution, the time horizon needs to be more carefully evaluated. Depending on the reason for project evaluation, the holding period should be set accordingly. In finance, the holding period could correspond to the time period necessary for orderly liquidation of the asset considered. For IS/IT projects, liquidation is most often available by stopping a project, which is normally possible at short notice or immediately. Due to the fact that IS/IT projects (normally) are not traded assets, this would mean exercising an abandonment real option, forfeiting any further benefits but also costs. While this analogy would lead to assume very short holding periods, the volatility of an IS/IT project's gains and losses over short periods of time will be small. Therefore longer holding periods should be considered in the context of IS/IT projects. If a single project is considered, the holding period could even be set to the assumed project length. For application within a larger organization performing several concurrent projects, evaluation of a start-up portfolio or similar as will be detailed in the next section, the holding period should necessarily be reduced to be in the area of one or several months, maybe a quarter.

For illustrative purposes, a first simple project will be considered. This project will, over its projected length of one year, necessitate costs of about 100 monetary units (MU), and is projected to generate positive cash flows of 140 MU with probability $p_1 = 0.4$, of 120 MU with probability $p_2 = 0.2$, of 100 MU with probability $p_3 = 0.2$, of 80 MU with probability $p_4 = 0.1$ and of 0 MU with probability $p_5 = 0.1$. No embedded options are considered at this stage. The resulting probability distribution for project value after one year therefore is discrete and is easily constructed. Setting a confidence level of 95% allows to easily determine the cutoff point in this distribution, leading to an absolute value-at-risk below zero of 100 MU, or a relative value-at-risk to the mean $E(P) = 8$ of 108 MU. While this seems straightforward and trivial in this simple case, stating these figures already offers additional information regarding risk for the

project, and might serve as an important complement to reporting only mean project value, or a measure like discounted cash flows.

Next, we will consider the case of a software growth option, implementing a web-based e-commerce system, embedded into a platform change from SAP R/2 to SAP R/3. All data for this option are taken from Taudes, Feurstein and Mild (2000). They give the spot price $S_0$ with 880,000 MU and volatiliy $\sigma=0.8$. Valuation of this American call option using the Black-Scholes formula gives 514,000 MU, with a delta of 0.7756. Using delta-normal valuation and 95% confidence level (corresponding to $\alpha=1.645$) in equation (4) results in a value-at-risk of 898,207 MU.

In the evaluation of single projects, value-at-risk measures can be computed both at the beginning (normally using project length as holding period), and also during the project for continuous monitoring. At the point of an investment decision at project start, value-at-risk measures allow for easy to understand, monetary quantification of associated risks, and therefore offer a good complement for other measures like net present value. On the downside, computing value-at-risk is either trivial and therefore offers little additional information, necessitates strong assumptions like normal distributions or gets complex if Monte Carlo simulation is employed. Using historical simulation will be mostly problematic due to missing large historical samples.

## 4. Value-at-risk for IS/IT project portfolios and risk management

There are numerous examples for when an IS/IT project portfolio needs to be evaluated regarding the contained risks. These include the classic case of a large software developing organisation that performs several projects. In that case, overall risk assessment is of high interest, especially if a diversification effect is in place or is strived for. The next possible application is for evaluating a portfolio of IS/IT start-ups, as held or being built by an investor. While this is more akin to a classical finance application, start-ups in this area can also be seen as IS/IT projects.

The last, and maybe the most often occurring possibility is a portfolio of an IS/IT project with several embedded options. In that case, an assessment of underlying risk factors is necessary. If only a portfolio of a project and an embedded option for example to defer investment, priced as an American call on the gross present value of the completed project (Trigeorgis 1998) is considered, there is only one underlying risk factor, project value, which eliminates diversification effects and reduces the associated covariance matrix $\Sigma$ in the delta normal method to a scalar, the risk factor's variance $\sigma$, with a vector $\chi$ of two delta-positions describing the exposure of both positions, project and option, to the risk factor (see equation 5). On the other hand, options on a different underlying asset, thus maybe depending on one or more other primitive risk factors, embedded in a project would necessarily lead to assessing the risk of a portfolio composed of one project and one or more options. In this case, diversification might be present, and needs to be included in the computation of the portfolio value-at-risk. In the second example given in the last section, a growth option for implementing a web-based e-commerce system was evaluated according to its value-at-risk on its own. As this option was embedded into a platform change from SAP R/2 to SAP R/3 together with four others, with these option values leading to a positive expanded (strategic) net present value of the platform change (Taudes et al. 2000), the whole portfolio of platform project and real options needs to be evaluated together. Simply evaluating each component separately and summing the resulting value-at-risks would negate any benefits from diversification. While two of the options implement EDI-based solutions, the others including the e-commerce system and the main platform project would be exposed and mapped to different risk factors.

For illustration, we will now expand on our treatment of the option presented above, complemented with the main platform project. Again, data are taken from Taudes, Feurstein and Mild (2000), although a volatility for the main project of $\sigma_{project}=0.2$ is introduced. Data for the web-based e-commerce system remain unchanged from last section. Furthermore, we presume the presence of two risk factors, with each position exposed to one of them, the option according to delta-normal method with delta 0.7756, the platform project with its full value at -416,500 MU. Lastly, a correlation of 0.3 is assumed between the risk factors. Using equation (5) at confidence level 95% corresponding to $\alpha=1.645$ gives

 

$$VAR_{div} = 1.645\sqrt{\begin{bmatrix} -416{,}500 & 0.7756\times 880{,}000 \end{bmatrix}\begin{bmatrix} 0.2^2 & 0.3^2 \\ 0.3^2 & 0.8^2 \end{bmatrix}\begin{bmatrix} -416{,}500 \\ 0.7756\times 880{,}000 \end{bmatrix}}$$
$$= 828{,}907.$$

The portfolio value-at-risk therefore is 828,907 MU, due to diversification smaller than the sum of individual value-at-risks (the undiversified value-at-risk) of

$$VAR_{undiv} = VAR_{project} + VAR_{option}$$
$$= \left(1.645 \times 0.2 \times \left|-416{,}500\right|\right) + 898{,}207$$
$$= 137{,}028 + 898{,}207 = 1{,}035{,}235.$$

In analysing portfolio value-at-risk, the change in value-at-risk due to addition of a new position can also be computed, termed incremental value-at-risk, as well as component value-at-risk, giving the reduction of the portfolio value-at-risk resulting from removal of a position. Due to diversification, both measures would in most cases be different than the individual value-at-risk of the position. This allows for in-depth analysis of components in a portfolio, or could even be used as a constraint for portfolio optimization (Yiu 2004).

One main point to consider when using value-at-risk to evaluate an IS/IT project and/or option portfolio is which primitive risk factors to choose, and how to map the positions to them, if the delta-normal method is to be applied. A survey of literature yields several risk factors commonly associated with IS/IT projects, including technological and organizational risk (Taudes et al. 2000). The most complete taxonomy to be found is by Benaroch, who distinguishes between firm-specific risks, including monetary, project, functionality and organizational risk, competitive risks and market risks including environmental, systemic and technological risk, and argues for real option analysis to assist in risk management by deliberately embedding options to address the various risks and thus optimally configuring the investment (Benaroch 2002). This line of research shows distinctive relationship with the value-at-risk approach argued for in this paper, with value-at-risk offering a way of quantifying risk reduction afforded by embedding certain options into the investment portfolio.

## 5. Conclusion

This paper has argued for adopting the value-at-risk approach in the evaluation of single IS/IT projects and also portfolios constructed from these projects and/or related real options. As has been detailed, value-at-risk is a common and accepted measure in the finance sector, and offers several advantages also in the area of IS/IT projects. While several approaches for computing value-at-risk exist, not all of these might be applicable for IS/IT projects, as large historical samples will mostly be absent. On the other hand, both Monte Carlo simulation and an analytical approach seem feasible.

Using small, illustrative examples, we have shown that value-at-risk can indeed offer additional information in evaluating single IS/IT projects or real options on such projects, offering an easy to interpret way of quantifying and comparing associated risks, and especially in evaluating IS/IT project and/or option portfolios, as this method explicitly accounts for diversification effects. In addition, the changes in risk due to changes in the portfolio, both from eliminating and adding new elements, can easily be determined, making value-at-risk a useful tool for risk management, complementing and extending the real options approach.

If value-at-risk is indeed adopted, many further enhancements are possible, including the introduction of risk adjusted performance evaluation of business units or project managers, using profit over value-at-risk for assessment. Naturally, many further issues still need to be investigated in the context of value-at-risk for IS/IT projects, especially the definition of primitive risk factors, the mapping of positions to these and others. Nevertheless, adopting value-at-risk might provide important additional information for IS/IT decision makers, and might constitute a necessary step towards IS/IT risk management.

## References

Allen, L., J. Boudoukh, and A. Saunders (2004) Understanding Market, Credit, and Operational Risk: The Value at Risk Approach, Oxford: Blackwell Publishing.

Basel Committee on Banking Supervision (1988) International Convergence of Capital Measurement and Capital Standards, Basel: BIS.

Basel Committee on Banking Supervision (1996) Amendment to the Basel Capital Accord to Incorporate Market Risk, Basel: BIS.

Basel Committee on Banking Supervision (2001) Consultative Document: Pillar 3 (Market Discipline), Basel: BIS.

Benaroch, M. (2002) “Managing information technology investment risk: A real options perspective”, Journal of Management Information Systems, 19(2), pp. 43–84.
Benaroch, M. and R.J. Kauffman (1999) “A case for using option pricing analysis to evaluate information technology project investments”, Information Systems Research, 10(1), pp. 70–86.
Benaroch, M. and R.J. Kauffman (2000) “Justifying electronic banking network expansion using real options analysis”, MIS Quarterly, 24(2), pp. 197–225.
Best, P. (1998) Implementing Value at Risk, Chichester: John Wiley & Sons.
Black, F. and M.S. Scholes (1973) “The pricing of options and corporate liabilities”, Journal of Political Economy, 81, pp. 637–659.
DeMarco, T. and T. Lister (2003) Waltzing With Bears: Managing Risk on Software Projects, New York: Dorset House.
Group of Thirty (1993) Derivatives: Practices and Principles, Group of Thirty.
Jorion, P. (2001) Value at risk: The new benchmark for controlling market risk, 2nd Edition, Boston: McGraw-Hill.
Jorion, P. (2002) “How informative are value-at-risk disclosures?”, The Accounting Review, 77(4), pp. 911–931.
Kevenides, H. A. (2002) “International real estate investment risk analysis”, Journal of Real Estate Portfolio Management, 8(4), pp. 61–73.
Luciano, E., L. Peccati, and D.M. Cifarelli (2003) “VaR as a risk measure for multiperiod static inventory models”, International Journal of Production Economics, 81-82, pp. 375–384.
Morgan Guaranty Trust Company (1994) RiskMetrics™ technical document, New York: RiskMetrics Group.
Pearson, N. D. (2002) Risk budgeting: Portfolio Problem Solving with Value-at-Risk, New York: John Wiley & Sons.
Pearson, N. D. and C. Smithson (2002) “VaR: The state of play”, Review of Financial Economics, 11, pp. 175–189.
Remenyi, D. (1999) Stop IT Project Failures through Risk Management, Oxford: Butterworth-Heinemann.
Sanders, D. R. and M.R. Manfredo (2002) “The role of value-at-risk in purchasing: An application to the foodservice industry”, Journal of Supply Chain Management, 38(2), pp. 38–45.
Santos, B. L. D. (1991) “Justifying investments in new information technologies”, Journal of Management Information Systems, 7(4), pp. 71–90.
Securities and Exchange Commission (SEC) (1997) Disclosure of Accounting Policies for Derivative Financial Instruments and Derivative Commodity Instruments and Disclosure of Quantitative and Qualitative Information about Market Risk Inherent in Derivative Financial Instruments, Release No. 33-7368, FRR No. 48, Washington: Government Printing Office.
Stambaugh, F. (1996) “Risk and value at risk”, European Management Journal, 14(6), pp. 612–621.
Tardivo, G. (2002) “Value at risk (VaR): The new benchmark for managing market risk”, Journal of Financial Management and Analysis, 15(1), pp. 16–26.
Taudes, A. (1998) “Software growth options”, Journal of Management Information Systems, 15(1), pp. 165–186.
Taudes, A., M. Feurstein, and A. Mild (2000) „Options analysis of software platform decisions: A case study”, MIS Quarterly, 24(2), pp. 227–243.
Trigeorgis, L. (1998). Real Options - Managerial Flexibility and Strategy in Resource Allocation, 3rd edition, Cambridge, Massachusetts: The MIT Press.
Yiu, K. (2004) “Optimal portfolios under a value-at-risk constraint”, Journal of Economic Dynamics and Control, 28, pp. 1317–1334.

 

# A Process Capability Approach to Information Systems Effectiveness Evaluation

**Sevgi Özkan**
**Informatics Institute, Middle East Technical University, Ankara Turkey**
sozkan@ii.metu.edu.tr

**Abstract:** While defining or measuring the effectiveness of the information systems (IS) function has proven complicated, further effort on refining IS assessment is essential for the effective management and continuous improvement of both the IS function and the organisation. In addition, an effort to investigate the relationships among the established IS assessment tools to better reconcile their existing differences is warranted. This paper aims to clearly differentiate the notions of 'Software' from 'Information Systems'. A new IS assessment model is proposed to provide a more holistic view on how IS quality may be assessed by means of a process capability understanding of evaluating IS effectiveness within the organisational context.

**Keywords:** Information systems quality, Information systems effectiveness, Assessment, Software process maturity, Process capability.

## 1. Introduction

This paper aims to clearly differentiate the notions of Software from Information Systems. Within these two different entities, until today Software has attracted the attention on the subjects of quality and assessment of that quality it possesses. Information Systems on the other hand, were not considered as entities that need to be assessed in order their level of quality to be depicted. Moreover, as most researchers referenced in this work point out, considering Software away from its natural context and framework, which is the Organization, is a severe mistake. Software together with the organizational context that it lies within, construct Information Systems. Therefore, the quality of Information Systems is mainly related to the artifact used, the Software, but not less importantly to the organizational context that this artifact is being used within.

Within the Information Systems literature, there are many models proposed to assess the quality of Information Systems. One of the more complete and better known is DeLone and McLean's model of information systems success (2003). This model has been used as a basis for empirical research, and has been refined and extended by many researchers. In this article, two immature models are introduced and elaborated. As concluded, these conceptual models fail to serve as widely accepted models, as they are newly formed, with few experimental and theoretical support and even fewer comments by other researchers. Built on the discussions, a new process capability based model is developed and presented.

Information Systems (IS) and Software (SW) are definitely two different entities. It can be argued that IS are the aim, whereas Software serves for that aim, so that IS tend to be the organizational context where the developed Software is used. IS therefore are built upon Software, and the quality of Software effects, if not determine the quality of the IS. The quality characteristics of Software differ from those of the Information Systems of which it is part of. The quality of Software focuses on the quality of the production of the object or artifact as widely used by many researchers, but the quality of IS focuses on the use of this object or artifact within the organizational environment. The borderline between a Software and an IS may be clear if Software is limited to programs, and IS is seen to be the organizational framework and context in which Software is used. However, this definition may be considered to be to some extent insufficient, inadequate and poor for the discussion about Software quality, as it obviously restricts the consideration to the technical characteristics of Software and leaves out the usage of Software, and the way usersexperience Software, and influence their opinion about its related quality (Von Hellens, 1997).

The Capability Maturity Models (CMMs) have become a de facto standard for assessing and improving Software processes, a model for judging the maturity of the Software processes of an organization and for identifying the key practices that are required to increase the maturity of these processes. It is an interesting and fascinating approach to apply the CMMs on IS in order to assess and improve the overall quality of the processes embodied in the IS. However it is necessary to alter, adjust and modify the maturity levels and key process areas of SW-CMM (or the updated version CMM-Integrated) to better match the needs of a possible IS CMM.

 

Reference this paper as:
Ozkan S (2006) "A Process Capability Approach to Information Systems Effectiveness Evaluation" *The Electronic Journal of Information Systems Evaluation*, Vol. 9, Iss. 1, pp 7-14, available online at www.ejise.com

Definitely, Software quality is an empty statement without some indication of its performance and applicability in the user environment. Therefore, the quality of Software emphasizes the quality of the production of the artifact called Software in an organizational context. These are related closely and largely to the definition of quality, that quality is contingent and resides in the user's perception of the product (Siakas, et al., 1999).

## 2. Software quality versus Information Systems quality

An IS can be defined technically as a set of interrelated components that collect or retrieve, process, store and distribute information to support decision making, coordination, control, analysis, and visualization in an organization (Laudon and Laudon, 2001). IS are socio-technical systems (Mumford, et al., 1984) of which information technology is one significant feature. They can be thought of as integrating an infrastructure and the various systems, which make use of that infrastructure (Galliers, 1994). IS are meaningful only when they are considered within a context, and the main distinction between a Software system and an IS is that Software is limited to the development process of a Software system, while an IS is seen to be the organizational context in which Software is used (Von Hellens, 1997). If we accept this difference and distinction, then we can argue that Software quality means development process quality not considering the usage of that Software, while IS quality will emphasize product quality assessed by the usage of Software in an organizational background. Due to the multidisciplinary character of IS a discussion about the necessity of a societal viewpoint in these days of globalization of the Software market, virtual global enterprises and cross-cultural teams follows with emphasis on Software quality and process improvement (Siakas, et al., 1999).

There are a number of approaches developed during the 1990's for the achievement of IS quality, however no approach yet provides a solution that is detailed enough in either a scientific or practical sense (Dahlberg and Jarvinen, 1997). Some of the existing approaches that were analyzed and evaluated within the research framework of this paper, have proven that they are mainly limited because they concentrate too much on the technical and control-oriented aspects of IS Quality management.

The movement that began as Quality Control, which refers to the maintenance of a predetermined level of quality, has developed into the notion of continuous Quality Improvement (Dahlberg and Jarvinen, 1997). It can be seen that most Software developers have ISO 9001 certification, yet this alone does not necessarily mean better value for the customer and therefore there is definitely a need for a *multi-perspective* quality model that includes both scientific and practical IS/IT purposes, and is linked directly to the IS practice (Dahlberg and Jarvinen, 1997). The key word in the research of Dahlberg and Jarvinen is definitely the multi-perspective quality model notation, which denotes that IS can not be assessed efficiently and effectively by concentrating on a single perspective but only if all possible viewpoints and perspectives are employed. What these perspectives are or should be, are later discussed in the paper, with respect to the existing IS quality framework.

The concepts, models and measures that work in other fields might be usefully applied to the IS field, but careful analysis and consideration should be given. Furthermore, IS Quality research should be done at real user organizations, so that the results will be both useful and accepted by IS developers. The aim of this study overall is to assess the effectiveness and quality of the Informatics Institute Online Information System, METU, and definitely the results of such an assessment would have been very helpful in better understanding the quality aspects of IS within the organizational context.

## 3. IS quality within the organizational framework

It is important for the IS developers to recognize that they are primarily engaged in a service-oriented business, rather than being in the business of producing high-quality Software (Von Hellens, 1997). Based on this finding of Von Hellens, organizations using IS, should be aware that the artifacts that they are using are not only Software but a service, and they should be treated as services. Von Hellens continues that, in detail, it has been observed that in many IS development organizations there is either a *covert or overt reluctance* to participate with the user in the requirements elicitation, development or implementation of a new system (Von Hellens, 1997). The absence of user input into this process leads to a decline in use quality (Garvin, 1987) in which users are assumed to have differing needs, and the degree to which a product satisfies those needs determines the quality of the product. This is mainly related to the definitions of quality given by Weinberg and Crosby, as "Quality is conformance to requirements" and "Quality is conforming to some person's requirements" (Weinberg, 1992). The reluctance on the part of

the developer to have sufficient user input is often due to the developer having a different view (i.e. manufacturing-based view) of product quality in which quality is defined in terms of engineering and manufacturing practice—usually 'conformance to specification' (Von Hellens, 1997).

## 4. Models of IS quality

There are several difficulties of developing a comprehensive model of IS quality. It should be obvious that a definition of IS quality needs to be holistic, one which encompasses all relevant contexts—both technological and organizational—so that the framework to be constructed takes into account the different work contexts and specific organizational needs that should be considered when evaluating an IS quality. The IS quality model therefore takes a broader approach to evaluating quality than does a Software Quality approach which does not take into account the organizational context in which the Software operates. The difficulties of developing a comprehensive model of IS Quality arise from an incapability and inability to settle the varied perspectives taken by the IS stakeholders (management, developer, user) each with their own idea of quality (Andersson and Von Hellens, 1997).

The following subsections introduce two previously developed IS Quality models: the SOLE Quality model and the IS Quality Process View.

## 5. The SOLE quality model

In an effort to develop a useful, general-purpose definition of IS quality, Andersson and von Hellens have developed the concept of *IS work quality* by using the SOLE (Software Library Evolution) quality model, originally developed by Eriksson and Törn (1991), with consideration of the management functions required to assure evolution quality and user support quality (IS work practices) (Andersson and Von Hellens, 1997).

The SOLE quality model aims to create divisions of quality classes that are consistent with the decisions made by key decision-makers during the Software life-cycle. SOLE identifies three divisions; business quality, use quality, and IS work quality. At each level the listed quality factors relate to individuals who are primarily interested in the quality of the respective quality factor. Therefore, this serves for the previous important notation of 'multi-perspectiveness', and coincides with the basis IS model—DeLone and McLean's (1992, 2003). The following Figure 1 displays the relationship between the three divisions within the overall IS quality framework.

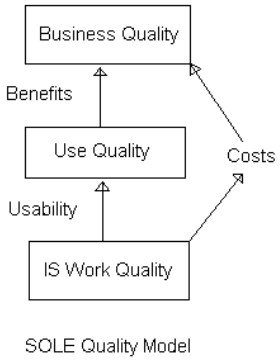


**Figure 1:** SOLE quality model (Eriksson & Törn, 1991)

The first part as shown in Figure 1, business quality, is the domain of senior and departmental management. It is seen in terms of costs and benefits. If benefits offset cost, then quality is considered to be good. Business quality refers to the quality of the activities that control the IS work all the way through the entire lifecycle of the product and service. The achievement and success criteria include whether deadlines and schedules are met with respect to the scheduled budget. It is the region and under the control of the Chief Information Officer (CIO) who is responsible for the efficient deployment of all types of assets, towards meeting the organization's previously settled information needs and goals. Therefore the project manager must keep continuing projects running efficiently and slickly with no unexpected shockers. Success criteria are the approval and acceptance of the deliverables (Andersson and Von Hellens, 1997).

The second part of the SOLE model, use quality**,** is defined by how well the system does what the user wants it to do. Users can be defined as the people who directly use the system by performing various work practices that prepare data and information for the system. The two main features of use quality in the SOLE model is what does the system do for the user, and how the interface is designed, otherwise called requirement quality and interface quality.

The third part of the SOLE model, IS work quality, is defined by the level of the performance of management, development, maintenance and operation of the IS, and the final products of Software and documentation. IS work quality considers the management tasks required to guarantee fruition and evolution quality and user support quality generally being concerned with all aspects of the way an IS serves the user.

 

Here it is important to emphasize that the three divisions within the SOLE IS Quality Model coincide with the DeLone and McLean's IS Success Model. 'Business quality' corresponds to the 'Net Benefits' *and* 'Service Quality' dimensions ; 'Use quality' to 'User satisfaction' *and* 'Information use' ; and 'IS work quality' to 'System quality' *and* 'Information quality' dimensions respectively.

The two dimensions of the SOLE model are referred to as the requirements dimension and the user interface dimension The requirement dimension includes such factors as "does the artifact meet the user's needs?", "is it secure?" and "is it easy to change according to new needs?". The user interface dimension is concerned with how easy the artifact is to use, whether effective help is available and that the artifact does not control the user's work performance (Andersson and Von Hellens, 1997).

Therefore based on these divisions and dimensions, the SOLE model shall be designed simply by a graph as:

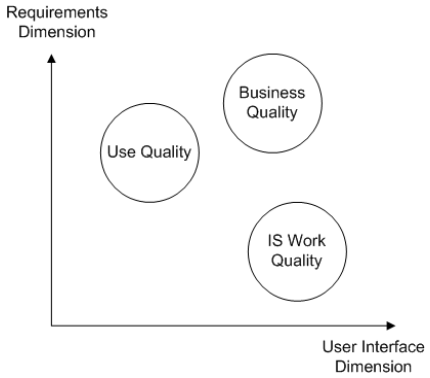


**Figure 2:** SOLE model

The most important voids of this model can be listed as:

- The model is primarily based on the Software Life Cycle. However the change over time (over the life cycle) can not be depicted or described without the notion of time. Therefore a third dimension can be considered to be the Time dimension.
- The model considers users as a single entity, however, users consists of different groups which view, understand, use and influence IS in different ways. Therefore, users can be considered as an alternative dimension for the model.

## 6. The IS quality process view by Olayele Adelakun

Another notion of quality as seen from some other point of view is that it is a multi-dimensional concept, which is context-dependent. It differs from the SOLE model and has somehow addressed the drawbacks of SOLE that have been addressed in the previous section. In this model, first of all, the quality of any IS can be said to be a function of three dependent variables (Adelakun, 1991). They are the type of IS, the Stakeholder group, and Time, as shown in Figure 3. The quality of any information system can be determined from this concept. Figure 3 below shows the quality model and how the three variables are related to the information systems quality.

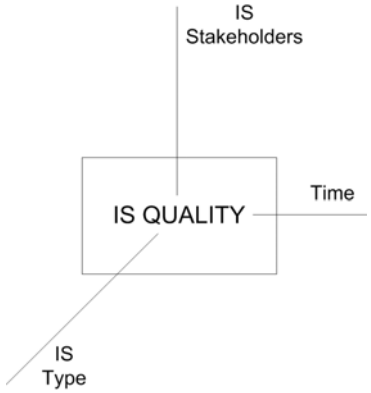


**Figure 3**: The quality model (Adelakun, 1991)

The IS quality process model was developed by Adelakun in order to operationalize the above framework. The three categories in the process model (IS planning, developers and users) cover the stakeholders' views in Figure 3. IS planning is a process of identifying IS that could be used to maintain and support a business strategy (Adelakun, 1991).

The business dimension of quality is defined as meeting or exceeding the stakeholders' expectations of the business benefits from the IS project (Adelakun, 1991). Examples of this business dimension of IS quality include: increasing productivity, improving customer services, reducing cost, compressing cycle time, and improving the accuracy of the delivery process.

It is even more important to communicate this business quality to the developer as the study by Adeyakun has proven that this activity is one of the most important activities in an IS project. It is not only essential to communicate this business advantage; it is also significant and critical to follow it up.

The IS development phase in Figure 4 usually focuses on the development of the Software object. In this phase of the IS project, as previously noted and justified, the quality focus is primarily technical. Many Software companies apply the Software section in the ISO 9001 quality standard or capability maturity model in the development of the Software product. These standards often view the IS development process as a linear one, starting with the functional requirements, programming, inspection and testing, delivery, and installation. Several researchers note that an unquestionable application of these standards in Software development can lead to serious development drawbacks. Many Software companies are realizing the limitation of these standards and are gradually moving away from them.

The last part of the IS quality process concentrates on the use which is primarily the user's view of quality, where concepts such as usability, quality-in-use etc. are applied here. Usually it is observed that the actual quality of the system can only be determined during its use (Adelakun, 1991).

## 7. The proposed PB-ISAM

Built on and inspired by the models that have been elaborated, a Process Based Information Systems Assessment Model (PB-ISAM), that would apply only to assessing and evaluating the quality of IS quality in terms of capability, has been proposed.

### 7.1 PB-ISAM goals

PB-ISAM has three goals:

- To determine the maturity of an IS.
- To improve the IS capability.
- To promote discussion on what it means to be a *mature* and hence an *effective* IS organization.

### 7.2 PB-ISAM structure:

The PB-ISAM, in accordance to the traditional CMM, should be a maturity growth model, consisting of five maturity levels. Each maturity level should describe a stage in the maturity of an IS organization. The lowest level would be level one, the initial level. Organizations at level one in SW-CMM are characterized by working in an ad hoc manner and by unpredictable performance. This characteristic of level 1 organization can be employed to be used in PB-ISAM, so that the same characteristics are applicable. Therefore, if IS are working or are delivered successfully, it is because of individual heroism. Organizations at level two, the repeatable level, should deliver, and use IS with a repeatable quality, in other words, they should repeat earlier successful performances in similar circumstances. For an organisation to be at level 2, it has to be at level 1 also. At the third level, the defined level, the aim should be the standardization of services. Organizations at level three should employ standard processes to select, develop, deliver and use IS and should have implemented organization-wide processes to train employees who useIS and manage IS related resources and problems. The fourth level, the managed level, should aim attaining quantitative control over the IS processes. And finally at Level five, the optimizing level, continuous process improvement of use and delivery of IS should be aimed.

As in SW-CMM, each maturity level (except for level one) should contain a number of key process areas. To reach a certain maturity level within the PB-ISAM, each of the key process areas of that level and lower levels have to be implemented by the organization whose IS has been assessed. Moreover, for a key process area to be considered implemented each of the goals of the key process area should be reached. As in traditional SW-CMM, a key process area hence should consist of goals and of activities, which are called key practices. An organization that implements all activities from a certain key process area is expected to also reach the goals of that key process area. This relationship between maturity levels, key process areas, common features and key practices is shown in Figure 4.

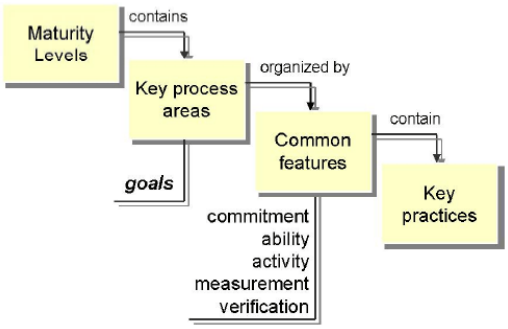


**Figure 4**: CMM structure

The CMM model distinguishes between five kinds of practices, called common features, which together, these five common features ensure that the goals of the key process area are reached. It is possible and appropriate to suggest such practices for the PB-ISAM. :

- Commitment to Perform: activities aimed at ensuring organizational and management commitment to the key process area activities.
- Ability to Perform: activities aimed at enabling the key process area.

 

- Activities Performed: the activities needed to get the job done.
- Measurement and Analysis: activities aimed at determining the status of the key process area.
- Verifying Implementation: activities aimed at verification of the implementation of the key process area.

However, in order to define the contents of PB-ISAM, it is important to take into account previously conducted studies in the area, and to use such studies as the guiding frameworks. Therefore, it is suggested that in order to develop a CMM for IS, the studies of Andersson, Von Hellens and Adeyakun, and DeLone & McLean, need to be taken into consideration. Initially, the guiding framework, can be formed by means of merging these studies to better depict and display the multi-dimension and multi-perspective attributes that a IS quality assessment model should embody. It can be a very interesting approach to place the SOLE model structure in the three-dimensional model of Adeyakun. Such an attempt can be simply graphically represented as shown in Figure 6 below. If DeLone and McLean's model were additionally merged into this model, the three quality dimensions would increase up to eight, resulting with a more comprehensive and complete model.

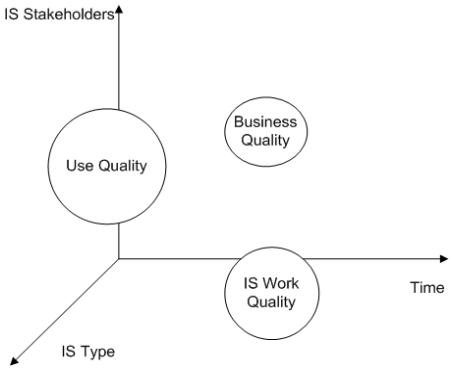


**Figure 6**: A proposed mixed framework

For a complete assessment of an IS, it is important to take into account all three IS stakeholders' perspectives: (1) IS planner, (2) IS developer, and (3) IS user. These three views are important since, as DeLone and McLean states in their latest paper (2003), different stakeholders in an organisation may validly come to different conclusions about the success of the same information system.

The capability (or maturity) level achieved by the Information System assessed may be presented as in Table 1 below.

**Table 1**: Information system success capability level ratings for the proposed assessment model (PB-ISAM)

| Maturity Level | IS Success Dimensions | Rating |
|---|---|---|
| Level 1<br>PRODUCTION<br>Performed | System Quality | Largely or fully |
| Level 2<br>PRODUCT<br>Managed | System Quality<br>Information Quality | Fully<br>Largely or fully |
| Level 3 RECEIPT<br>Established | System Quality<br>Information Quality<br>Information Use | Fully<br>Fully<br>Largely or fully |
| Level 4<br>INFLUENCE ON<br>RECIPIENT<br>Predictable | System Quality<br>Information Quality<br>Information Use<br>User Satisfaction | Fully<br>Fully<br>Fully<br>Largely or fully |
| Level 5<br>INFLUENCE ON<br>SYSTEM<br>Optimizing | System Quality<br>Information Quality<br>Information Use<br>User Satisfaction<br>Net Benefits (Organizational & Workgroup Impact)<br>Service Quality | Fully<br>Fully<br>Fully<br>Fully<br>Largely or fully<br><br>Largely or fully |

This capability model is proposed by merging the DeLone and McLean's IS success dimensions; and the SW-CMM and ISO 15504 assessment methods. For instance, as an interpretation of this table, if we take an IS which is found to be at maturity level 3 after assessment by any of the three stakeholders; according to the proposed table, this would mean that the criteria for levels 1 and 2 are fully accomplished by that IS. That is, the indicators for system quality and information quality are achieved at least to the percentage of 86% or higher; *and* the indicators for information use are achieved at least to the percentage of 51% or higher. Being at level 3 also indicates that other success dimensions which belong to higher levels, level 4 and 5, are therefore either not achieved (0% to 15%) or are partially achieved (16% to 50%). The rating scale proposed shall be defined according to the ISO 15504 specification as shown in Table 2.

**Table 2**: Indicator rating scale adopted from ISO 15504 (1998) specification

| Percentage interval | Rating |
|---|---|
| 0%-15% | Not achieved |
| 16%-50% | Partially achieved |
| 51%-85% | Largely achieved |
| 86%-100% | Fully achieved |

## 8. Conclusions

This study overall, aims to distinguish and identify existing models of IS quality assessment ; and

based on these established models to propose a process based model ; and finally to accomplish an assessment based on this model in an organization, the METU Informatics Institute Online Teaching System. In this paper, the first two steps of these total three steps were successfully accomplished in great extent. However, the final step is still to be completed.

Until recently, the practitioners and researchers in the general Software domain have considered and focused primarily to the quality of the artifact being developed, on the Software. This artifact, as a laboratory object, was analyzed and depicted usually away from its natural environment, the organizational context. Merging the organizational context with the artifact, we obtain IS, which are multi-dimensional and entities with multiple perspectives. Considering the quality of Software, is definitely considering only the manufacturing process of that product, leaving out the context within which that product will operate. Of course the manufacturing quality of a product is important, but until recently, in the Software area, all attention was focused only in this "manufacturing" process.

This study has shown that some important awareness is being developed within several areas of influence, on the subject of IS quality. Two models were briefly introduced and discussed in this paper. As it has been elaborated, these two models, lack extensive and detailed consideration and are definitely incomplete with respect to many points. On the other hand, they are compatible within themselves and with other generally accepted evaluation approaches. There are no contradictions or real overlaps, hence they can be complementary in their positive effect on the organization. Information systems, by definition (DeLone and McLean, 1992), are integrated systems for providing information to support operations, processes, management analysis and decision-making functions within an organisation. Therefore IS quality shall comprise the requirements of the business organization, the users, and the IT personnel (Ozkan and Bilgen, 2003). In that regard, the paper in general supports the claim that the use of a systematic IS based performance evaluation approach would improve organizational effectiveness.

## References


Adelakun, O., (1991) *Planning and Managing Strategic Information Systems Project: A Case study on IS quality process in ABC,* Turku Centre for Computer Science TUCS, and Turku School of Economics and Business Administration, Institute of Information Systems Science

Andersson, T., Von Hellens, L., (1997) *Information systems work quality*, Information and Software Technology 39, pp. 837-844.

Dahlberg, T., Jarvinen, J., (1997) *Challenges to IS quality*, Information and Software Technology 39, pp. 809-818.

DeLone, W. H., McLean, E. R. (1992) Information systems success: The quest for the dependent variable. *Information Systems Research,* 3(1), 60-95.

DeLone, W. H., McLean, E. R. (2003) The DeLone and McLean Model of Information systems success: A Ten-Year Update *Journal of Information Systems,* Spring, Vol.19, pp.9-30

Eriksson, I., Törn, A., (1991) A model for IS quality*, Software Engineering Journal*, pp. 152-158, August.

Galliers, R.D. (1994) *"Relevance and Rigor in Information Systems Research: Community, in Business Process Reengineering - Information Systems Opportunities and Challenge*s", Proceedings of the IFIP TC8 Open Conference on BPR, Elsevier, Holland.

Garvin, D., (1987) *Competing on the eight dimensions of quality*, Harvard Business Review, Nov-Dec, pp. 101-109.

ISO/IEC TR 15504 (1998) Information Technology, Software Process Assesment, Part 2: A reference model for processes and process capability

Laudon K. C., Laudon J. P., (2001) *Management Information Systems*, 7th Ed. Prentice Hall.

Mumford, E., Hirschheim, R., Fitzgerald, G., Wood-Harper, T., (1984) *Research Methods in Information Systems*, Procedings of the IFIP WG 8.2 Colloquim, Manchester Business School, 1-3 September.

Niessink F., Clerc V., Van Vliet, H., (2002) *The IT Service Capability Maturity Model*, IT Service CMM Release L2+3-0.3, Status: Draft, December 6.

Ozkan, S. (2003) *Information Systems Success Dimensions: An Integrated Approach* 4th International Conference on Management Sciences, Istanbul Technical University, Istanbul, Turkey, 5-8 March.

Ozkan, S., Bilgen, S. (2003) "*Notes Towards IS Assessment: A Comparison of two Models within the Context of the Internet"* in proceedings of the IADIS International Conference WWW/Internet 2003, Algarve, Portugal, Vol.2, pp.1215-1219

Siakas, Kerstin V., Georgiadou, Elli. (1999) *Process Improvement: The Societal Iceberg*, EuroSPI99.

Vidgen, R. T., Wood-Harper, A. T. & Wood, J. R. G. (1993) A Soft System Approach to Information Systems Quality*, Scandinavian Journal of Information Systems*, Vol. 5, pp. 97-112.

Von Hellens, L.A., (1997) Information systems quality versus software quality: a discussion from a managerial, an organisational and an engineering viewpoint*, Information and Software Technology* 39, pp. 801-808.

Weinberg, G. M., (1992) "*Quality Software Management",* Addison Wesley.


.

 

 

# Evaluating e-Commerce Success – A Case Study

**Shaun Pather[1], Dan Remenyi[2] and Andre de la Harpe[1]**
**[1]Cape Peninsula University, South Africa**
**[2]Cape Peninsula University, South Africa and Trinity College Dublin, Ireland**
pathers@cput.ac.za
dan.remenyi@tcd.ie
andre@kalahari.net

**Abstract:** The business community in the past decade has been characterised by debate over the value or effectiveness of e-Commerce and how this type of technology needs to be implemented. During this period the business world has witnessed many examples of failures of Internet based business. There is little doubt that the high failure rate in Dot.Coms had much to do with misconceptions regarding the ease with which e-Commerce could be implemented. Unrealistic expectations caused tried and tested business rules to be abandoned as hyperbole over took sound business sense. Although it is clear today that the Internet and the Web can facilitate business processes to add value to organisations, this technology has to be managed with considerable care. This paper reports on a case study conducted in kalahari.net, a well known South African e-Tailing business.

This case study highlights several valuable lessons to do with the evaluation of an e-Commerce investment and how to ensure its success. Specifically the case study closely examines aspects of kalahari.net's IS management policy, and identifies a set of preliminary e-Commerce success dimensions.

**Keywords:** e-Business, e-Commerce, Internet business, web-facilitated business, Information Systems Management, business evaluation, IS success

*In brick and mortar organizations, applications support the business, but in e-Businesses applications are the business (Kroll, 2001).*

## 1. Introduction

In the ten years, approximately, since e-Commerce became a major issue there have been many attempts to create and operate successful businesses facilitated by this technology. The establishment of a very large number of Dot.Com businesses during the second half of the 1990s reflects the high hopes which were placed in this technology. Most of these attempts have failed and the reasons for such failures are well catalogued (See for example Ames, 2001; Carton, 2001). The number of successes has been relatively small and the lessons, which can be learnt from these organisations, are not yet well known. What has become clear though, is that basic business principles still hold (Remenyi et. al. 2004).

Over the last decade, a multitude of studies have focused on various aspects of the practice of e-Commerce. Several of these studies have focused on measures and frameworks for evaluating the success of the IS function. DeLone and McLean's review of the academic and trade journals over a seven year period (1996-2002) found that "most of the articles were conceptual in nature…" (DeLone & McLean, 2004: 35). As a contribution towards operationalizing e-Commerce success metrics, this paper reports on an ongoing investigation into e-Commerce success factors. The paper presents a single case study of a well known South African e-Commerce venture, kalahari.net[1], which is regarded as one of the successes in the e-Commerce environment in that region (Financial Mail: 2004). In particular the case study examines the complexities involved in managing kalahari.net's information systems, and highlights a preliminary set of indicators of success of the IS function.

## 2. Background to e-Business in South Africa

South Africa was not immune to the e-Commerce hype. According to the Department of Trade and Industry, expectations that the Internet would boost SA's economy and revolutionise the market by allowing small firms to compete equally with larger rivals did not materialise. The general manager of information technology and communications is quoted as saying that "Many of the initial hopes of the internet revolution have been disappointing" (Stones, 2002). As well as the disappointments of the SME sector, larger organisations which participated in the e-Commerce gold rush also ran into problems. There are many examples of this. The following are some of the well known examples in South Africa.

- Broadcast Interactive Group, an internet venture with the backing of several radio stations, closed before it was properly off the ground.

[1] kalahari.net® is a Trademark and this is recognised by the authors.


Reference this paper as:
Pather S, Remenyi D and de la Harpe A (2006) "Evaluating e-Commerce Success – A Case Study" *The Electronic Journal of Information Systems Evaluation*, Vol. 9, Iss. 1, pp 15-26, available online at www.ejise.com

- In July 2001, Woza, a successful independent online content company closed down after its main investor, Bytes Technology Group pulled out – even though it claimed a page impression rate of 5.5 million a month.
- The Shoppingmatrix.com, which set out to mainly retail DVDs and music CDs shut down after alleged cash flow problems.
- The SPAR national supermarket chain closed down its online shopping site due to extremely poor sales via the site.
- The banking venture Blue Bean and Twenty20 were also South African e-Commerce ventures that did not last very long (with the latter recently being relaunched).

The Department of Trade and Industry, claims that the biggest disappointment of the Internet had been its failure to empower small businesses through its capacity of allowing them to communicate more easily with customers and trading partners, and to close the gap between big and small companies. They claim that among the problems faced by start-up online companies in South Africa, businesses underestimated the necessity of having a trusted brand name to secure online sales (Stones, 2002).

A survey of online retail activity in South Africa (Goldstuck: 2002) reports that the failure rate of e-Commerce in South Africa was 35% (2000-2002) and this was predicted to grow to 40%. This survey points out that online retail in South Africa is "at a very early stage of its market penetration, and remains deeply immature in its implementation."

Many factors contributed to the demise of the online retail market during this period. In general it can be stated that the business world underestimated the complexities and importance of many aspects of business including marketing, finance, human resources as well as not properly appreciating the challenges offered by the technology itself. Thus organisations like Boo.Com failed as much from technological and IS management blunders as it did from general business incompetence.

In light of the foregoing, research into e-Business in South Africa is timely in assisting practitioners to obtain an understanding of the complexities surrounding this business paradigm. The authors have chosen to conduct a detailed study of the well-known South African business called kalahari.net. The objective of this case study is to evaluate this organisations success and to identify some of the key characteristics of its operation which have lead to this success. Although this case study is wide ranging in scope its main thrust is related to how kalahari.net manages and evaluates its IS function.

## 3. Research methodology

The approach to this case study draws mainly on interpretivist methods. Cognoscente of recent criticisms of the value of management research (Starkey & Madan: 2001) and that of the relevance of certain types of empirical research in IS (Benbasat & Zmud: 1999), the authors adopted a case study approach (Yin: 1994).

The rationale of using a case study was to allow an in-depth examination of a real world problem based on an existing company already engaged in e-Commerce. Case study research according to Harrison (2002: 177) is more aptly described as a strategy than a method. It sets out to address the understanding of a phenomenon within its operating context. Of necessity, case study research is about making sense of the complexities of a real-world working environment and this is the approach taken here.

The primary sources of data for the case study were interviews with knowledgeable informants from within kalahari.net, one of whom became a co-author of this paper. The interviews took place during June and August 2004 and during March 2005. The interview transcripts were analysed using qualitative content analysis (Henning, 2004: 104-109) to reduce the data through a process of coding. Motivations for choosing a qualitative approach to this investigation are provided by Babbie & Mouton (1998:270).

The authors made several visits to the premises of kalahari.net. In addition to this we also reviewed a number of public documents of the holding company, and carefully examined the web site. This also included registering as a customer, and making a purchase. The authors also generated a "complaint" to kalahari.net's customer support section, to determine how this was handled by the system.

## 4. An overview of KALAHARI.NET®

kalahari.net is a South African based business referred to by its owners and managers as an e-Tailer[2]. It is a web and Internet facilitated

[2] An e-Tailer is an online retailer… and in the B2C sector, the business model focuses on sales to the individual customer (Laudon & Traver, 2003: 71).

business which sells products such as books, CDs, DVDs, videos, software, hardware, wine, and health care goods. These products are sourced from South Africa, and elsewhere. In addition to these products kalahari.net has a number of online partners through which products such as ticketing solutions to theatres, cinemas and major events may be bought.

kalahari.net is operated as a business unit of Via Afrika which is a wholly owned subsidiary division of Naspers Limited. Naspers is listed on the Johnannesburg Stock Exchange and Nasdaq in New York. Naspers Limited, is today a R10 billion turnover enterprise and has R8 billion assets (Naspers, 2004). Via Afrika controls a number of different businesses operating as independent business units in book publishing and distribution, niche retail and entertainment, and in private education. In 1998 Naspers decided to take advantage of the opportunities offered by the Web and the Internet and launched through its subsidiary, Via Afrika, a number of internet businesses which included 24.com (now mweb.co.za), kalahari.net, fin.24.com and news24.com.

As mentioned above kalahari.net was first envisaged as a book selling business. The idea for this came from the CEO of Naspers. John van Relihan, who was responsible for Via Afrika's book club division, grabbed this opportunity. At this point they owned the largest book club in Africa. With so much publicity concerning the apparent success of Web based businesses such as Amazon.com it was thought that this could be emulated in South Africa and Via Africa could do it.

Some "venture capital" money was allocated by the holding company to fund the operation. John van Relihan set up an independent team and they began working towards the creation of an e-Commerce operation.

This was the period of extraordinary hype concerning the Web and it was generally thought that it was not difficult to set up an e-Commerce operation and that it could be achieved in a short period of time. The dictum emanating from American business schools and consultants in this period was that with as little as $50,000 and within 60 days an e-Commerce website could be up and running. This type of thinking omitted the issue of on-going costs and revenue and the breakeven period. As there was no-one in the Via Africa group with e-Commerce experience this type of exaggeration appears to have been believed.

As a result of this thinking there was virtually no preparation for kalahari.net. There was no business case prepared, although it was generally thought that a breakeven situation would be reached in two years. No rational for suggesting a period of two years to reach the breakeven point has been offered and even today after six years, breakeven has not been achieved. Furthermore there was no risk analysis performed. The development of the website was rushed and a launch took place in October 1998 approximately one month after the decision was taken to get into this business.

Not surprisingly the Website attracted little business. The problems with the first attempt to make kalahari.net an e-Business were typical of the many errors made by start up Dot.Com of that period.

The marketing plan was ill conceived. The technological issues and challenges, especially relating to information systems architecture, were not really understood. The web-site interface is described as being "horrendous", with long download times, and poor information on the site. The crucial internal data loading process took weeks and resulted in unreliable product data. There was no fulfilment process in place. The kalahari.net team was made up of three business oriented people, and approximately ten newly graduated IT students. The sourcing policy was not well thought through. The funding was not well conceived or planned – Via Africa had what they refer to as an "open book" basis for funding kalahari.net. In addition, there were inadequate internal controls with which to management the businesses.

Failure was certainly staring them in the face. The sum of the potential loss was not big in the Naspers world but the failure of an e-Commerce venture in the full light of the public media was a most unattractive prospect.

As a result of this predicament management at both kalahari.net and Via Africa decided to re-launch the business. Via Afrika appointed Susan van der Schijff who had a direct marketing background to take over the reigns at kalahari.net in March 1999. Susan had been the product developer for the book club and she had a much better understanding as to what an e-Business was about. The re-launch took place in October 1999 and in so doing the kalahari.net management ensured that:-

- A more knowledge and experienced team was put in place. The largely inexperienced IT staff were dismissed, and a new team of only

 

four people were recruited. This included a seasoned IT manager as well as a web-site designer.

- The back office systems were reorganized to become more responsive to the needs of a web-facilitated business and the fulfilment processes were redesigned.
- The web-site itself was substantially overhauled.
- The relationships with suppliers were strengthened to ensure that more accurate product data was provided.
- The direct marketing experience of the new general manager, was used to implement new strategies to attract and retain customers. This included diversifying the product base.

Thus, in 1999, kalahari.net was re-launched in a much more thoughtful and professional way, keeping a close watch on all the important business variables. Since the re-structure and the re-launch of Kalahari.net the business has grown from strength to strength, as indicated by its increasing turnover shown in Figure 1.

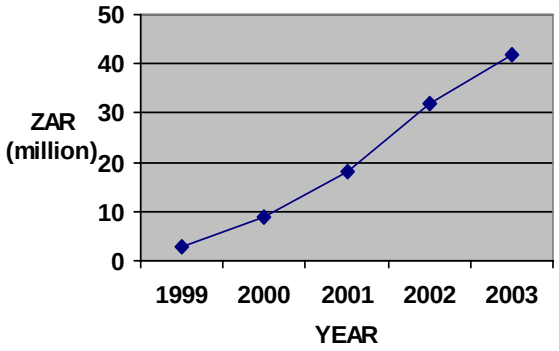


**Figure 1:** kalhari.net turnover

Over the past 6 years kalahari.net has become known as one of the best recognised e-Commerce brands in South Africa. It is not only well established but it has become a role model in South Africa. It was recently rated as the best site (Figure 2) from among of a possible 1000 websites (including Amazon.com) in a Financial Mail survey (Financial Mail, 2004). It was identified as the subject of this study as it offers many lessons for both well established Internet facilitated business as well as for those who wish to enter this market for the first time.

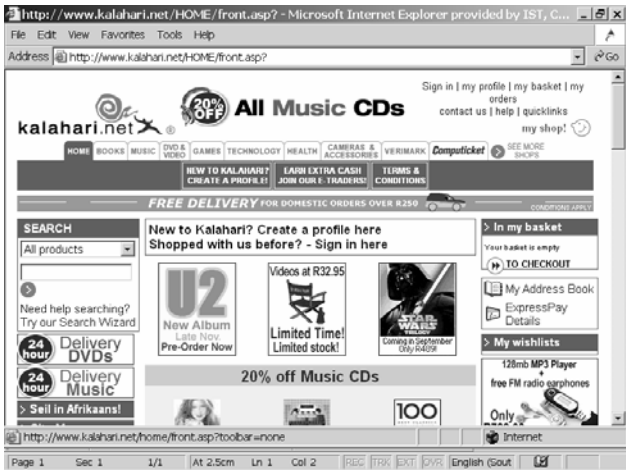


**Figure 2:** Kalahari.net ® Home Page (Screen shot shows partial page only)

## 5. An evaluation of Kalahari.net

As mentioned above kalahari.net is regarded as a successful e-Commerce or e-Business in South Africa. According to members of staff the reasons why kalahari.net is considered as such are:-

- Its year on year growth;
- It's the biggest B2C e-Commerce website in SA;
- It has 200,000 registered customers;
- It has received accolades from the South Africa press especially the prestigious Financial Mail;
- It has no domestic competition;
- It was the first e-Business to be fully compliant with the Electronic Communications and Transactions (ECT) Act;
- It is close to breakeven and intends to start making profits in the 2005-2006 period.

The company is regarded as a success despite the fact that it has not yet broken even. Breakeven is anticipated soon. However kalahari.net has the financial backing of a more substantive parent company which may probably be able to sustain it for quite some time if that was to become necessary. It is therefore problematical to really call kalahari.net a commercial success. It certainly is a public relations and awareness success and it is very beneficial to the Naspers Limited group to have a business which is so highly recognised in South Africa and which is so well regarded. But the objective of business is ultimately profit – or at least not making losses - and this has not yet been achieved. Therefore care needs to be taken with the use of the word success. If kalahari.net was an independent operation where the owner managers had to go to the financial market for

funding it is questionable as to whether it would have survived the hiatus in the financial markets.

Our evaluation of kalahari.net is that the work of the past 5 years has positioned the business so that if the current growth performance is sustained and if costs are keep under control it will become a profit generator in the near future. But referring to kalahari.net as a business success when it has not yet reached its breakeven point after six years is not a description we would readily wish to use.

## 6. Meeting the IS management challenge

As mentioned above one of the major objectives of this research was to understand how kalahari.net managed the IS function (comprising 4 major systems – See Figure 3) which is regarded as a core aspect of any e-Commerce business. The interviews reveals some very interesting IS management issues. Some of these are recognised as tried and tested IS management practices, however there are some innovative ideas in use at kalahari.net as well.

Before discussing the detail of some of the management processes used by the IS function in kalahari.net it is useful to point out how critical they perceive their IS function to be for their business. Besides the fact that the website has to function without error 24/7 they also rely heavily on a wide range of other information. In the words of management,

> *"Without a database there is no website and with no website there is no business. Data is needed to allow us to feed the site. Information such as how mnay items are available on the website, what the stock availability is, our pricing, the number of days to deliver and the number of days for products to arrive are essential. We monitor the supply chain closely. Where the stock is coming from? How much is international or how much is local? The weight, the volumetric mass, delivery dates, when will the customer receive an order, how many customers did receive on time, how many did not, to mention only some of the issues."*

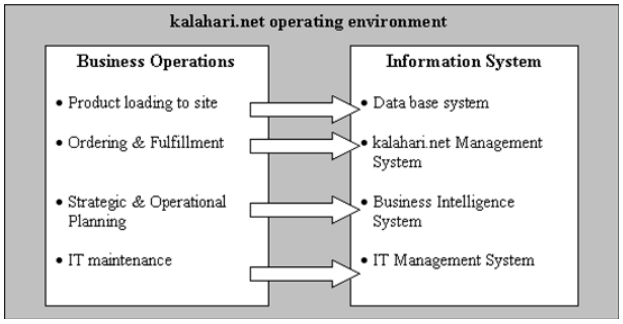


**Figure 3:** There are 4 central Information Systems at kalahari.net

In addition, the IS function is required to report weekly on issues such as, cash flows, order received per day, order deliveries per day, cost per order income, postage cost per order, exception reports on suppliers' costs.

There are also reports required on operational issues such as website down time, downloading of website time for customers etc. The management of kalahari.net clearly believe that their business is fuelled by information.

The following sections provide an overview of key areas of the organisation's IS management and operations. We view these as having a central role to play in achieving success with their IS.

### 6.1 Aligning the IT and business stakeholders

When the kalahari.net website initially commenced operations in 1998 the relationship between business managers and the embryonic Information Technology (IT) department was at a very low ebb. Thus in the initial period of kalahari.net's operations, the IS were plagued by serious problems such as bugs, and generally inadequate performance. This is hardly surprising when one considers the fact that the website was developed and was up and running within one month by what was in effect very young and inexperienced people.

Due to the rush to be in business the initial attitude of business managers was *"Get something going, get it on the web",* without proper consideration as to the implications of their requests. This is clearly an unsatisfactory approach to IS. However the IS function responded as best it could by trying to implement these requests without fully understanding what was really required. This led to unsatisfactory systems which in turn resulted in distrust between the two parties and an eventual a substantial breakdown in communications.

After this rather messy start the senior management made a strategic decision to bring

 

together in a much more functional way the business and IS stakeholders. In working towards this they firstly outsourced the IS department in total – the management of kalahari.net felt that the extra burden placed on managing IS personnel placed a burden on their abilities to keep focused on the business. In doing so, a very important condition was placed on outsourcing - the outsourced partner was required to base its personnel at kalahari.net's headquarters in Cape Town. Secondly controls embodied in IS development and maintenance methodologies were introduced for all aspects of IS work.

## 7. Maintaining the business and IT partnership

The philosophy behind kalahari.net attitude to maintaining a sense of partnership between the business and the IS function was to ensure continuous dialogue between these two groups. This was affected through regularly scheduled meetings.

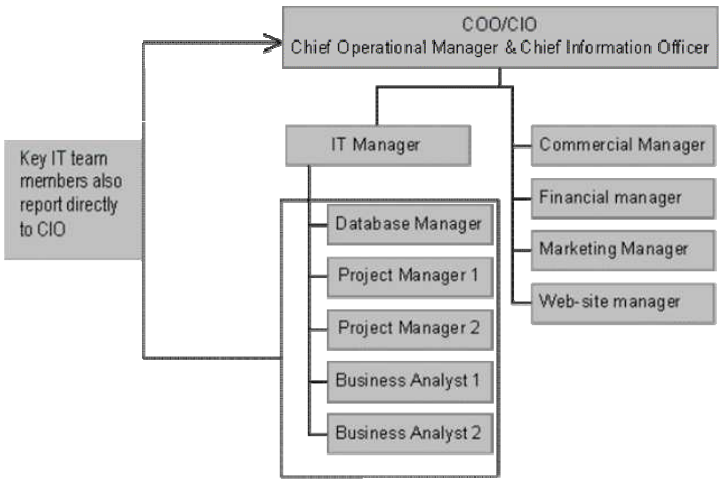


**Figure 4:** kalahari.net Business Task Team

The IS-Business partnership is maintained using a three-pronged strategy:

The first step in establishing this dialogue was to establish a weekly business-prioritising meeting. A task team of business stakeholders (see Figure 4) meets weekly with IT managers to address immediate, short-term and medium term priorities. These team meetings had a positive effective in fostering greater co-operation between the Business and IS, as well as between different managers of the business.

Secondly, attention was focused on systems development methodologies. IBM Rational Unified Process®, or RUP®, methodologies were introduced. RUP®, is a comprehensive, Web-enabled set of software engineering best practices that provide guidance for streamlining the team's activities (Kroll, 2001). Business analysts at kalahari.net have given RUP the thumbs up, and favour its ability to provide sets of project documentation that is meaningful to both IT and business stakeholders. In this way all the role-players can participate meaningfully systems development phases. As a result of the implementation of this new methodology, business stakeholders became fully involved in all phases of the systems development life cycle (see Figure 5).

Thirdly, key IT team members became involved in decision making at all levels. They are involved in making IT project decisions as well as in routine business meetings at which issues such as gross profit, budgets, turnover etc. are discussed.


**Figure 5:** RUP® Systems Development Cycle Source IBM [http://www-306.ibm.com/software/awdtools/rup/]

### 7.1 Database management is central to operations

Central to e-Tailing is the management of a large database. kalahari.net is linked to approximately 400 supplier databases all over the word – these range from 2 million products down to only 2 products. A primary kalahari.net database houses all product information (See Figure 6). Extreme care is taken to ensure the accuracy of the data. There is no room for errors e.g. a DVD that should retail for R899 should not be sold for R89.99.

Suppliers are totally responsible for providing the data to the primary databases each day. Since it is not possible to manually check such a huge stock database, exception reporting takes place to identify errors e.g. selling price is lower that the cost price.

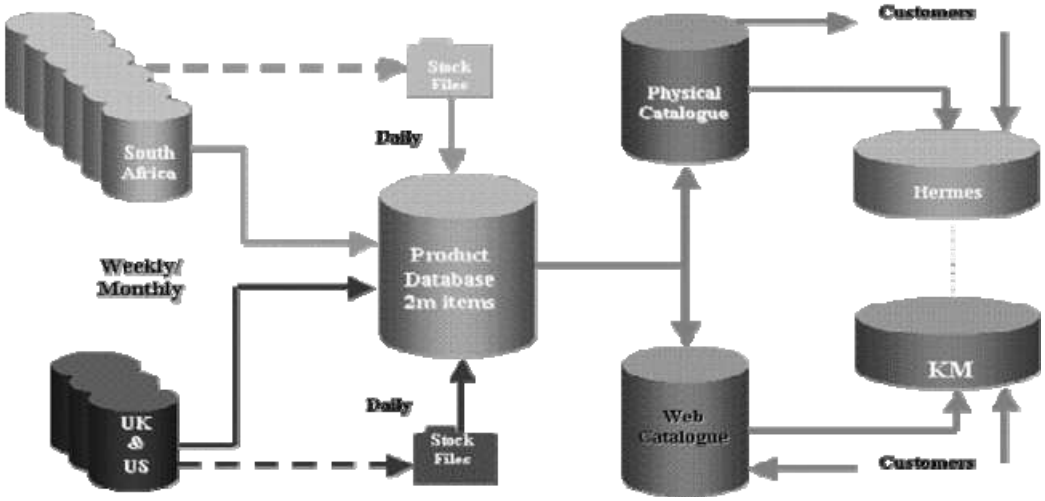


**Figure 6**: Ensuring suppliers update the database with accurate data is an all important business process. Source [ kalahari.net®]

In addition to exception reporting, a suppliers report is maintained to determine how many suppliers have actually sent through an updated data file, at what time did the report come through, and how long did it take to upload.

### 7.2 Ordering and fulfilment

The kalahari.net Management System (KMS) which is essentially a CRM system handles all ordering and fulfilment processing. Central to ensuring that these key processes function smoothly, is once again, the integrity of the data being handled.

KMS is used to monitor the purchasing cycle. Orders to suppliers are compiled twice daily, and depending on the facilities available at the supplier end, these are either electronic (EDI, FTP, email) or manual (via fax). All transactional data relating to orders in the pipeline are handled by KMS. This is processed and is available as a report which gets prioritised each morning at the start of the day. Based on the data and supplier reports, customers are contacted regarding delivery of products on time. kalahari.net also utilises manual methods such as telephone or fax to verify orders to suppliers.

### 7.3 Efficient delivery systems

Suppliers and couriers have deadlines and must adhere to agreed SLAs e.g. overseas suppliers have an hour after receiving a FTP file to report on fulfilment problems – thereafter they have up to 24 hours to have products transported to a designated courier. Within 2 days the products are shipped to kalahari.net's distribution centre near Cape Town International airport. Thereafter the shipping agents at the airport have 1 day to clear customs and excise and to transport the goods to distribution centre where streamlined processes ensure minimum delays in delivery to the customer.

### 7.4 Essential web-site features

There are several key features on the Web-site that are viewed by kalahari.net's management as central to success.

The first of these is the site's search engine. The use of the search facility by a customer usually marks the commencement of a potential transaction. Consequently the search facility is continuously tested. In addition to this all searches conducted are produced as

 

reports in order for the Marketing department to monitor the interests of shoppers.

Secondly, downloading time of the web pages is considered crucial. kalahari.net prides itself on providing its customers with a download time of 8 seconds or lower and considerable resources have been spent on design and technical infrastructure to make this a reality. Monitoring of the download time takes place through continuous testing of the loading time.

Thirdly, merchandising is viewed as being important. Quick turn around on updating stock items – IT must be able to support updating of products – they to achieve updating within 10 minutes.

Fourthly, with regards the payment gateway kalahari.net ensures that within 2 seconds a customer will receive feedback on verification of card, as well as authorisation of payment.

### 7.5 Business intelligence for management decision making

At the core of both strategic and operational decision-making is a Business Intelligence System (BIS) (Figure 7). Information from BIS is used for both strategic as well as day-to-day planning. BIS is an IS responsibility with a manager assigned to the BI system reporting to the CIO.

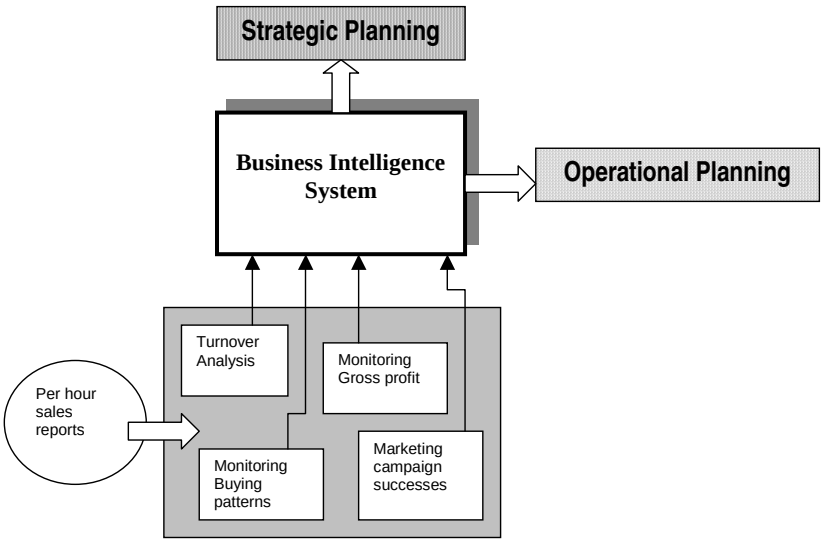


**Figure 7**: Business Intelligence System

As one of the core functionalities, the BIS provide hourly and daily sales reports, which are either used operationally or presented as detailed turnover analysis report to senior management. Turnover is also monitored per product category which can be viewed at any time of the day. This report assists in monitoring feedback on marketing, doing customer profiling, and in identifying areas in which the kalahari.net brand needs to be strengthened. Gross profit is also monitored daily in order that managers make decisions speedily when profits fall below expected margins.

### 7.6 Trust and privacy

In South Africa, the Electronic Communications and Transaction (ECT) Act was promulgated in 2002. The Act introduces a number of regulations to SA's e-commerce community. Among other things, the Act provides for the protection of consumers in terms of privacy and the security of electronic payments (South Africa, 2002).

In 2002 a B2C e-Commerce Readiness Survey was conducted by South African based Buys Attorneys and Trust Online, examining the websites and legal notices of 607 South African websites. Buys and Trust Online examined legal notices, privacy policies, terms and conditions of use, disclaimers and security policies. kalahari.net was found to be the only company that was 100% compliant with the ECT Act. As a result, in August 2002, the company went public, with a claim that it was the first South African e-Tailer to be compliant with the requirements of the ECT Act and was also fully in support of the efforts made to protect the consumer (e-Strategy, 2002).

## 8. Key elements of the business model

kalahari.net's management is emphatic that traditional business rules apply in the way they conduct business. However they are aware of the key differences between managing a brick and mortar retailing operation and that of an e-Commerce shop front.

First and foremost a brick and mortar-retailing model is essentially a collection model – the business procures products from suppliers and the customer goes to the business to "collect" the goods. However e-Tailing is a delivery model – customers don't collect but they receive.

Secondly, e-Commerce is a direct marketing business. All direct marketing principles are applicable whereas brick and mortar use distribution marketing. This is a very critical difference to the brick and mortar environment.

Thirdly, the demand on speed in the e-Commerce environment is much greater – customers are not tolerant if products are not available in the same way that they can be with a brick and mortar business.

Fourthly packaging in e-Commerce is crucial as the business is responsible for delivery of goods.

Fifthly, the database is a major difference e.g. a brick and mortar retailer is not as concerned about customer databases in the same way as the e-Commerce business. Accurate data pertaining to delivery addresses, email addresses, correct products etc. are of paramount importance.

Lastly, a very sound relationship with suppliers is required to ensure that the suppliers understand the importance of supplying accurate product data.

## 9. Towards an evaluation of the IS function

The way in which the IS function at kalahari.net evolved is interesting for several reasons.

During the initial period it appears that the IS requirement for an e-Business was substantially underestimated. Via Africa charged into this business with no real understanding of what IS infrastructure was required to create an organisation like kalari.net and to establish an industrial strength online business.

The IS team which developed the early website were clearly not adequately experienced from either a technological or from a business point of view. The comments made by kalahari.net concerning the IS function's lack of communications and understanding, although present in many organisations, seems to have been chronic in this case.

The termination of their employment and the outsourcing of the IS function seems to have been an inevitable consequence of the lack of planning undertaking in the pre-launch phase and the poor management during the first year of operation. Disposing of the IS team in this way this could not be regarded as a sign of successful management and it is likely that it was not confidence building for the other members of the organisation either.

At the re-launch of the business there is clearly a new understanding of the role of IS in such an online retailing organisation. The IS function has effectively moved to the centre of the stage. IS is perceived as being a major focus of business attention and the IS staff is brought right into the centre of business decision-making.

The regime of regular meetings of business and IS colleagues and quick follow up of decision and solutions to problems demonstrates a whole new understanding of how to use this technology. The use of modern systems methodologies that caters for the involvement of all stakeholders is also an indication of the new attitude towards IS in the business.

The following provide some indication that in the current setup of the IS function is performing at a satisfactory level:

- There is a very small error rate recorded on the product database.
- The payment gateway, and checkout process works smoothly.
- There are a number of payment options, all of which work well. This is due also in part to a strong relationship with the relevant 3rd parties.
- The web-site was voted as the best of a 1000 sites by the public. This is an indication that it is meeting the needs of its clientele.

 

- The synergies between the four major IS (and their associated portfolio of applications) discussed in this paper, appears to support the attainment of the business objectives - the bottom line of which is to get closer to break-even.

Like the evaluation of the success of the kalahari.net business, in general it is too early to pronounce on the success of the IS function - but it is nonetheless easy to see a much clearer understanding of the importance of IS in their business and a much more intensive approach to integrating IS and business functions.

## 10. Future requirements

Current management of the IS infrastructure occurs on a needs basis - with the impact on turnover being the most common indicator of success/failure. We would rather suggest that the business requires an integrated assessment procedure. Such a procedure must provide a framework to evaluate the various components of the IS infrastructure, in a way that gives management the ability to monitor the impact of the technology (and thus their investment) on the attainment of specific business objectives. Such an assessment should be conducted periodically, so as to aid medium term and longer-term decision-making.

As a starting point towards an integrated assessment procedure, a summary of critical e-Commerce functions that require IS support are identified. These are presented in Figure 8. which highlights a preliminary set of IS related metrics that were identified during the investigation. These metrics are mapped onto Angehrn's ICDT model (Angehrn, 1997). The ICDT model identifies four virtual spaces on the internet (Information, Commuication, Transaction and Distribution spaces), and is one useful way of differentiating the applicability of metrics. Furthermore these four virtual spaces lend themselves to the customer buying cycle (van der Merwe & Bekker, 2003) viz. need recognition, gathering information, evaluating information, making a purchase. The organization of metrics in this way will allow kalahari.net management to monitor critical areas of the business, in terms of identified needs e.g. an increase in page hits, but decrease in purchases may prompt a closer look at items associated with the *Virtual Transaction Space.*

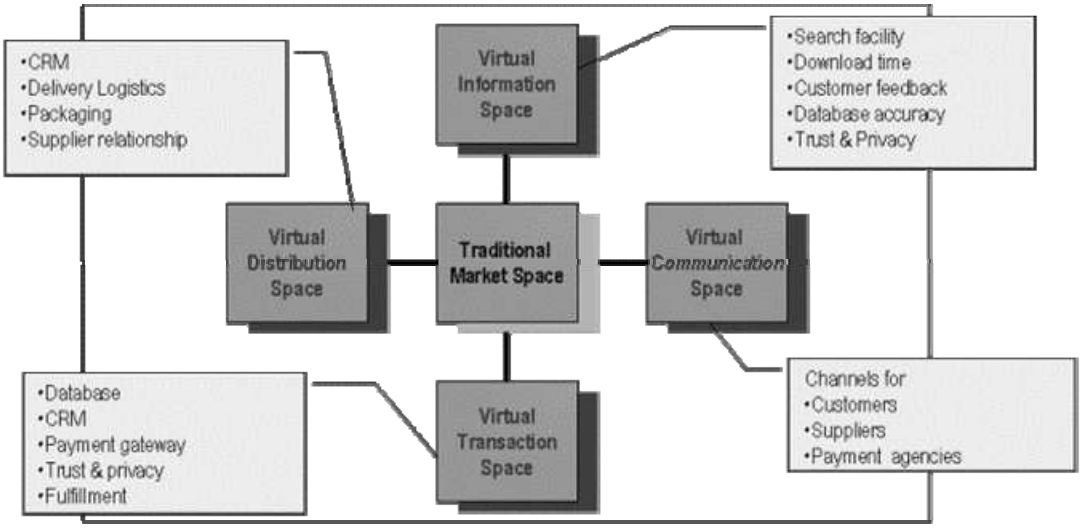


**Figure 8**: Preliminary metrics for evaluating e-commerce success

## 11. Lessons from the case study

From the discussions with the management of kalahari.net it is apparent that they believe that there are several special issues on which they have to focus which are different to a traditional bricks and mortar business. These issues as well as other pertinent success criteria, highlighted in the foregoing discussion, are presented in Figure 9, as a set of critical success factors.

In particular the following areas of business operations are deemed to be critical to the success of the e-Commerce business:-

- At all costs the electronic shop-front must be kept open and active 24/7/365 – this includes ensuring the availability of a hot-site. Site under construction notices and other distractions by which the shop-front is not available are extremely counter

productive to the business. The electronic shop-front needs to be as attractive as possible and have as wide an appeal as possible.

- Agility and flexibility - the IS infrastructure must allow for the business to be agile in responding to product updates to the web-site, advertising of specials etc. Changes have to be made immediately. Any substantial lead-time involved represents lost revenues.
- An effective, efficient, database infrastructure is essential for a successful online retailing business. An on-line shop lives by its twin databases. These are the product database which represents the inventory in the shop front and needs to be both up-to-date and error free and the client database which represents the main market opportunities. The product database needs to be seamlessly connected to the procurement and delivery systems to ensure that the correct items are delivered on time.
- IT and business stakeholders need to think together, and engage in joint decision making across all areas of business – including "non-IT" issues. There is just no room for a "culture-gap".


**Figure 9**: Critical success factors of the business

Make every effort to ensure the privacy of customer data. In the e-Tailing environment, the volume of successful transactions, is directly dependent on the trust the customers have in your systems' ability to offer secure transactions.

The web-site requires: minimum download time, accurate and smooth searching facilities, minimal click through rates, a quick payment facility, pre-orders facilities for purchasing of products prior to product launch, and visible security features. The electronic shop-front needs to centre around the customer.

## 12. Conclusion

In this paper we have described a South African e-Business operating in the B2C environment. A set of preliminary metrics for evaluation, as well as Critical Success factors, has been identified. The key aspects of the findings of the case study indicate a few important lessons regarding both the management of IS and that of the business. In addition to the lessons already outlined, the following need to be highlighted:

- One of the key issues relating to the creation, deployment, and maintenance of IS infrastructure in an e-Business relates to the relationship between business and IT stakeholders. The case study demonstrates that e-Business requires the bringing together of these stakeholders in not only the day-to-day management of the business, but also in key-decision making structures. Furthermore, the utilisation of systems development methodologies that are friendly to all stakeholders has been shown to be a contributory factor to this relationship.
- The use of a modern IT management tool is crucial to keeping the IS infrastructure functioning. Systems need to be monitored vigilantly, and have to be kept as lean as possible to ensure reduced workloads on both systems and people.
- Business is business, be it electronic or brick and mortar. However, the common element is the financial statements. Money does count, and kalhari.net has demonstrated that all eyes are on problems that directly affect turnover and gross-profit. Indeed these are monitored and reported daily, and business reacts with agility, if either of these are not meeting targets. The company's information system is always the first stop in reacting to a falling bottom line.
- In the electronic on-line trading environment, database technology is the most central aspect of the IS infrastructure. Extreme care, and diligence is taken to ensure that product data, and customer data, is accurate. There is no room for errors,

and successful transactions depends entirely on the way in which this is managed.

It is not our intention to offer these findings as generalizations for the e-Business sector in South Africa, and elsewhere, but rather to offer these findings as an initial set of criteria that could be used to adapt business models, and improve management practice. Many of the issues raised are mostly pertinent to the B2C environment, but issues around procurement of goods are important to the B2B sector as well.

Future work includes, further investigation and collection of empirical evidence of e-Commerce success metrics in order to verify and extend those identified; comparison of findings in this case study to others; and the development of more comprehensive framework for measuring the performance of IS in e-Commerce businesses.

## References


Ames,S. (2001). "Dot-coms making a comeback?". *ZDNet News*. December 27, 2001,[online] http://zdnet.co..co./2100-1106-277420.html [Accessed July 2002].

Angehrn, A.A. (1997). "The Strategic Implication of the Internet", *Proceedings of the 5th European Conference on Information Systems*.

Babbie, E., and Mouton, J. (1998) *The practice of social research*, Oxford University Press, Cape Town.

Benbasat, I., and Zmud, R.W (1999) "Empirical Research in Information Systems: The Practice of Relevance", *MIS Quarterly*, Vol 23, No.1, pp3-16.

Carton, S. (2001). *The Dot.Bomb Survival Guide: Surviving (and Thriving) in the Dot.Com Implosion*, McGraw-Hill, Boston.

DeLone, W.H., and McLean, E.R. (2004). "Measuring e-Commerce Success: Applying the DeLone & McLean Information Systems Success Model", *International Journal of Electronic Commerce*, Vol. 9 (1), pp. 31-47.

estrategy. (2002) "Kalahari.net first SA e-tailer to be compliant with ECT Bill", [online], http://estrategy.co.za/News.ASP?pklNewsID=7796&pklIssueID=189

Financial Mail. (2004) "Virtual Community: The Web is best at linking people together", 27 August 2004, pg 2.

Goldstuck, A. (2002) "The Goldstuck Report: Internet Access in South Africa, 2002", [online], http://www.theworx.biz/access02

Hendrick, K., and Heiman, R.V. (2001) "Achieving ROI with Rational ClearCase", *IDC,* January 2001, [online], http://whitepapers.zdnet.co.uk/0,39025945,60071840p-39000565q,00.htm

Henning, E. (2004) *Finding your way in qualitative research*, Van Schaik, Pretoria.

Kroll, P. (2001) "The RUP: An industry-wide Platform for Best Practices", *The Rational Edge e-zine for the rational community*, [online], http://www.therationaledge.com/content/dec_01/f_TheRUP_pk.html

Laudon, K.C. and Traver, C.G. (2003) . *E-Commerce: business. technology. society*, Pearson, Boston.

Naspers Limited. (2004) "2004 Annual Report", [online] http://www.naspers.com/Financials/ar2004eng/index.asp

Remenyi, D., Grant, K., Pather, S. (2004). "It was a Shock when Boo went Under – The Legacy of the e-Bubble: Lessons for Managers", Journal of General Management, Vol 29(3), pp. 24-36.

South Africa. (2002) *Electronic Communications and Transaction Act, Act 25 of 2002,* Government Printer, Pretoria.

Starkey, K., and Madan, P. (2001) "Bridging the Relevance Gap: Aligning Stakeholders in the Future of Management Research", *British Journal of Management*, Vol 12, pp.3-S26.

Stones, L. (2000) "Internet Fails to Put SA Business on the Global Map" Business Day: 13 September 2002,[online], http://allafrica.com/stories/200209130288.html

Van der Merwe, R., and Bekker, J. (2003). "A framework and methodology for evaluating e-Commerce Web-sites", *Internet Research: Electronic Networking Applications and Policy*, Vol 13 (5), pp.330-341.

Yin,R. K. (1994) *Case Study Research: Design and Methods*, Sage, London.


 

# Stock Price Reaction to Investments in Information Technology: the Relevance of Cost Management Systems

**Narcyz Roztocki[1] and Heinz Roland Weistroffer[2]**
**[1]State University of New York at New Paltz, USA**
**[2]Virginia Commonwealth University, Richmond, USA**
roztockn@newpaltz.edu
hrweistr@vcu.edu

**Abstract:** The identification of conditions and factors under which investments in Information Technology (IT) can be expected to yield tangible returns is the subject of many productivity studies. Event study methodology, which examines the reaction in the stock price to announcements of different types of IT investments, is one approach to this kind of research. In the research presented in this paper, we use event study methodology to investigate the effect of cost management systems on payoffs from IT investments. The motivation for our research is based on the assumption that companies possessing reliable cost management systems, such as Activity-Based Costing (ABC), are less likely to make expensive mistakes when investing in IT. Furthermore, the companies that use ABC and thus know the costs of their operation, are better able to single out those IT projects which positively impact the bottom line and competitiveness. In our study, we use a sample of three companies that are adopters of ABC, to examine the impact of 81 IT investment announcements on stock prices.

**Keywords:** Activity-based costing, cost management systems, event study methodology, information technology productivity paradox

## 1. Introduction

Information Technology (IT) plays a crucial role in today's business environment, and a substantial portion of corporate budgets is dedicated to IT (Weill *et al.* 2002). However, investments in IT do not always produce the desired pay-offs; companies often do not experience the expected tangible benefits from their investments. This lack of tangible benefits derived from IT investments, also called the Productivity Paradox of IT (Brynjolfsson 1993), has been widely discussed.

Many published studies related to the productivity of IT focus on the identification of conditions and attributes under which investments are most likely to result in tangible benefits. For example, investments which cannot easily be duplicated are seen as promising to increase competitiveness and profitability (Clemson & Row 1991). But surprisingly few publications deal with the issue of quality of data needed for cost analysis.

In many organizations, a proper estimation of costs needed for cost benefit analysis represents a major challenge, as traditional cost management systems are still widely used (Ness & Cucuzza 1995). Distorted cost estimates in turn often lead to poor strategic decision making (Johnson 1987, Johnson & Kaplan 1987). Activity-based costing (ABC), a more modern cost management method, produces more reliable cost estimates by looking at activities and multiple cost drivers when tracing overhead (Cooper 1987a, 1987b, 1988, 1989). In many instances, companies using ABC were able to achieve substantial improvements in their cost structures and show higher profits (Cooper & Kaplan 1991). In addition, ABC implementation resulted in better decision-making processes (Cooper & Kaplan 1988).

For all these reasons, it may be assumed that companies extensively using ABC will make different investment decisions than those relying on traditional accounting methods. It others words, firms using ABC can be expected to follow different strategies, when investing in IT, from those not using ABC. Firms using ABC may prefer IT investments which better support their cost reduction effort and therefore produce more tangible benefits.

The objective of the research presented in this paper is to gain an understanding of the effects of cost management systems, such as ABC, on IT-related investments. The remainder of our paper is structured as follows: First, we briefly review prior event studies dealing with the impact of IT investments. Next we present our research hypotheses and the motivation and background leading to these, followed by an explanation of our data collection and analysis methodology. We conclude with a discussion of our results and possible directions for future research.



Reference this paper as:
Roztocki N and Weistroffer H R (2006) "Stock Price Reaction to Investments in Information Technology: The Relevance of Cost Management Systems" *The Electronic Journal Information Systems Evaluation* Volume 9 Issue 1, pp 27-30, available online at www.ejise.com

## 2. Previous event studies of IT investment announcements

A fundamental assumption in the field of finance is that in an efficient market, stock prices fully reflect all available information. Thus, when new, unexpected investment information is released, the stock prices adjust accordingly. If the content and implication of the released information are perceived as adding business value to the shareholders, the stock prices are expected to increase (Fama 1970, 1991, Fama *et al.* 1969). Recently, several studies of stock price reaction to IT investment announcements have been conducted, to investigate the fundamental question: "*does IT matter?*" (Carr 2003).

Dos Santos *et al.* (1993) examined stock price reaction in the context of industry and innovation. With respect to industry, the expectation was that IT would have a larger impact on financial firms than manufacturing firms, due to the information intensity of the financial industry. However their results could not confirm this. With respect to innovation, the assumption is that the introduction of new technology or technology enabled products will result in a competitive edge until the technology becomes routine within the industry. The results of the study did show positive excessive stock price returns for innovative IT investments. Im *et al.* (2001) examined stock price reaction in the context of industry, firm size and time period. They found no price reaction for larger firms, and positive price reactions for smaller firms. They also found that prices reacted more positively to newer announcements of IT investments than to older announcements. Chatterjee *et al.* (2002) investigated whether investments that target IT infrastructure rather than applications induce positive price reactions, which seemed to be confirmed by their analysis. Dehning *et al.* (2003) found positive, excessive returns to announcements of IT investments by firms making transformative investments.

## 3. Hypotheses and motivation

No previous event study has looked at the effect of cost management systems on excessive stock price reaction to IT investment announcements. Whereas traditional cost accounting systems lump costs into a few overhead pools or categories, with Activity-Based Costing (ABC), these categories are divided into scores of activities, and costs are allocated to each of these activities. While this approach requires more data collection, it also results in more accurate cost estimation. Our argument is that organizations possessing reliable cost managements systems, such as ABC, are less likely to make expensive mistakes when investing in IT, and that this is recognized by the market:

> *H1: Announcements of IT-related investments in companies using ABC will result in positive reactions in stock price.*

As previous studies have shown, the stock market generally tends to react positively to innovative or transform IT investments, i.e. investments that do more than merely automate existing processes and thus increasing efficiency. Automate IT investments tend not to lead to excessive stock price reaction, as these are easily duplicated and thus provide only short-term competitive advantage. However, companies using efficient cost management systems could longer enjoy cost benefits since their cost reduction efforts are less likely to be duplicated by companies using inefficient traditional cost allocation systems. Thus our assumption is that with companies using ABC, there will be a positive reaction in the stock market following automate IT investment announcements:

> *H2: Announcements of automate IT-related investments in companies using ABC will result in positive reactions in stock price.*

Though generally, innovative transform investments are seen as positive by the stock market, such investments also are seen as more risky, and in companies using ABC will result in a less favorable stock market reaction than automate investments:

> *H3: Announcements of transform IT-related investments in companies using ABC will result in a less positive reactions in stock price than announcements of automate investments.*

## 4. Methodology

### 4.1 Data collection

We started our selection of firms with a published list of early adopters of ABC (Gordon & Silvester 1999). In order to assure that these companies are still using ABC, we searched the Internet for press reports and other references about these companies and their usage of ABC. We then selected three companies where we can reliably assume that they still use ABC.

Next, for this selected group of ABC users we searched Lexis-Nexis database for announcements about IT-related investments. Each announcement was examined for its relevance and coded as "automate or "transform", using the coding rule developed by Dehning *et al.* (2003). Six announcements did not fit into either category. Table 1 shows the number of announcements for each of the three companies,

 

and the number that were considered to refer to "automate" and to "transform" investments.

**Table 1:** Number of Announcements

| | Total Events | Total "Automate" | Total "Transform" |
|---|---|---|---|
| Parker Hannifin | 20 | 12 | 7 |
| Honeywell | 34 | 18 | 16 |
| United Technologies | 27 | 13 | 9 |
| Overall | 81 | 43 | 32 |

### 4.2 Data analysis

Similarly to earlier event studies of IT investment announcements (see for example Dos Santos *et al.* (1993)), we calculated the *cumulative standardized excess return (CSAR)* for the company stock for each IT investment announcement over two days, day 0 and day -1. The day of releasing the news about a given IT investment is defined as day 0; however, if the news is released on a day when stock markets are closed, the first following business day is defined as day 0 while the day -1 is the business day before the announcement day. The reason that the day before the announcement is included is that generally the stock market receives the investment information the day before the official public announcement. We used a 200 days (day –201 through –2) estimation period in calculating the CSARs.

## 5. Results

Table 2 shows the results of our investigation. The overall average CSAR is -.09729, (not significantly different from zero) thus not supporting our Hypothesis 1. However, the average CSARs for each company for automate investments is positive, supporting our Hypothesis 2. Also, both the average CSAR and the number of positive CSARs are larger for automate investments than for transform investments for all three companies, supporting our Hypothesis 3.

**Table 2:** Average cumulative standardized excess returns

| | Average CSAR | Average CSAR "Automate" | Average CSAR "Transform" | Percentage Positive | Percentage Positive "Automate" | Percentage Positive "Transform" |
|---|---|---|---|---|---|---|
| Parker Hannifin | -0.15774 | 0.29390 | -0.91587* | 45 % | 67 % | 14 % |
| Honeywell | -0.14443 | 0.32015 | -0.66708** | 41 % | 72 % | 6 % |
| United Technologies | 0.00685 | 0.41395 | -0.17694 | 41 % | 54 % | 33 % |
| Overall | -0.09729 | 0.34117* | -0.58365** | 42 % | 65 % | 16 % |

* Significant at the 5 % level
** Significant at the 1 % level

## 6. Conclusions and discussion

In this study, in order to examine the possible effect of cost management systems on investments in IT, we investigated the change in stock price of three companies known as early adopters of ABC. Analysis of our data suggests that even for companies using ABC, overall, investors do not associate IT investments with a substantial change in future cash flows, which is manifested in a lack of abnormal returns in the full sample. In other words, our findings indicate that the mere adoption of ABC in a company will not necessarily convince the investors that most investments in IT will generate business value.

However, our results also suggest that for companies using ABC, investors do associate automate IT investments with higher future cash flows, which can be seen in the positive abnormal returns for this kind of IT investments. In contrast, investors seem to expect transform IT investments, in companies using ABC, to have no positive effect on future cash flows.

Though we hypothesized that the stock price reaction to transform IT investments, in companies using ABC, would be less positive than to automate IT investments, the negative stock price reaction is surprising, since it is generally assumed that such transform IT investments do increase business value.

Our findings do indicate that accounting practices in a company may have a substantial effect on the success rate of IT investments. Regarding the evaluation process of the automate IT investments, the use of a more reliable cost management system, such as ABC, appears to significantly increase confidence in these investment decisions. The negative reaction to transform investments implicates that the evaluation of IT investments is a complicated

issue and much more research in this area is needed.

The results of our study also provide further evidence that event study methodology, a common research tool in corporate finance and strategic management, is also very promising in the field of IT.

## 7. Limitations and directions for future research

One limitation related to our methodology is the lack of accounting for confounding news, which are not related to the IT investments but may have impact on stock prices. In spite of our efforts to identify and to exclude possibly contaminated announcements, it is always possible that such news did exist. In addition, we did not extensively study the timing of the announcements. It is, for example, possible that some of the announcements would receive a different investors' reaction if a major competitor also just implemented a similar IT system.

Another limitation, of-course, is the limited number of data studied; we only looked at three companies using ABC. A future study may look at a larger number of companies using ABC as well as companies using traditional accounting systems, and compare the results between the two over the same time period.

Finally, as in earlier event studies, we examined the stock market reaction to announcements related to IT investments and not to the outcomes of the projects which were announced. It is well known and documented that some investments will be realized and some are intended but not realized (Minzberg 1978). A future study could examine this issue.

## Acknowledgements

An earlier version of this paper was presented at the 11th European Conference on Information Technology Evaluation (ECITE 2004), Amsterdam, Netherlands, November 11-12, 2004. We are very grateful to the participants of this conference who provided helpful comments.

## References

Brynjolfsson, E. (1993) The Productivity Paradox of Information Technology. *Communications of the ACM* 36, 67-77.

Carr, N.G. (2003) IT Doesn't Matter. *Harvard Business Review* 81, 41-49.

Chatterjee, D., Pacini, C., & Sambamurthy, V. (2002) The Shareholder-Wealth and Trading-Volume Effects of Information-Technology Infrastructure Investments. *Journal of Management Information Systems* 19, 7-42.

Clemson, E.K. & Row, M.C. (1991) Sustaining IT Advantage: The Role of Structural Differences. *MIS Quarterly* 15, 275-292.

Cooper, R. (1987a) The Two-Stage Procedure in Cost Accounting - Part One. *Journal of Cost Management* 1, 43-51.

Cooper, R. (1987b) The Two-Stage Procedure in Cost Accounting - Part Two. *Journal of Cost Management* 1, 39-45.

Cooper, R. (1988) The Rise of Activity-Based Costing - Part One: What is an Activity-Based Cost System? *Journal of Cost Management* 2, 45-54.

Cooper, R. (1989) The Rise of Activity-Based Costing - Part Three: How Many Cost Drivers Do You Need, and How Do You Select Them? *Journal of Cost Management* 2, 34-46.

Cooper, R. & Kaplan, R.S. (1988) Measure Cost Right: Make the Right Decisions. *Harvard Business Review* 96-102.

Cooper, R. & Kaplan, R.S. (1991) Profit Priorities from Activity-Based Costing. *Harvard Business Review* 69, 130-135.

Dehning, B., Richardson, V.J., & Zmud, R.W. (2003) The Value Relevance of Announcements of Transformational Information Technology Investments. *MIS Quarterly* 27, 637-656.

Dos Santos, B., Peffers, K., & Maurer, D. (1993) The Impact of Information Technology Investment Announcements on the Market Value of the Firm. *Information Systems Research* 4, 1-23.

Fama, E.F. (1970) Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance* 25, 383-417.

Fama, E.F. (1991) Efficient Capital Markets: II. *The Journal of Finance* 46, 1575-1617.

Fama, E.F., Fisher, L., Jensen, M.C., & Roll, R. (1969) The Adjustment of Stock Prices to New Information. *International Economic Review* 10, 1-21.

Gordon, L.A. & Silvester, K.J. (1999) Stock market reaction to activity-based costing adoption. *Journal of Accounting and Public Policy* 18, 229-251.

Im, K.S., Dow, K.E., & Grover, V. (2001) Research Report: A re-examination of IT Investment and the Market Value of the Firm - An Event Study Methodology. *Information Systems Research* 12, 103-117.

Johnson, H.T. (1987) The Decline of Cost Management: A Reinterpretation of 20th-Century Cost Accounting History. *Journal of Cost Management* 1, 5-12.

Johnson H.T. & Kaplan R.S. (1987) *Relevance Lost: The Rise and Fall of Management Accounting*. Harvard Business School Press, Boston.

Minzberg, H. (1978) Patterns in Strategy Formulation. *Management Science* 24, 934-948.

Ness, J.A. & Cucuzza, T.G. (1995) Tapping the Full Potential of ABC. *Harvard Business Review* 73, 130-138.

Weill, P., Subramani, M., & Broadbent, M. (2002) Building IT Infrastructure for Strategic Agility. *MIT Sloan Management Review* 44, 57-65.

 

# A Framework for the Evaluation of Business Models and its Empirical Validation

**Jean-Paul Van Belle**
**Department of Information Systems, University of Cape Town, South Africa**
jvbelle@commerce.uct.ac.za

**Abstract:** This article describes a proposal for a framework to evaluate and compare enterprise models. It suggests three major categories for grouping the model evaluation criteria: syntactic, semantic and pragmatic analysis. The paper draws on a wide literature to present a large selection of criteria and to operationalise their measurement by means of several possible metrics. As an empirical validation test, a selection of metrics for eight of the criteria has been calculated for fifteen large enterprise models. Their interpretation supports the usefulness and validity of the overall framework. Various attempts at deriving a composite overall quality score are discussed, but there is less confidence in the validity of this approach.

**Keywords:** Enterprise modelling, evaluation framework, system analysis metrics.

## 1. Introduction

The trend towards model-driven development and the growing automation of much of systems programming, increases our reliance on high-quality models. Furthermore, the move from functional "silo" applications to enterprise-wide systems has widened the scope of business domain models to integrated enterprise-wide models. This move is amplified by the need to build enterprise-wide data warehouses to satisfy the managers' desire for integrated information. Modelling, and enterprise modelling in particular, has become serious business.

But, despite the abundance of evaluation and comparison frameworks for development and modelling *processes* (*conf.* methodology engineering), there appears to be a relative dearth of guidance available on how to evaluate the actual *output* of the modelling activity: how does one evaluate a completed enterprise model?

This article proposes a comprehensive, systematic yet interdisciplinary framework for the evaluation of domain models. It also addresses the concern of empirical testing of the proposed framework. This is done by selecting a representative sample of criteria within the framework – those that could most easily be automated – and testing them against fifteen medium-sized enterprise models. The intent is not to present a final and conclusive set of criteria for the evaluation of models, but rather to illustrate the wide range of possible metrics as well as the obstacles which must be given consideration when comparing models.

## 2. Research objective

The purpose of this paper is the empirical validation of a comprehensive framework for the analysis and evaluation of enterprise models. No satisfactory framework for enterprise model evaluation was found in the literature, although a number of candidate frameworks for the evaluation of other "intellectual products" exist. These are used as inputs to guide the building of an integrated and comprehensive framework. Since the development of a theoretical framework does not contribute towards science unless it is accompanied by a verifiable and methodological testing and evaluation of the framework itself, an empirical validation of the framework is also presented. For this, a "test bed" was constructed, consisting of fifteen publicly available enterprise models from different reference disciplines.

The framework includes *intrinsic* qualities (absolute measures that can be computed for one specific model) and *comparative* qualities (relative measures that compare models). Some of these entail a ranking or judgement (better, worse) whereas other measures merely differentiate (e.g. model A is more like model B, whereas models C, D and E form a separate family). Because of practical and methodological reasons, the emphasis was on *static* models. However, a parallel and completely independent research effort developed a similar framework to cater for dynamic models (Taylor 2003). Also, although the framework focuses on enterprise models, it is quite possible to use the framework for the evaluation of models from *other domains* such as embedded systems, or specific functional areas within the enterprise.

Finally, the framework was developed from an *interdisciplinary* perspective. Since enterprise models themselves originate from such diverse sources as systems theory, computer science, ERP, accounting, linguistics and systems engineering, the proposed framework sought to incorporate contributions from those areas and

 

Reference this paper as:
Van Belle J-P (2006) "A Framework for the Evaluation of Business Models and its Empirical Validation" *The Electronic Journal Information Systems Evaluation* Volume 9 Issue 1, pp 31-44, available online at www.ejise.com

others, such as construction architecture, complexity theory and aesthetics.

## 3. Prior research

A rich body of literature exists on evaluation criteria for models. These stem from a variety of reference disciplines: methodology engineering, systems engineering, ontology research, modelling, etc. Criteria can be arranged in a flat (i.e. unstructured sequential) list, in a hierarchical tree structure or in the form of a framework based on some theoretical structuring concept or theory.

Many authors have suggested lists of criteria to evaluate models (e.g. Benyon 1990:66, Korson & McGregor 1992, Claxton & McDougall 2000, Halpin 2001). Additional criteria are those that define high-quality data, such as those proposed by Orli (1996) and Courtot (2000). Ontology researchers have also proposed their own criteria for enterprise models (e.g. Fox 1998, van Harmelen 1999, Noy 2001). Williams (1996) lists 45 requirements for enterprise reference architectures.

A few authors go beyond unstructured lists and organise their evaluation criteria into frameworks, usually in the context of evaluating the quality of modelling approaches and methodologies (e.g. Khaddaj 2004, Brazier 1998). Structures which organise the different criteria can also be matrix presentations of quality factors; these are often used in a software engineering context: McCall's quality factors, the very similar Böhm model and Gillies' hierarchical quality model (Böhm 1976, Gillies 1997).

Most frameworks suffer from the *grounding problem*: they lack an underlying theoretical or philosophical basis for the framework dimensions. Although it is acknowledged that the frameworks may still be valuable and valid – as long as they are based on the principles of soundness and completeness – some authors (e.g. Frank 1999) stress the value and importance of a strong theoretical grounding for an evaluation framework. The framework proposed below roots itself firmly in the discipline of semiotics and draw on a fundamental distinction which has proved valuable in many different contexts.

## 4. The model evaluation framework

The proposed framework will be developed in the following sequence. Firstly, the two major dimensions, or the conceptual structuring principles, will be explained and motivated. Next, the framework will be "populated" with evaluation criteria. Finally, some sample metrics will be presented for the various criteria.

### 4.1 Framework dimensions or structuring principles

The first and major dimension is grounded in linguistics, information and communication theory and semiotics. The key distinction used in the framework is the fact that all models – or indeed any informational object – have a syntactic, semantic and pragmatic aspect (see Table 1).

- *Syntax* refers to the type of constructs and the legal ways of combining them i.e. the physical or logical tokens which represent the information..
- *Semantics* refers to the meaning: the sense of the information obtained by interpreting the token or signifier.
- *Pragmatics* refers to the context i.e. considerations, issues and background information which influences or moderates the interpretation of the information.

**Table 1:** Main classification dimension of proposed analysis framework.

| Classification concept | Related terms and mappings |
|---|---|
| Syntax | Symbols, form, shape, structure |
| Semantics | Meaning, denotation, sense |
| Pragmatics | Background, situation, context |

The *syntactic* analysis deals with the structural model relations, i.e. shape and form of the entities and their relationships and groupers. It treats all entity and relationship names, therefore, as mere "alphanumeric labels". The *semantic* analysis of models is concerned with the intrinsic *meaning* of the model i.e. the relationship with and mapping to the underlying domain reality that the model represents. The essence of semantic analysis is to unravel the *meaning* of the name (label, word, token) used for a specific model element (entity, relationship, grouper). Put another way, semantic analysis is concerned with the correspondence (mapping, projection, validity) between the model (abstract or intellectual construct) and the underlying domain (reality). Whereas syntactic analysis is fairly technical and easy to automate, semantic analysis involves the more tricky matters of meaning and interpretation and thus lends itself not quite as easily to objective and/or automated analysis. *Pragmatic* model analysis requires the consideration of information regarding the use, environment or context of the model i.e. information outside the model. The analysis techniques falling under this heading include the face validity, degree of use, authority of model author, availability, cost, flexibility, adaptability, model currency, maturity and support. Most analysis relies on the searching and ranking of certain specific information details, often involving a degree of subjective interpretation and an understanding of business issues.

Interestingly, Fabbrini (1998) used the same distinction between semantics, syntax and pragmatics in the context of assessing the quality of a software architecture, but he interpreted the terms quite differently and did not operationalise their approach in any practical way. Equally, Brinkkemper (1996:214) makes reference to the doctoral research of John Venable whereby, in his search for a theoretical basis for the CoCoA methodology, he distinguished his criteria for the evaluation of conceptual data models between those relating to semantic concepts, syntactic constructs and more generic requirements such as views and modularization. These are by no means the only explicit references of the three categories in the context of information systems analysis: the fairly well-known publications from both Benyon (1990) and Stamper (1987) made references to these distinctions, and most computer science research in formal (programming) languages uses the distinction between syntax and semantics extensively. In particular, Stamper proposes his *semiotic framework* to classify information on six different levels: three on the human information level: social world, pragmatics, semantics; and another three on the IT platform: "syntactics", empirics and physical world (1995). He also maintains that too much of the research focuses on the "syntactic" elements. Finally, a parallel but independent research effort concerning the quality of process models uses the same framework distinctions between syntactic, semantic and pragmatic issues [Taylor 2003].

A second organising principle, summarized in Table 2, is not as clear-cut and is presented here mainly as an ordering mechanism within each column of the framework. It will not be explored further and is left for future research purposes. It is proposed that, within each category above, measures range from "absolute" to "relative". For some criteria factors, an absolute value can be established ("more is better") whereas for others, no ideal value can be determined *ex ante*, since the desired target value depends on factors not directly related to the intrinsic nature of the model. It must be recognized that this distinction is not a categorical classification but rather a continuum: there is a fuzzy cross-over area in the middle where it is relatively difficult (or arbitrary) to decide whether a given quality measure is absolute or relative.

**Table 2:** Possible second dimension for proposed analysis framework.

| Classification concept | Related terms and mappings |
|---|---|
| Absolute measures | Theoretical; "das Model an Sich"; the model as object; objective standards; intrinsic qualities; technical factors; "Conforms to specification"; computer science; academic. |
| Relative measures | Applied; "das Model für Uns"; the model as subject; subjective standards; extrinsic qualities; business factors; "Fit for purpose"; information systems; practitioner. |

## 4.2 Populating the framework with detailed criteria

Table 3 lists the proposed model evaluation criteria within the framework structure. Criteria are grouped into conceptual clusters, which can be sub-divided if one wishes. Note that only simple or primitive evaluation criteria are listed here. Composite concepts consist of sub-criteria and can therefore often not be allocated to one specific framework cell. An example is *usability*, which includes all of the pragmatic and many semantic and syntactic criteria. Another example is model *dependability*, as discussed by Barbacci (1995).

The various criteria are drawn from different sources, which often ascribe different meanings to certain criteria and, conversely, different authors sometimes use different terms to describe a similar concept. To indicate this overlap in meaning, criteria were grouped into "clusters".

**Table 3:** Populated framework for model analysis.

| | Syntactic | Semantic | Pragmatic |
|---|---|---|---|
| **Absolute** | Size | Genericity: universality & technical independence | Validity: authority & user acceptance |
| | Correctness; error-free; integrity; consistency | Completeness (domain coverage); conciseness; efficiency | Flexibility; expandibility; portability; adaptability |
| | Modularity; structure; hierarchy. | Expressiveness<br>Similarity and overlap with other models | |
| **Relative** | Complexity; density | Perspicuity; comprehensibility; understandability; self-descriptiveness | Price; cost; availability<br>Support |
| | Architectural style | Documentation | Purpose; goal; relevance; appropriateness |

### 4.3 Proposed metrics for the evaluation criteria.

The empirical validation of the framework requires that the each of the evaluation criteria can be measured or calculated. In order to operationalise the framework, many measures were adopted from the rich literature whilst others were newly developed. Although many more were calculated than shown, only the ones that were found to display some validity are listed in Table 4 which also presents the final proposed version of the framework.

**Table 4:** Summary of validated framework metrics and measures.

| | Criterion | Suggested metric / measure |
|---|---|---|
| SYNTACTIC | Size | CASE (concept) count and adjusted CASE count |
| | Correctness; error-free; integrity; consistency | Syntax error, consistency and standards level score |
| | Modularity | Number of groupers, group levels and diagrams |
| | Structure; hierarchy | Multiple inheritance; mean inheritance depth, reuse ratio. |
| | Complexity; density | Relative connectivity; average fan-out; plot of Fruchterman-Reingold (for similar-sized models); harmonic mean of fan-out; fan-out distribution (chart); fan-out model signature. |
| | Architectural style | Layout aesthetics |
| SEMANTIC | Genericity | % mapping to domain |
| | Coverage | Domain coverage score; core concept coverage |
| | Completeness | Ranking of absolute lexicon coverage |
| | Efficiency; conciseness | Relative lexicon coverage |
| | Expressiveness | Average expressiveness score |
| | Similarity & overlap with other models | Plot of similarity coefficients; most similar neighbours; similarity dendogram; most important concepts. |
| | Perspicuity; comprehensibility; understandability; readability | Normalised rank-adjusted weighted perspicuity count based on user lexica |
| | Documentation | Completeness, extensiveness, readability (Flesh Reading Ease score) |
| PRAGMATIC | Validity: authority & user acceptance | Academic author citations |
| | Flexibility; expandability; adaptability | Composite flexibility score |
| | Currency; maturity | Descriptive table & taxonomy |
| | Purpose; goal; relevance; appropriateness | Descriptive table |
| | Availability | Medium & status |
| | Cost | Purchase cost |
| | Support | Tool & vendor support, user base |

Formatted: Bullets and Numbering

### 4.4 Sample of enterprise models

In order to demonstrate the feasibility of using the framework to evaluate real-world, industrial-strength systems, fifteen medium-sized to large enterprise data models were captured. Models had to have at least 100, preferably more than 200 entities and needed to be publicly available. They are grounded in a wide variety of reference disciplines, but the prototypical example of an enterprise data model is the model underlying Enterprise Resource Planning (ERP) systems.

The following gives a brief overview of the generic enterprise models which were selected, grouped according to their reference discipline (Van Belle, 2002a). The database with the models is available in XML format for research purposes on request.

Two ERP models underlying the leading integrated enterprise applications were captured: the *SAP* R/3 (Sheer, 1998) and *BAAN* IV. The latter was "re-engineered" from Perreault (98). Four generic data model libraries were captured from their respective published books: *Hay* (1996) and *Silverston* (1997; 2001); Marshall's *BOMA* (Marshall, 2000) and *Fowler*'s analysis patterns (Fowler, 1997). Two smallish academic enterprise reference models were found: *Purdue*'s Reference Model for CIM and *ARRI*'s Small Integrated Manufacturing Enterprise Model in IDEF0 notation (Williams, 1991) in DFD notation. As an example of a data warehousing model, *Inmon*'s set of high and mid-level data models was gleaned from his website. Two framework derived models were *AKMA*'s Generic DataFrame and IBM's San Francisco "*SanFran*" (predecessor of WebSphere). Finally, three enterprise ontologies were selected: the Enterprise Ontology developed by the *AIAI* in Edinburgh (Uschold, 1998), *TOVE* from EIL in Toronto and a subset of the *CYC* Upper Ontology was created containing all organisation and enterprise-related concepts.

 

Although the meta-model used for capturing the above models contains 7 meta-entities and numerous meta-attributes, the only three meta-entities used below are the meta-concepts of "*entity*" (sometimes called *object* or *concept*; their graph-analytic equivalent is a *node* or *vertex*), the "*grouper*" construct and "*relationship*" (equivalent to a *connection, link, edge* or *arc*). The latter can be subdivided into hierarchical structuring relationships of the "IS-A" type (reflecting specialisation/generalisation) and "proper" domain relationships reflecting some semantic or domain-inspired link between entities.

## 4.5 Syntactic analysis

The syntactic analysis deals with the purely structural aspects of the model. Its analysis techniques are mainly derived from software engineering and computer science. This includes a large variety of standard syntactic metrics relating to size, grouping, layering, inheritance structure, and network visualisation, as well as some less standard metrics such as interface aesthetics (i.e. the visual beauty or tastefulness of graphically drawn models). Perhaps the most obvious criterion is model size. Many *size* measures were calculated of which three are listed in table 5: the total number of entities or classes in the model, the CASE size or concept count (the number of entities + relationships + grouper elements) and an expanded concept count whereby additional meta-modelling elements such as attributes are also taken into account. Although the measures are very correlated coefficient, the expanded CASE size is preferred because it favours the more fully specified models above the shallower models. For instance, the badly specified *Inmon* model now drops its relative ranking, although there is still no accounting for model documentation and description. It is perhaps somewhat surprising to see the SAP and BAAN models rank behind the data models published by Hay and Silverston but both the CYC and TOVE ontologies are indeed very sizeable.

A second criterion is model *correctness.* Models were rated using a composite "correctness score" consisting of a score for the amount of errors, the degree of (in)consistency and the use of and adherence to notational standards. A typical correctness issue is two diagrams showing different cardinalities for a given relationship between the same entities .

Formatted: Bullets and Numbering

The "scoring system" used for both errors and consistency problems gave 3 for no problems or errors, 2 for minor, 1 for medium and 0 for major problems. In addition, a score between 0 up to 2 was allocated for adhering to stringent standards for naming, diagramming etc. The combined "correctness score" thus ranges from 0 to 8; a higher the score indicating a more correct model. Not surprisingly, the well-validated SAP model achieves the highest score with most of the other well-known models following immediately after. Some lesser-known individual research models obtain relatively low scores.

**Table 5:** Syntactic model analysis.

| | Size | | | Correctness | | | | Complexity | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Nr of Entities | CASE Size | Expanded CASE Size | Accuracy | Consistency | Standards | Combined Score | Cyclomatic Complexity | Relative Connectivity | Average Fan-Out | De Marco's Data Bang | Average Data Bang | (Harmonic) Mean Fan-out |
| AIAI | 94 | 270 | 510 | 2 | 3 | 2 | 7 | 30 | 1.82 | 3.32 | 220 | 4.99 | 1.81 |
| AKMA | 82 | 565 | 769 | 1 | 2 | 2 | 5 | 6 | 1.15 | 2.18 | 160 | 2.85 | 1.54 |
| ARRI | 128 | 430 | 790 | 2 | 2 | 1 | 5 | 79 | 2.09 | 3.31 | 592 | 4.97 | 1.81 |
| BAAN | 328 | 1086 | 1927 | 2 | 2 | 2 | 6 | 377 | 2.29 | 5.24 | 2018 | 8.7 | 2.23 |
| BOMA | 183 | 552 | 770 | 3 | 2 | 2 | 7 | 65 | 1.68 | 3.00 | 557 | 4.35 | 1.81 |
| CYC | 777 | 2623 | 4537 | 2 | 2 | 2 | 6 | 511 | 2.32 | 3.60 | 3507 | 5.49 | 1.94 |
| Fowler | 120 | 375 | 579 | 2 | 2 | 2 | 6 | 37 | 1.67 | 2.76 | 372 | 3.92 | 1.71 |
| Hay | 291 | 1292 | 3465 | 2 | 3 | 2 | 7 | 491 | 3.13 | 6.17 | 2470 | 10.5 | 2.42 |
| Inmon | 427 | 2429 | 2670 | 1 | 1 | 1 | 3 | 17 | 1.08 | 2.14 | 682 | 3.03 | 1.22 |
| NHS | 269 | 751 | 1460 | 0 | 3 | 2 | 5 | 48 | 1.7 | 2.54 | 622 | 3.6 | 1.52 |
| Purdue | 106 | 343 | 866 | 0 | 0 | 1 | 1 | 136 | 2.11 | 5.03 | 711 | 7.99 | 3.82 |
| SAP | 396 | 1218 | 1917 | 3 | 3 | 2 | 8 | 285 | 1.97 | 3.73 | 1851 | 5.64 | 2.30 |
| SanFran | 109 | 332 | 532 | 1 | 0 | 0 | 1 | 74 | 1.68 | 3.47 | 520 | 5.25 | 1.95 |
| Silverston | 267 | 1269 | 2235 | 2 | 3 | 2 | 7 | 114 | 1.51 | 3.08 | 950 | 4.55 | 1.76 |
| TOVE | 564 | 1937 | 2042 | 2 | 1 | 2 | 5 | 678 | 2.28 | 4.51 | 3876 | 7.19 | 2.33 |

Model *complexity* has many interpretations: Edmonds (1999) calculates 48 different syntactic complexity metrics. In an attempt to minimize the influence of "size", model complexity in this framework refers to the average "density" of the model network. Of the various *complexity* metrics which were calculated for the models; Table 5 lists only McGabe's (1976) cyclomatic complexity, relative connectivity (including the inheritance structure), average fan-out; and De Marco's data bang (Shepperd 1995). However, none of these appear to convey the subjective feeling of model network density well. However, *the frequency distribution of the entity fan-outs* for each model were found to yield a distinctive and characteristic signature of the underlying model complexity (see Van Belle, 2002b). The most descriptive statistic for this average "density" of the model network turned out to be the *(harmonic) mean of the fan-out distribution*.

Inmon's low complexity is typical of data warehousing conceptual models whereas the ontologies and financial models score relatively high. Since it is relatively easy for a small, compact model to achieve relatively high complexity (e.g. Purdue), the harmonic mean fan-out is best compared among similarly sized models e.g. SAP and Baan exhibit similar complexity as do the fairly comparable BOMA, Fowler and Silverston.

## 4.6 Semantic analysis

Semantic model analysis refers to the relationship of a model with the domain it is representing. This type of analysis is grounded in linguistics, ontology research and lexicography. Much of the analysis concentrates on similarity, correspondence and cluster analysis. It proved to be a challenge to eliminate subjectivity from the metrics thus preference was given, where feasible, to automatically calculated or computer-generated measures.

Perhaps the most straightforward criterion is the *expressiveness* of a model. This depends on the richness of the modelling language used and is determined by looking at the number of (different) modelling language constructs used by each model. An expressiveness score can be calculated as the weighted index of the number of meta-model attributes covered in a model. In our case, the metric used the following expressiveness qualities (unless otherwise specified, a weight of 1 was applied): degree of formality (weight of 3); diagrams; directed graph; use of generalisation; depth of inheritance tree; multiple inheritance; number of grouper levels (x2); entity definitions; entity examples; entity attributes; relationship names; relationship role names (x2); relationship cardinalities; relationship types; definitions for relationships/groupers; and the use of additional constructs such as constraints.

It is no surprise that the three ontologies in the sample (CYC, TOVE and AIAI) – which use semantically very rich languages - and the object-oriented models (BOMA, SanFran) all score very high. The exact composition of the expressiveness metric can be modified to suit the ultimate requirements of the model analysis.

Model *perspicuity* and *readability* refer to the extent to which the model can be understood or comprehended by the intended users or readers of the model and how self-describing the model is. The perspicuity analysis was based on matching all model element names against common domain vocabulary lists. Although several business lexicons were investigated, the use of Someya's (1999) well-validated corpus-based business language list annotated with word frequency statistics yielded the most valid results, especially if slightly more sophisticated wordlist preparation and matching algorithms are used. The GPC ("Gross Perspecuity Count") measures what percentages of model element labels exist in the business word list, whereas the NRAWPC Normalizes for the size of the model, is Rank-Adjusted for word use frequency and applies a Weighting to concatenated or multiple word labels.

Formatted: Bullets and Numbering

**Table 6:** Semantic and pragmatic model analysis.

| | Expressiveness | | Perspicuity | | Completeness | | Authority | | Flexibility | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Model** | **Raw Expressiveness** | **Weighted Expressiveness** | **GPC** | **NRAWPC** | **Completeness1** | **Completeness2** | **Google Page-Rank™ for model** | **PageRank™ for organisation URL** | **Digitally available?** | **Customizable/ reusable** | **Implementation independence** | **Overall flexibility score** |
| **AIAI** | 10.5 | **13.0** | 85% | **68%** | 80 | **272** | 7 | **7** | Yes | Some | Low | **1.50** |
| **AKMA** | 8.8 | **10.0** | 94% | **75%** | 74 | **233** | 5 | **5** | No | No | High | **1.00** |

 

| | Expressiveness | | Perspicuity | | Completeness | | Authority | | Flexibility | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **ARRI** | 7.3 | **8.5** | 86% | **77%** | 121 | **346** | 3 | **5** | No | No | Med. | **0.50** |
| **Baan** | 6.3 | **8.0** | 95% | **81%** | 235 | **636** | 0 | **7** | Yes | Some | Med. | **2.25** |
| **BOMA** | 9.7 | **12.0** | 91% | **77%** | 156 | **452** | 0 | **4** | Yes | Some | High | **2.25** |
| **CYC** | 9.5 | **12.0** | 89% | **74%** | 590 | **1143** | 6 | **7** | Yes | Some | Med. | **2.25** |
| **Fowler** | 6.8 | **8.5** | 88% | **68%** | 100 | **336** | 0 | **7** | No | Yes | High | **2.00** |
| **Hay** | 9.3 | **12.0** | 93% | **76%** | 201 | **574** | 5 | **5** | No | Yes | High | **2.00** |
| **Inmon** | 6.5 | **7.5** | 91% | **76%** | 356 | **840** | 4 | **6** | No | Some | Med. | **1.00** |
| **NHS** | 8.3 | **9.0** | 86% | **70%** | 144 | **398** | 4 | **6** | Yes | No | Med. | **1.50** |
| **Purdue** | 6.5 | **7.5** | 93% | **79%** | 116 | **383** | 4 | **6** | No | Limited | Med. | **0.75** |
| **SanFran** | 7.7 | **9.5** | 90% | **76%** | 99 | **310** | 5 | **9** | Yes | Yes | High | **3.00** |
| **SAP** | 8.8 | **10.5** | 94% | **82%** | 236 | **632** | 0 | **8** | Yes | Some | Med. | **2.25** |
| **Silverston** | 8.8 | **11.5** | 95% | **81%** | 141 | **461** | 0 | **5** | Yes | Yes | High | **3.00** |
| **TOVE** | 9.5 | **12.5** | 77% | **60%** | 226 | **571** | 4 | **6** | Yes | Some | Med. | **2.00** |

It is again encouraging to find the measure validated by the high ranking of the well-known ERP models as well as the published data models. Models with obscure or even obtuse language score very low. It is, of course, possible to adopt a more simplistic approach by using a more general and more easily computable readability index such as a Flesh Reading Ease or Flesch-Kincaid Grade Level score. However, the results of this were found to be far less valid (see Van Belle 2004).

A third criterion attempts to measure the semantic equivalent of the syntactic size metric: model *completeness.* It measures how much of the domain has been covered. It requires an accepted and complete description of the model domain against which each model can be mapped. Since this is hardly ever available, an approximation will normally be used. Here we used the same business lexicons as were used for the perspicuity analysis and calculated how many distinct concepts within the business lexicon are covered – the more words or concepts that were covered, the better. The analysis was then enhanced by using an intermediary translation process with synonyms as found in WordNet ("completeness2") to enable the mapping of meanings instead of word tokens. Not surprisingly, the larger models tend to be more complete although some interesting observations can be made. The ERP and data warehousing models cover most of the business domain. But the second-largest (syntactic size) model, *Hay*, drops dramatically in position when looking at completeness. This is a strong motivation in favour of complementing standard syntactic measures with semantically based metrics.

## 4.7 Pragmatic analysis

Pragmatic model analysis, as defined above, is concerned with criteria which cannot be assessed purely on the basis of the information contained within the model, but which require the consideration of information regarding the use, environment or context of the model i.e. information outside the model. Unfortunately, most pragmatic criteria involve a substantial degree of subjectivity. Also, by definition, many criteria depend on the actual purpose for which the evaluation is carried out. For this reason, the discussion is limited to the two pragmatic measures which are believed to be more objective and universally applicable. In a real-world model analysis, several other criteria are likely to be important and should be included on a as-needed-basis, e.g. cost, tool support etc.

Model *authority* refers to the acceptance of the model by practitioners in the field. Depending on the publishing medium of the model, metrics that are reasonably easy to collect are the relative sales ranking for book-based models (e.g. by Amazon.com), and popularity of a web page by counting the number of external hyperlinks to the page for web-based models, e.g. using the Google PageRank™ system. Another pragmatic proxy used in the commercial world is *authoritative validity* i.e. the reputation of the authoring person, team or organisation. This can be measured by the number of author citations which measures and ranks the academic standing or authority of the lead author associated with the model relatively accurately, especially within the same reference disciplines. Table 6 lists the PageRank metrics for both the webpage listing the model as well as the home page of the author (organisation). The highest authority is accorded to large commercial organisations such as IBM (SanFran) and the ERP systems, as well as the better known research organisations (ontologies).

Model *flexibility* is concerned with how well models can be changed or *adapted* to different situations. A composite flexibility measure was calculated which awarded a score for three aspects of flexibility: the model's availability in digital format, its customisability or reusability and its implementation independence. Models designed specifically as templates for model

**Formatted:** Bullets and Numbering

development such as Silverston and SanFran scored maximally. The ontologies and ERP models also scored high, whereas the more obscure academic models (ARRI, Purdue) tended to score well below average.

## 4.8 Composite model evaluation score

It is a natural question to investigate whether an overall score can be calculated, combining the values of the various component metrics to capture, perhaps, the essence of overall model quality. In an attempt to achieve this, those metrics which appeared to be the most valid or representative measure for each criterion in this research were selected from tables 5 and 6. These were expanded CASE size, correctness, harmonic mean of fan-out for complexity, weighted expressiveness, NRAWPC for perspicuity, completeness2, Google PageRank™ for website authority and the flexibility composite score as explained above. In each of these eight measures, each model was given a ranking from 1 to 15 against the other models based on its scores, giving rise to a vector of 8 ranking scores for each model.

This vector of rankings for a given model was processed in four different ways:

- The *average* ranking could be calculated.
- A model's *median* ranking was calculated
- The number of times that a model was ranked in the "bottom *half*" of the models was deducted from the number of times that it featured in the "top half" of the models.
- The number of times that a model was ranked in the top *quartile* minus the number of times it found itself in the bottom quartile.

Formatted: Bullets and Numbering

Note that this procedure ignores relative weighting of criteria as well as the absolute differences in the underlying scores. However, this section is merely intended as a first-cut evaluation of the overall feasibility of calculating a composite score. Table 6 details the results. The section with "calculated scores" gives the raw values of the above four calculations. Since these scores are not normalised, they provide little information. Therefore, the last four columns convert the raw scores into a relative position by ranking each model from 1 to 15.

**Table 7:** Composite model ranking.

| | Calculated (or Raw) Scores | | | | Overall Ranking based on … | | | |
|---|---|---|---|---|---|---|---|---|
| **Model** | **Average Rank** | **Median Rank** | **#[R<8] - #[R>8]** | **#[R<5] - #[R>11]** | **Average Rank** | **Median Rank** | **#[R<8] - #[R>8]** | **#[R<5] - #[R>11]** |
| **AIAI** | 8.3 | 9 | -1 | 0 | **9** | **9** | **10** | **8** |
| **AKMA** | 11.3 | 11.5 | -7 | -4 | **15** | **14** | **15** | **15** |
| **ARRI** | 10.0 | 10.5 | -5 | -1 | **14** | **13** | **13** | **9** |
| **BAAN** | 5.1 | 4 | 6 | 3 | **4** | **3** | **2** | **5** |
| **BOMA** | 6.9 | 6.5 | 2 | 2 | **7** | **6** | **7** | **6** |
| **CYC** | 4.4 | 3 | 6 | 5 | **2** | **1** | **2** | **2** |
| **Fowler** | 9.6 | 11.5 | -2 | -3 | **12** | **14** | **11** | **14** |
| **Hay** | 4.9 | 4 | 6 | 4 | **3** | **3** | **2** | **3** |
| **Inmon** | 9.1 | 9.5 | 0 | -2 | **10** | **10** | **8** | **12** |
| **NHS** | 9.9 | 9.5 | -5 | -2 | **13** | **10** | **13** | **12** |
| **Purdue** | 9.1 | 9.5 | -2 | -1 | **10** | **10** | **11** | **9** |
| **SAP** | 3.6 | 3.5 | 8 | 6 | **1** | **2** | **1** | **1** |
| **SanFran** | 8.1 | 8 | 0 | -1 | **8** | **8** | **8** | **9** |
| **Silverston** | 5.5 | 5 | 4 | 4 | **5** | **5** | **5** | **3** |
| **TOVE** | 6.8 | 6.5 | 4 | 1 | **6** | **6** | **5** | **7** |

The first observation is that the rankings are relatively robust, regardless of which procedure is used. In fact, it was found that even including additonal measures (i.e. more than 8 metrics) does not necessarily affect the relative ranking of a model significantly (Van Belle, 2004).

A more detailed evaluation can be made by looking at the overall rankings of the models, *regardless of reference discipline*. The *top scoring* models are: Scheer's SAP reference model, CYC's enterprise sub-ontologies, Hay's data model and Baan's reference model. From a subjective evaluation of these models, based on their study throughout this research, a strong case can be presented that they indeed represent the best overall models in the database, not only in their respective disciplines but also from an overall quality perspective. It is a particularly interesting vindication (and validation) of the framework that these models originate from three fairly different reference disciplines, yet the framework allows an interdisciplinary comparison in spite their fundamentally different modelling

 

approaches and philosophies. They are followed by a set of "close seconds", namely Silverston's data models, the TOVE ontology and Marshall's BOMA patterns.

Similarly, the models ranked at the *bottom of the scale* are indeed the "worst" models from a qualitative perspective. AKMA is a fairly shallow model. ARRI is a tiny model, focussing on a manufacturing environment. NHS is a vertical model which is very specific to the health care industry. Although large, Inmon appears to be an inconsistent and extremely shallow model. Again, it is interesting that the framework manages to identify problematic models, regardless of the underlying reference discipline.

An alternative way of analyzing the rankings is by comparing the relative positions of models *within the same reference discipline*. Between the two ERP models, SAP fairly consistently beats or matches Baan across various criteria. It must be stressed that the Baan model here is not the original conceptual model but a model re-engineered from a description of its relational implementation. This guarantees a quality handicap so the results should not be read as reflecting on the actual ERP packages. Nevertheless, there is support for the contention that the SAP model has a better theoretical foundation as well as represents significantly more analysis effort.

Comparing the data models, it must be admitted that the quality difference between Silverston and Hay is a tough call. However, of the more "pattern-like" models, BOMA is definitely cleaner than SanFran, bearing in mind that the author experienced significant methodological problems in capturing SanFran. The lagging position of Fowler, however, is very debatable and it may well have been penalized because of its highly conceptual nature.

The comparative scores for the ontology-based models correspond perfectly with the amount of ontology engineering and analysis effort invested in each of the models. CYC represents by far the most effort, followed by TOVE and then by AIAI. However, although smaller, AIAI is a more homogenous and conceptually higher level model, which is perhaps not fully reflected in the score. Within the CIM models, it must be recognized that although ARRI is a much cleaner, more correct and rounded model than Purdue, the latter is a better representation of the enterprise domain and represents significantly more modelling effort than ARRI.

This section illustrates that that the interpretation or even overall value of a composite index is limited. This reinforces an early comment from the field of system engineering metrics: "[C]alculating and understanding the value of a single overall metric for […] quality may be more trouble than it is worth. The major problem is that many of the individual characteristics of quality are in conflict; added efficiency is often purchased at the price of portability, accuracy, understandability, and maintainability." (Böhm 1978:ix)

## 5. Conclusion

The overall research objective was to present and empirically validate a framework for the comparative evaluation of enterprise models. This research suggests that the proposed framework is a highly productive and valid approach for evaluating enterprise models. As shown in Addendum 1, almost all of the evaluation criteria which have been suggested in the literature can easily be accommodated by the framework.

The most useful dimension within the framework is the separation of syntactic, semantic and pragmatic criteria or factors. Each of these sets of criteria has a very distinct tone, reflecting a certain paradigmatic approach. Syntactic analysis has a strong pure science and engineering flavour, drawing heavily from computer science metrics, systems engineering graph-theoretical and even architectural concepts. Semantic analysis relies mainly on lexicography and computational linguistics, as well as more conceptual information sciences such as meta-analysis or information frameworks. Finally, pragmatic analysis, focused on practical business or commerce issues such as support, pricing, purpose, organizational impact etc. The framework thus brings together the basic constituent reference disciplines of information systems.

The framework was validated empirically using a sample consisting of fifteen relatively large enterprise models. Eight criteria were selected from the framework for which a number of possible metrics, including some newly proposed ones, were calculated.

Overall, it was quite easy to generate fairly valid metrics for the framework criteria. Interestingly, it was possible to calculate fairly objective metrics for semantic and pragmatic analysis, although care has to be taken with the interpretation of the results.

A combined overall model ranking, intended to represent some type of composite quality index,

appears to have some but fairly limited face validity.

Apart from the value of the framework to *classify* existing criteria, the framework should also be seen as an ideal way for the *creative design or generation* of new criteria or measures. Indeed, many metrics suggested in this research were inspired by thinking about model evaluation along the dimensions identified by the framework.

There is still considerable scope for future research and development of the framework. The development of more refined or alternative metrics, especially for the pragmatic analysis, would be useful. A more theoretically valid approach to combining individual criteria into composite metrics, such as model quality or usability, is also still an open research question, although this may well prove to be an illusive goal. More emphasis could also be placed on structural relationships (such as inheritance) and grouper constructs. Finally, the framework should be empirically validated within other modelling domains.

**Formatted:** Bullets and Numbering

## References


Avison, D.E. & Fitzgerald, G. (1995) Information Systems Development: Methodologies, Techniques and Tools. McGraw Hill: London.

Barbacci, M.R.; Klein, M. H.; Longstaff, T. A. et al. (1995) Quality Attributes. Technical Report 95-TR-021, Software Engineering Institute.

Benyon, D. (1990) Information and Data Modelling. Blackwell: Oxford, UK.

Böhm, B. et al. (1978). Characteristics of Software Quality. Elsevier North-Holland: New York.

Brazier, F.M.T. & Wijngaards, N.J.E. (1998) "A Purpose Driven Method for the Comparison of Modelling Frameworks".In Proceedings of the Eleventh Workshop on Knowledge Acquisition, Modeling and Management, Banff, Canada, 18-23 Apr 1998.

Brinkkemper, S.; Lyytinen, K. & Welke, R.J. (Eds.) (1996) "Method Engineering: Principles of Method Construction and Tool Support" In Proceedings of the IFIP TC8, WG8.118.2 Working Conference on Method Engineering, Atlanta, Georgia, 26-28 Aug 1996.

Chapman & Hall, London.Courtot, T. (2000). What to Look for in Packaged Data Models/Databases. In Proceedings of the Meta Data Conference, Arlington, Virginia, 19-23 Mar 2000.

Claxton, J.C. and McDougall, P.A. (2000). Measuring the Quality of Models. The Data Administration Newsletter, 2000:14. [Online] http://www.tdan.com/i014ht03.htm.

Edmonds, B. (1999) Syntactic Measures of Complexity. Doctoral Thesis, University of Manchester.

Fabbrini, F.; Fusani, M. & Gnesi, S. (1998) "Quality Evaluation based on Architecture Analysis" In Proceedings of the International Workshop on the Role of Software Architecture in Testing and Analysis (ROSATEA'98), Marsala, Italy, 30-Jun to 3-July 1998.

Fowler, M. (1997) Analysis Patterns. Addison-Wesley: Reading (MA).

Fox M.S. & Gruninger M. (1998) "Enterprise Modelling" The AI Magazine, Fall 1998: 109-121.

Frank, U. (1999) "MEMO: Visual Languages For Enterprise Modelling" Arbeitsberichte des Instituts für Wirtschaftsinformatik (Universität Koblenz-Landau) Nr 18.

Gillies, A. (1997). Software Quality: Theory and Management. Thomson: London.

Halpin, T. & Bloesch, A. (1999) "Data Modeling in UML and ORM: A Comparison." The Journal of Database Management, 10 (4): 4-13.

Hay, D.C. (1996) Data Model Patterns. Dorset House: New York.

Khaddaj, S. & Horgan G. (2004) "The Evaluation of Software Quality Factors in Very Large Information Systems" Electronic Journal of Information Systems Evaluation, 7 (2): 43-48.

Korson, T. & McGregor, J.D. (1992) "Technical Criteria for the Specification and Evaluation of Object-Oriented Libraries" Software Engineering Journal, 7 (3): 85-04.

Marshall, C. (2000) Enterprise Modelling with UML. Designing Successful Software through Business Analysis. Addison-Wesley, Reading (MA).

McGabe, T.J. (1976) "A Software Complexity Measure" IEEE Trans. Software Engineering, 2 (Dec 1976): 308-320.

Ngo, D.; Chek L., Teo, L. et al. (2000) A Mathematical Theory of Interface Aesthetics. Unpublished working paper. [Online] http://www.mi.sanu.ac.yu/vismath/ngo/.

Noy, N.F. & McGuinness, D.L. (2001) Ontology Development 101: A Guide to Creating Your First Ontology. SMI technical report SMI-2001-0880.

Orli, R.; Blake, L.; Santos, F. & Ippilito, A. (1996) Address Data Quality and Geocoding Standards. Unpublished report. [Online] http://www.kismeta.com/Address.html.

Oxford. (1979) The Oxford Paperback Dictionary. Oxford University Press: Oxford.

Perreault, Y. & Vlasic, T. (1998) Implementing Baan IV. Que: Indianapolis, Indiana.

Scheer, A.-W. (1998) Business Process Engineering. Reference Models for Industrial Enterprises. Springer-Verlag: Berlin (2nd ed).

Shepperd, M. (1995) Foundations of Software Measurement. Prentice-Hall: London.

Silverston, L., Inmon W.H. & Graziano, K. (2001) The Data Model Resource Book. A Library of Universal Data Models For All Enterprises. J. Wiley: New York (2nd ed).


 

Someya, Y. (1999). A Corpus-based Study of Lexical and Grammatical Features of Written Business English. Masters Dissertation, Dept of Language and Information Sciences, University of Tokyo.
Stamper, R. (1997) "Semantics. Critical Issues" in Information Systems Research, Boland, R.J. & Hirschheim, R.A. (Eds.), J. Wiley: Chichester, 43-78.
Taylor, C. & Sedera, W. (2003) "Defining the Quality of Business Process Reference Models" in Proceedings of the 14th Australasian Conference on Information Systems (ACIS), Perth, 1-3 Dec 2003.
Uschold, M., King, M., Moralee, S. & Zorgios, Y..(1998) The Enterprise Ontology. The Knowledge Engineering Review, Vol. 13. Special Issue on Putting Ontologies to Use.
Valente, A. & Breuker, J. (1996). Towards Principled Core Ontologies. In Proceedings of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems, Banff, Canada, Nov 1996.
Van Belle, J.P. (2002a) A Survey of Generic Enterprise Models. Proceedings of the 32nd Annual SACLA Conference, Port Elizabeth (South Africa), June 2002.
Van Belle, J.P. (2002b) Towards a Syntactic Signature for Domain Models: Proposed Descriptive Metrics for Visualizing the Entity Fan-out Frequency Distribution. Proceedings of the SAICSIT Conference, Port-Elizabeth (South Africa), Sep 2002.
Van Belle, J.P. (2004) "A Framework to Evaluate the Quality of Information System Models." Ingénierie des Systèmes d'Information, Special Issue on IS Quality, 9 (5) (in publication).
Van Harmelen, F. & Fensel, D. (1999) "Practical Knowledge Representation for the Web" in Proceedings of the IJCAI-99 Workshop on Intelligent Information Integration, Stockholm, Sweden, 31 Jul 1999.
Williams, T.J. (1996) "The Needs of the Field of Integration" in Architectures for Enterprise Integration. Bernus, P.; Nemes, L. & Williams, T. (ed). Chapman & Hall, London: 21-31.
Williams, T.J. (ed). (1991) A Reference Model For Computer Integrated Manufacturing (CIM). A Description from the Viewpoint of Industrial Automation. CIM Reference Model Committee, International Purdue Workshop on Industrial Computer Systems, the Instrument Society of America: Research Triangle Park (North Carolina).

## Addendum: Evaluation Criteria Mapped Against the Framework

The following table gives an overview of some of the model evaluation criteria found in the literature and how they are mapped into the proposed model evaluation framework categories (left three columns).

| Model characteristic | Benyon 1990 | Fox 1993 | Fox 1998 | Valente 1996 | Courtot 2000 | Chen 1998 | Orli 1996 | Claxton 2000 | Witt 1994 | Crockett 1991 | Powel 1996 | Halpin 2001 | Frank 1999 | Korson 1992 | V.Harmelen 1999 | Brazier 1998 | Williams 1996 | Noy 2001 | Swartout 1996 | McGall 1997 | Boehm 1997 | Gillies 1997 | Syntactic | Semantic | Pragmatic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Abstract basic concepts | | | | ■ | | | | | ■ | | | ■ | | ■ | | ■ | | | | | | | ▨ | ■ | ▨ |
| Adheres to standards | | | | | | ■ | | | | | ■ | | | | | | ■ | ■ | | ■ | | | ■ | ▨ | ▨ |
| Aesthetics/appeal | | | | | | | | | | | | | | | | | | | | | | ■ | ■ | ▨ | ▨ |
| Alternatives/comparison | | | | | | ■ | | | | | | | | | | | | | | | | | ■ | ■ | ■ |
| Architecture | | | | | ■ | | | | ■ | | | | | | | | | | | | | | ■ | ▨ | ▨ |
| Availability | | | | | ■ | | | | | | | | | | | | | | | | | | ▨ | ▨ | ■ |
| Basis for communication | ■ | | | | | | | ■ | | ■ | | | | | | | ■ | | | | ■ | | ▨ | ■ | ▨ |
| Competent/problem oriented | | ■ | | | | | | | | | | | | | | ■ | | | ■ | | | | ▨ | ▨ | ▨ |
| Computer manipulation | ■ | ■ | ■ | | | | | | | ■ | | | | | | | | | | | | | ▨ | ▨ | ■ |
| Conciseness | | | | | | | | | | | | | | | | | | | | ■ | ■ | | ■ | ▨ | ▨ |
| Consistency (model) | | | | | ■ | ■ | ■ | ■ | | ■ | | | | ■ | | | ■ | | | ■ | ■ | | ■ | ▨ | ▨ |
| Consistency (representation) | ■ | | | | | | ■ | | | ■ | ■ | | | ■ | | | | | | ■ | ■ | | ■ | ▨ | ▨ |
| Controlled vocabulary | | | | | | | | | | | | | | | | | | ■ | | | | | ▨ | ■ | ▨ |
| Coverage/Domain (extent) | | | | | ■ | | | | | | ■ | | | ■ | ■ | ■ | | | | | | | ▨ | ■ | ▨ |
| Docs different levels | | | | | | | | | | | | | | ■ | | | | | | | | | ▨ | ■ | ▨ |
| Docs indexed alpha/keyw/struct | | | | | | | | | | | | | | ■ | | | | ■ | | | | | ▨ | ■ | ▨ |
| Docs organised & structured | | | | | | | | | | | | | | ■ | | | | | | | | | ▨ | ■ | ▨ |
| Documentation | | | | | | | | | | | ■ | | | ■ | | ■ | | | | ■ | | | ▨ | ■ | ▨ |
| Economics/costs | | | | | | | | | | | | | | | | | ■ | | | | | ■ | ▨ | ▨ | ■ |
| Effect on business/integration | | | | | ■ | | | | | | ■ | | | ■ | | | | | ■ | | | | ▨ | ▨ | ■ |
| Efficiency (representation) | | ■ | ■ | | | | | ■ | | | | | | ■ | | | ■ | | | ■ | ■ | ■ | ■ | ▨ | ▨ |
| Efficiency/minimality (constructs) | | | ■ | ■ | | | | | | | | ■ | ■ | | | | ■ | | | ■ | ■ | ■ | ■ | ▨ | ▨ |
| Executability | | | | | | | | | | | | | | | | ■ | ■ | | | | | | ▨ | ▨ | ■ |
| Expressiveness | | | | | | | | | | ■ | | ■ | | | | | ■ | | ■ | | | | ▨ | ■ | ▨ |
| Extensible/Customizable/Modifiable | | ■ | | | ■ | | | | | | ■ | | | ■ | | | ■ | ■ | ■ | ■ | ■ | ■ | ▨ | ▨ | ■ |