UNIVERSITEIT AMSTERDAM

VU Research Portal

The relevance of knowledge about existing information systems

Fischer, S.; de Vos, A.F.

1993

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Fischer, S., & de Vos, A. F. (1993). *The relevance of knowledge about existing information systems*. (Serie Research Memoranda; No. 1993-45). Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam.

Download date: 22. May. 2021

1993
45

Serie Research Memoranda

The Relevance of Knowledge About Existing Information Systems (Assessing the readiness for reverse engineering)

S. J. Fischer
A. J. van der Vos

vrije Universiteit   *amsterdam*

The relevance of knowledge about existing information systems [1]
Assessing the readiness for reverse engineering

S.J. Fischer
< v67csf01@vm1.sara.nl >
Faculty of Economic Sciences, Business Administration and Econometrics
Department of Information Technology and Information Management
Vrije Universiteit Amsterdam
The Netherlands

A.J. van der Vos
< ajvdvos@cs.vu.nl >
Faculty of Mathematics and Computer Sciences
Department of Information Systems
Vrije Universiteit Amsterdam
The Netherlands

Abstract
One of the critical issues in managing Information Technology (IT) nowadays is the management of existing information systems. Due to several causes, the decision as to whether knowledge about existing systems can and should be extracted has become a crucial one. This article discusses this problem area, and based on a problem definition outlines an approach to assessing the feasibility of knowledge extraction. The approach has two components: an overall, managerial part focusing on the readiness of the software organization for the new technology, and a specific, technical part focusing on assessing feasibility of knowledge extraction for individual systems. Both these components were defined and tested as part of a reverse engineering project at a large banking and insurance company in the Netherlands.

1. Introduction

One of the most widely used models for assessing an IT department of an organization is Stage Model of Richard Nolan (Gibson and Nolan, 1974, Nolan, 1973, 1979, 1987). Despite the fact that the model has been proved theoretically and empirically invalid (Lucas and Stutton, 1977, Kling and Kraemer, 1984), the model seems to be a valid tool for identifying critical issues for managing IT. Nolan discusses these issues along four dimensions: the level of planning and control, the portfolio of existing systems, the level user participation, and type of IT activities.

In this article the focus is one of the dimensions that Nolan describes, the dimension of the type of IT activities. Nolan makes a distinction between developing new information systems, making adaptions to existing systems, and integrating different information systems. As an organization reaches a higher level along the Nolan curve, attention shifts from new information systems, to existing information systems, and finally to integration. However, although the type of IT activity is a critical issue for managing IT, current literature does not offer any evidence with respect to this shift in attention.

Thus, based on current literature it is difficult to show the relevance of knowledge about existing information systems. The relevance of this knowledge increases as attention shifts to adapting existing systems, and also to integrating systems. The Nolan model does not provide us with any guidelines as to decide whether or not knowledge about information systems should be acquired or not. Although new research efforts have started to emerge in the areas of assessing IT management, e.g. the Maturity Framework of the Software Engineering Institute (see Humphrey, 1987, 1988, 1989), these efforts have not yet provided us with explicit guidelines.

So, when conducting research on the relevance of extracting knowledge about existing systems, new concepts have to be defined because existing literature does not provide for it. In this article a detailed discussion is presented with respect to the feasibility of knowledge extraction. The results are derived from a large research effort that focused on assessing the relevance of extracting knowledge about existing

1

systems. Because knowledge extracting is tightly linked to reverse engineering and reengineer-ing - reverse engineering is the most common method for knowledge extraction (see Reckoff, 1985, Chikofsky, 1990) - the research is also focused on these activities.

This article is organized as follows. First, the research problem is presented in detail. Several determinants for the relevance of extracting knowledge from existing systems are discussed. An overview is given of an approach to assessing the feasibility of knowledge extraction. The approach has two components: an overall, managerial part focusing on the readiness of the software organization for the new technology, and a specific, technical part focusing on assessing feasibility of knowledge extraction for individual systems. In the third section, the first component is discussed. Four different types of situations or paradigms are identified that allow a distinction to be made in IT departments where knowledge extraction is relevant, and IT departments where it is not. Based on this discussion, section four describes a framework for assessing the feasibility of knowledge extraction and reverse engineering for individual information systems. Based on these discussions, section five presents several recommendations and conclusions.

2. A description of the research problem
In May 1990, a banking and insurance company in the Netherlands started a reverse engineering project. The objective of the project was to clarify the potentials and weaknesses of the reverse engineering in the organization. The final results of the project were presented in October 1992. The project had three major areas of attention:

O method - the method/approach to reverse engineering.
O technology - availability of tech-nology to be used for reverse engineering.
O organization - potential use of reverse engineering in the IT organization.

To identify the potential use of reverse enginee-ring in the IT organization, two separate research efforts were conducted. The first effort focused on the overall, managerial issues, i.e. identifying the readiness of the organization for applying reverse or re-engineering. It is discussed in the next section. The other effort focused on the economical and technical issues of reverse engineering, i.e. to value information systems according to their reusability. It is discussed in section four.

The problem definition for these two efforts was defined in the following way. A specific business problem is considered: an increase commercial capabilities of the organization. With respect to information systems, commercial capabilities can be enhanced by making information systems more flexible, and by increasing the maintainability of information systems. Maintainability and flexibility can be achieved by renovation of complete systems, or by building new systems in order to replace existing systems. It is assumed that the first solution, total renovation, is too difficult for a large organization to realise. Thus, the focus should be on replacing existing systems.

With respect to replacement, two potential problem areas exist:
O (during information systems planning) it is difficult to assess the feasibility of replacement. It is difficult to relate the requirements of end-users to that which already exists.
O (during implementation of information systems) insecurity exists as to the replacement: the functionality of the existing system is not always known.

Based on these potential problem areas it can be concluded that the focus of replacement should be on *controlling the existing situation*. In order to control the existing situation, knowledge about function and structure of information systems is important. A way in which to acquire functional and structural knowledge is reverse engineering.

The problem definition was defined in such a way as to be characteristic for large organizations. It was derived from a large survey research project among companies using software development tools in the Netherlands (see Kuster et. al., 1992). The survey showed that the assumption of Nolan, i.e. the shift in attention from developing new systems, to existing systems and integration is not clearly visible. However, the percentage of the total number of man years spend by the IT organiza-tion on maintaining information systems, decre-ases with an increase in sophistication of opera-tional activities and planning activities. Hence

2

the definition of the two problem areas.

3. The overall readiness for knowledge extraction

Organizations are shaped based on a number of implicit and explicit assumptions. Usually, these assumptions are made by the founder of a company. Management of Information Technology (IT) is shaped in much the same way. It is based on several assumptions how to carry out information systems development, maintenance, information services, and how to plan for and control all these activities. Based on these assumptions, it is possible to identify several perspectives or metaphors that exist in IS and CS literature. Examples of perspectives are the management accounting philosophy, which stems from cybernetics and the metaphor of organizations as machines, organization development, political perspectives and social perspectives (see Morgan, 1986). Dealing with perspectives means dealing with issues that are not orthogonal. Consequently, several perspectives will usually be valid at the same time. We will not focus on metaphors but on paradigms. Because paradigms hold 'meta-theoretical assumptions about the nature of the subject of study' (Burrell and Morgan, 1979), no two paradigms can hold at the same time.

The concept of control can be used to make a distinction between several paradigms. With respect to control systems, Anthony identifies activities at a strategic level, tactical level and operational level (1964, 1972). Strategic activities address long term aspects, e.g. developing a strategy for IT. The tactical level represents activities at a shorter time frame, such as planning for various activities. The operational level holds most of the IT activities, i.e. short term activities as developing information systems, maintenance, project management, and services.

Figure 1 shows strategic, tactical and operational activities along two dimensions. The horizontal dimension represents the emphasis on operational activities performed. The vertical dimension represents the emphasis on the set of strategic and tactical activities. Strategic and tactical aspects were taken together to contrast management aspects to the more technical aspects of the operational activities.

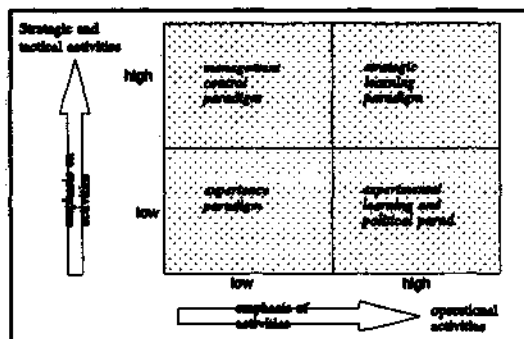Based on these two dimensions, four different paradigms are identified: experience control,

Figure 1. Four paradigms or approaches for managing IT

management accounting, experimental learning and strategic learning. The paradigm of experience represent few emphasis on strategic, tactical, and operational activities. Because the low level of emphasis on activities, practitioners have to rely on their expertise. Consequently, it will be very difficult to implement a proper control system; usually this system will be incomplete.

IT organizations that rely on the paradigm of management control are relatively sophisticated in strategic and tactical activities, but not in operational activities. Thus, experts pay a lot of attention to formulating a strategy for IT, but few attention to developing and maintaining information systems. Because of the focus on planning, the culture of the IT organization will be task and process oriented, tightly controlled and normative. The organization will rely on a cybernetic model of control, on a management control system.

The political and experimental learning paradigm is characteristic for organizations that emphasize the operational activities as compared to the strategic and tactical activities. Few emphasis is laid on the formulation of a strategy of IT, but developing and maintaining information systems is highly sophisticated. So, on the one hand state-of-the-art technology is used to arrive at a sophisticated level of systems development, but on the other hand the formulation of a strategy for IT is rather ad hoc. There are several explanations for this paradox. For example, IT is critical for survival of the organization, but the external environment is too dynamic to develop a strategy for use of IT.

3

The best the organization can do is always to use the state-of-the-art technology and rely on the ad hoc formulation of a strategy (see, for example, Informatie planning in de praktijk, 1992). Another explanation may be that the IT organization is very political, which reduces the relevance of plans and strategies. Regardless of the explanation for the paradox, the control process for the IT organization will be political: since no detailed plans and strategies exist, control will be result oriented.

The fourth paradigm, that of strategic learning, is characterized by a high level of emphasis on all activities, i.e. strategic, tactical and operational activities. Since we are dealing with non-routine, non-industrial type processes (see Cusumano, 1991), the control process will be a homeostatic one (see Hofstede, 1978), based on a learning system. Strategies are not necessarily planned, and can emerge from various places in the organization. The organization is self regulating, homeostatic.

When relating these different approaches or paradigms to controlling the existing situation, and to the relevance of extracting knowledge about existing systems, several conclusions can be made. Empirical research showed that as more emphasis is paid to operational and to strategic/tactical activities, the relative amount of time spent on maintaining information systems decreases (Fischer, 1992d). However, as discussed earlier, the shift in attention that the Nolan model assumes does not hold.

Based on the framework presented above a detailed study was conducted on the readiness for reverse engineering in the banking and insurance company. About thirty interviews were carried out at several key figures in the IT organization. The framework of figure 1 was fleshed out by using stage models to describe the horizontal and vertical dimensions in detail. It was concluded that the IT organization relied heavily on a management control paradigm, and as a result was ready for the new technology, but on a small scale. In addition, several critical factors for success were identified. For a more detailed discussion of the framework, see Fischer (1992a, 1992b, 1992d).

4. Assessing the feasibility of reverse engineering
Having assessed the potential for reverse engineering and knowledge extraction on a

conceptual level, the second effort of the project Revival was to develop and test a framework for assessing the feasibility of reverse engineering for individual systems.

The framework is based on the concept of software quality. When looking at characteristics of software quality, three main characteristics can be identified (Boehm et. al., 1978):

O *adaptability*, which focuses on the difficulty with which it is to adapt software. What is the total cost of adapting an information system to a new situation?

O *portability*, which focuses on the difficulty of changing the platform on which an information systems works. How many changes have to be made to run an information system on another platform?

O *usability*, which focuses on how well users are capable to work with the information systems.

Software quality can be regarded as the interaction between each of these three characteristics. Based on the characteristics, measuring software quality can be approached in several ways: in a top-down way, i.e. based on a total quality definition, in a bottom-up way, i.e. based on a definition of quality by the end-user, or in an intermediate way, i.e. based on a combination of user requirements and definition of quality.

When applying the intermediate approach to software quality, which we will call *relative quality*, the three main characteristics of quality can be related to external entities. In this way the abstract factors of Boehm are considered from a practical point of view. As figure 2 shows, adaptability is related to software, portability to hardware, and usability to the software organization.

The shaded areas in figure 2 represent the interaction between the external entities. They allow several lower level quality aspects to be identified. For example, reusability is the result of the interaction between software and the software organization. Thus, extraction of knowledge and the potential for reverse engineering can best be described in terms of software and software organizations, in other words, in terms of usability and adaptability.

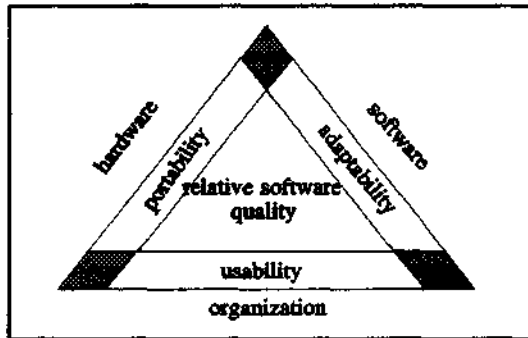What can be concluded from this discussion is

4

Figure 2. The relation between the three main aspects of relative software quality

that the feasibility of knowledge extraction, based on relative quality, is measured in terms of software and software organization. For both these aspects frameworks exist that identify several levels of maturity. For software, a classification can be used of several possible architectures of software (see Kusters, 1991, Vinig and Achterberg, 1991), ranging from unstructured programs to repository based programs. For software organization, the framework presented in the previous section may be used, or the Maturity Model of the Software Engineering Institute. In the Revival project, the Maturity Model was used to flesh out the framework.

A detailed approach to assessing the relative software quality of individual systems was developed in the Revival project. For a discussion, see Van der Vos and Fischer (1992). The model developed was tested in two practical cases. Based on interviews and existing code and documentation of the information system, a qualitative measurement of reusability was carried out. The measurements were consistent with the impressions of the users of the information systems, and provided them with a proper mechanism to compare different information systems. It allowed experts to decide among these systems whether knowledge extraction is feasible or not.

6. Summary and conclusions

In this article the aspect of knowledge extraction and reverse engineering was discussed. An approach was presented to assessing the feasibility of knowledge extraction. This approach consisted of two components. The first

component studies the overall readiness of an IT organization for the new technology, the second component on assessing the possibility of knowledge extraction or reverse engineering for individual information systems. For both these components detailed frameworks were defined, and tested as part of a large project at a large banking and insurance company in the Netherlands.

Although the components were qualitative in nature, they provided valuable insight into the problem area, and will therefore be used in other projects at the same organization. What can be concluded from this is that a qualitative evaluation is not necessarily undesirable, as long as it is uses knowledge of experts in maintenance and of end-users. Because feasibility study requires the integration of aspects of computer science, information systems and economics, it is almost impossible to define a totally objective approach to assessing feasibility. Such an objective approach should not be regarded as a panacea. Instead, it is better to apply multi-discipline research to arrive a qualitative model that is easy to understand and use for experts, and which allows them to value systems and technology based on their own experience.

References

- Anthony, R.N., (1964), *Management Control Systems*, Irwin, Homewood Ill.
- Anthony, R.N.,J. Dearden, R.F. Vancil, (1972) *Management Control Systems, Text, Cases and Readings*, rev. ed., Irwin, Homewood Ill.
- Basili, V.R., (1990) "Viewing maintenance as reuse-oriented software development", *IEEE Software*, No. 1, January, pp. 19-25.
- Boehm, B.W., et. al., (1978) *Characteristics of software quality*, North Holland.
- Burrell, G., G. Morgan, (1979) *Sociological Paradigms and Organizational Analysis*, Heinemann, London.
- Cavano, J.P., J.A. McCall, (1987) *A framework for the measurement of software quality*, November, Proceedings of the software quality and assurance group.
- Chikofsky, E.J., J.H. Cross, (1990) "Reverse engineering and design recovery: a taxonomy", *IEEE Software*, No. 1, January, pp. 13-17.
- Cusumano, M.A., (1991) *Japan's Soft-*

ware Factories - A Challange to U.S. Management, Oxford University Press.

- Fischer, S.J., (1992a) "A control perspective on information technology", Serie Research Memoranda, Research Memorandum No. 13, Vrije University Amsterdam.

- Fischer, S.J., (1992b) "A control perspective on information technology - towards business redesign of the IT organization", Information Systems and Organizational Processes, Third Australian Conference on Information Systems, ISOP92, Wollongong, Australia.

- Fischer, S.J., (1992c) "Boiling the frog or seducing the fox - organization aspects of implementing CASE-technology", to be presented at IFIP WG 8.2., Social aspects of information systems, 1993, Noordwijkerhout, The Netherlands.

- Fischer, S.J., (1992d), "Identifying different paradigms for managing Information Technology", Serie Research Memoranda, Research Memorandum, Vrije University Amsterdam.

- Gibson, C.F., R.L. Nolan, (1974) "Managing the Four Stages of EDP Growth", Harvard Business Review, pp. 76-88.

- Jolink, D. (ed.), (1991) Reverse Engineering (Reverse Engineering; in Dutch), Samson Bedrijfsinformatie, Alphen aan de Rijn/Deurne.

- Kusters, R.J., (1991) Software ontwikkel-omgevingen (Software development environments; in Dutch), Technical University Eindhoven, The Netherlands.

- Kusters, R.J., G.M. Wijers, et. al., (1992) Het gebruik van CASE-tools in Nederland (The use of CASE-tools in the Netherlands; in Dutch), Kluwer Bedrijfswetenschappen.

- Hofstede, G., (1978) "The poverty of management control philosophy", Academy of Management Review, July, pp. 450-461

- Humphrey, W.S., (1988) "Characterizing the Software Process: A Maturity Framework", IEEE Software, March, 73-79.

- Humphrey, W.S., (1989) Managing the Software Process, Addison-Wesley.

- Humphrey, W.S., W.L. Sweet, (1987), A Method for Assessing the Software Engi-

neering Capability of Contractors, Software Engineering Institute, Carnegie Mellon University, CMU/SEI-87-TR-23, ESD/TR-87-186, September.

- Informatieplanning in de praktijk - kunst of kunde?, (1992) symposium department of Information Systems, Vrije University Amsterdam, 21 mei 1992.

- Kling, J.L., K.L. Kraemer, (1984) "Evaluation and Organizational Information Systems: An Assessment of Nolan's Stage Model", Communications of the ACM, 27(5), May, pp. 466-475.

- Lucas, Jr., H.C., J.A. Sutton, (1977) "The Stage Hypothesis and the S-Curve: Some Contradictory Evidence", Communications of the ACM, pp. 254-259.

- Morgan, G., (1986) Images of Organizations, Sage Publications.

- Nolan, R.L., (1973) "Managing the Computer Resource: A Stage Hypothesis", Communications of the ACM, pp. 399-405.

- Nolan, R.L., (1979) "Managing the Crisis in Data Processing", Harvard Business Review, pp. 115-126.

- Nolan, R.L. (1987), "Managing the advanced stages of computer technology: key research issues", The Information Systems Research Challenge, Proceedings, (ed. F. Warren McFarlan), Harvard Business School Press, Boston, pp. 195-216.

- Reckoff, M.C., (1985) "On reverse engineering", IEEE Transactions on systems, man and cybernetics, March-April, pp. 244-252.

- Vinig, G.T., J.S. Achterberg, (1991), "CASE-technology - From software development life cycle (SDLC) to software engineering life cycle (SELC)", Serie Research Memoranda, Research Memorandum, Vrije University Amsterdam.

- Vos, A.J. van der, S.J. Fischer, (1992) "A qualitative approach to determining the feasibility of reverse engineering", Serie Research Memoranda, Research Memorandum, Vrije University Amsterdam.