

2005

Using process modelling to capture requirements for a digital library

Michael Johnstone
Edith Cowan University

Follow this and additional works at: <https://ro.ecu.edu.au/ecuworks>



Part of the [Computer Sciences Commons](#)

Johnstone, M. (2005). Using process modelling to capture requirements for a digital library. In Information, libraries and elearning : Proceedings of the Inaugural Transforming Information & Learning Conference. This Conference Proceeding is posted at Research Online. <https://ro.ecu.edu.au/ecuworks/2705>

Using Process Modelling to Capture Requirements for a Digital Library

Abstract

This paper describes the results of an action research study which explored how IT professionals used process modelling to capture requirements and design artefacts concerning a digital library system. The intention was to produce a web-based, non-relational (XML) database system to store publications of various types and to allow public domain access to the system. Traditionally, such systems are constructed with a data-oriented design, with little attention being paid to process concerns. In this study, process, described by business rules, was defined first which provides advantages as many business systems (including library systems) are often process-oriented.

The team involved in this study modelled a reasonably complex system, articulating 168 business rules across 18 major functions covering many aspects of the operation of a digital library system. The team was able to produce a usable requirements specification that was used as the major input into the design phase of the system being developed. The design phase considered both process and network perspectives explicitly before modelling a data-oriented view of the proposed system. It is argued that this multi-perspective view led to a better-designed system than would have been obtained with a single perspective method.

An issue for this study was that third year software engineering students were used in place of professional software developers. Using students instead of practitioners raises an interesting question concerning the authenticity of this study vis-à-vis action research. The fundamental question being tested in study was to ascertain whether a process modelling method could be used to model the domain of digital libraries i.e. does it contain a necessary set of constructs to satisfactorily model the problem domain? It was considered that the environment of student projects in a university with regular supervisory oversight offered advantages for the monitoring of this question more so than the alternative. As regards the authenticity, this study was a real project for a real client.

Introduction

Cathro (1999, para. 1) states that the term "digital library" first appeared in the early 1990s, as universities and other institutions began to build discipline-based collections of information resources in digital form, and to provide access to these collections through local and wide area networks.

Since then, examples of digital libraries with varying degrees of sophistication have flourished, except in the field of information systems (with the notable exception of the ACM collection, which

is a user-pays system). Venable (1999) highlighted the problems facing the information systems community with respect to the sourcing and distribution of resources (see figure 1) and proposed a public domain information systems digital library (ISDL) as a solution.

This paper presents the results of an attempt to prototype a public domain digital library system suitable for the information systems discipline.

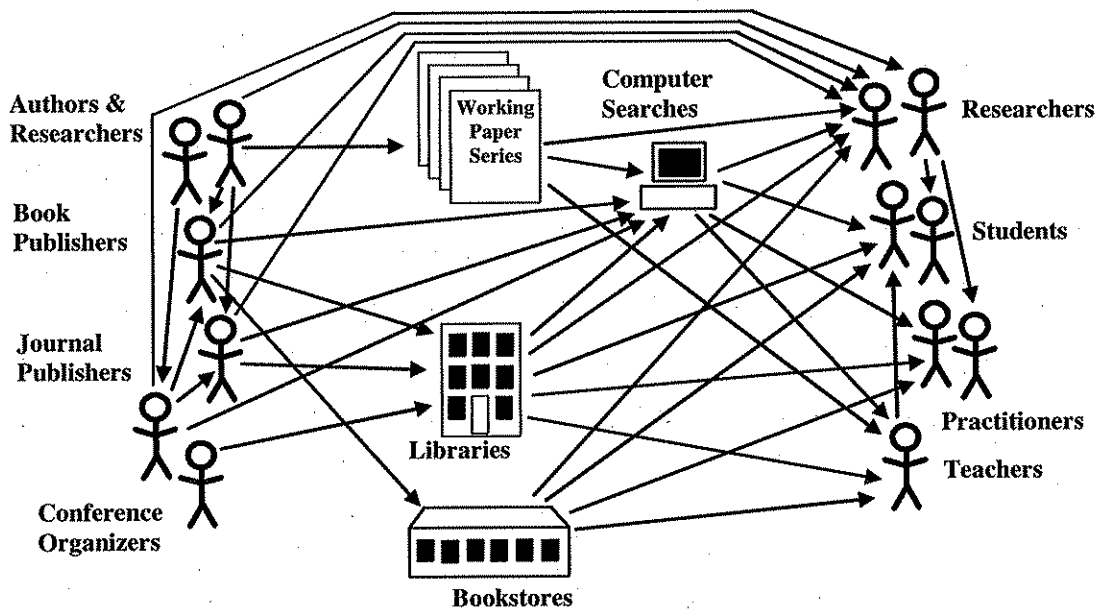


Figure 1: Current Locations of Resources and Major Actors in Information Systems Research (taken from Venable, 1999).

Requirements Modelling with Business Rules

The "software crisis" of the 1960s resulted from the tension between the introduction of powerful third generation

computer hardware and the failure of software development to keep pace with what is now known as Moore's Law. The result was that the software development methods that existed at the

time were not able to successfully produce large software systems of sufficient quality. Systems were delivered late, cost much more than expected, performed poorly and were difficult to maintain.

One response to this crisis was a focus on control of process (which spawned many systems development approaches e.g. Prototyping, the Spiral Model and latterly, approaches based on reusability such as the Rational Unified Process). Concurrently, there was also a drive to develop better methods for creating and maintaining software, to the extent that Jayaratna (1994) now believes that there are over 1000 extant methods for building software.

This does not mean, however, that research into development methods should cease, in fact, quite the contrary. There has been a shift away from established prescriptive methods such as SSADM to more agile methods, due, in part, to a reaction against mechanistic methods of development that may not actually assist in solving real business problems.

The re-emergence of business process modelling as a tool for understanding the functions within organisations, coupled with the necessity to maintain state information in web-based transactions suggests that process-oriented techniques may yet be useful in modelling business systems. Melão and Pidd (2000), for example, call for a pluralistic approach to modelling and suggest several perspectives on business processes (viz. deterministic machines, dynamic systems, feedback loops or social constructs). Earlier work has been completed in modelling web-based systems using state-based

models. The work described here continues that research by exploring how practical, acceptable and efficient this method is for inexperienced clients (in this case, undergraduate students).

There are many definitions of what exactly constitutes a business rule. It is important that the definition of a business rule captures something of the context and nature of a business rule as well as identifying its constructs. McDermid (1998, p20) defines a business rule as *'...an explicit state change context in an organisation which describes the states, conditions and signals associated with events that either change the state of a human activity system so that subsequently it will respond differently to external stimuli or reinforce the constraints which govern a human activity system'*, a definition that will be used from here onwards.

One model used to capture business rules is that of McDermid (1998). This model, called the Business Rules Diagram (BRD) method, is a state-based model which uses notation similar to, but more powerful than, flowcharts. As an information systems analysis method, the BRD method is positioned between the use case approach of Jacobson et al. (1992) and more complex object models.

A business rules diagram (see figure 2 for an example) uses combinations of four atomic constructs; states, events, conditions and signals. States (represented by circles) reflect the status of an object of interest at any given time, so for example a manufacturing work order might occupy the states not started, in progress or completed. Events (rectangles) are actions carried out internally by an organisation. They

are considered to be instantaneous occurrences that model an organisation's business processes e.g. cancel work order.

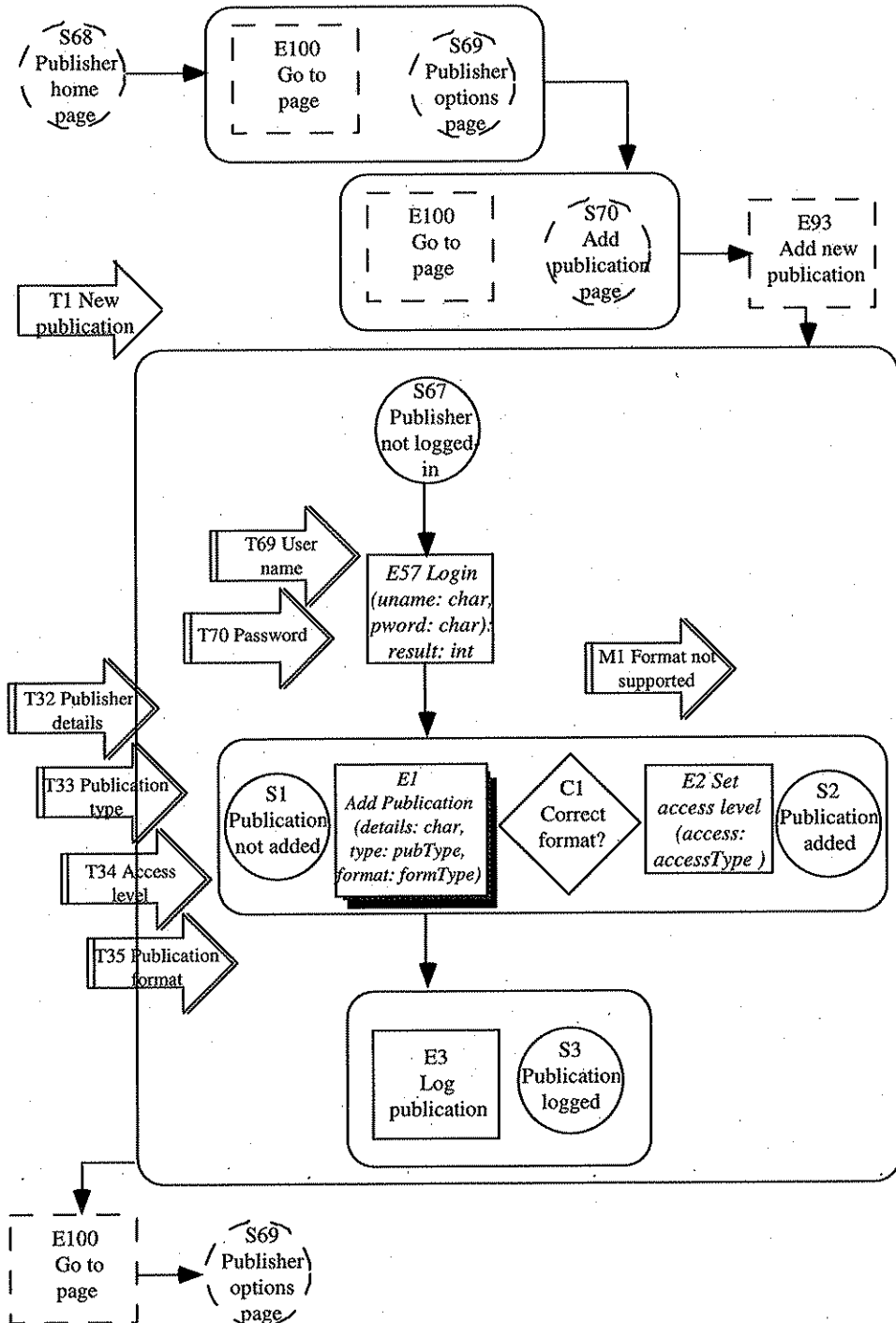


Figure 2: A Business Rules Diagram Representing the "Add Publication" Use Case from the Study.

Conditions (diamonds) define the criteria by which objects of interest in the business move from one state to the next as events take place. Sometimes, several conditions must be met in order for an event to take place, thus increasing complexity. It is argued that modelling conditions without the context of states and events (and vice versa) is far less powerful or useful. Lastly, signals (thick arrows) either enter or leave the human activity system. Signals that enter the system will typically initiate activity within the system and so these are called triggers. Triggers may be external, such as a customer sending an order, or internal, such as one department sending a document to another department, both of which then trigger off some activity. Further, a trigger may be a time trigger, e.g. an activity beginning at the start of the day or the end of the month. Those signals that leave the system serve the purpose of informing those outside the system of what has occurred inside the system and therefore are referred to as messages. An important additional construct is the Harel blob (Harel, 1988), represented by a rounded rectangle, which encapsulates other constructs and is used to model selection or simultaneous action.

The Action Research Study

In this study the client was a university academic with many years of systems development experience. The participants in the study were a systems analyst with 17 years experience in information systems development (the researcher), two method users (university students) and the client. The method users were two third-year undergraduate software engineering

students with little work experience in IS development. The problem domain was that of digital libraries. The intention of the method users was to produce a web-based, non-relational (XML) database system to store publications of various types and to allow public domain access to the system.

This study was reasonably complex, with the team articulating 168 business rules across 18 major functions covering many aspects of the operation of a digital library system. These functions included adding a publication, searching (both full-text and metadata), building a collection, indexing a collection and the ability to annotate collection items. In this system, a publication is considered to be any information systems resource stored on the ISDL, retrievable by anyone while a collection is a user-defined series of (usually similar) publications. Table 1 show some sample business rules from the system. Notice that the rules can be generic (e.g. rules 1.1 and 1.3), specific (e.g. rule 1.4) or concerned with navigating the system (rule 1.5). Rule 1.2 indicates that there are particular actors (users) that access the system and that those actors may have different requirements of the system or differing constraints placed upon their use of the system.

Initially, a use case diagram was generated by the team which assisted in classifying the business rules and served as a top-level context diagram throughout the remainder of the study. The method users were also able to produce a usable requirements specification that was used as the major input into the design phase of the system being developed.

ID	Business Rule
1.1	Any individual can use the system
1.2	A book publisher can add a book or an edited book
1.3	Publications can exist in different formats
1.4	Formats of publications are limited to: text, word, pdf, postscript, rtf, latex, html, xml
1.5	After authenticating, a publisher is taken to the publisher options page

Table 1: Sample Business Rules from the "Add Publication" Use Case.

The researcher acted as a facilitator and thus provided guidance only when asked by members of the team, therefore reducing (but not entirely removing) the effect of an imposed world-view from the researcher about what constitutes "good practice" or "the solution to the problem". In order to assist in tracking the progress of the studies, the team kept minutes of project meetings, dated the production of BRD artefacts and the researcher kept separate field notes containing reflections about the products and social processes that were the outputs of the studies. The method users were interviewed separately by the researcher at two key intervention points, the completion of the analysis phase and the completion of the design phase of the system. These interviews were semi-structured and taped for later analysis by the researcher.

From a purely technical perspective, the method users chose to use an open source native XML database, thus

electing to favour flexibility and extensibility over speed of access. The major advantage of such a system is that it would preserve the XML hierarchy but such technologies are still immature (compared to relational systems) and thus scalability concerns would be an issue.

In order to ensure a suitable degree of rigour (the lack of which is often a criticism of action research), the guidelines for conducting interpretive research proposed by Klein and Myers (1999) were followed in this study. The study used the research framework of Venable and Travis (1999), which applied the framework of Nunamaker et al. (1991) to the creation and refinement of systems development methods as a kind of system to be developed during IS research and extended Nunamaker's framework to include action research as a method of in situ investigation (see figure 3).

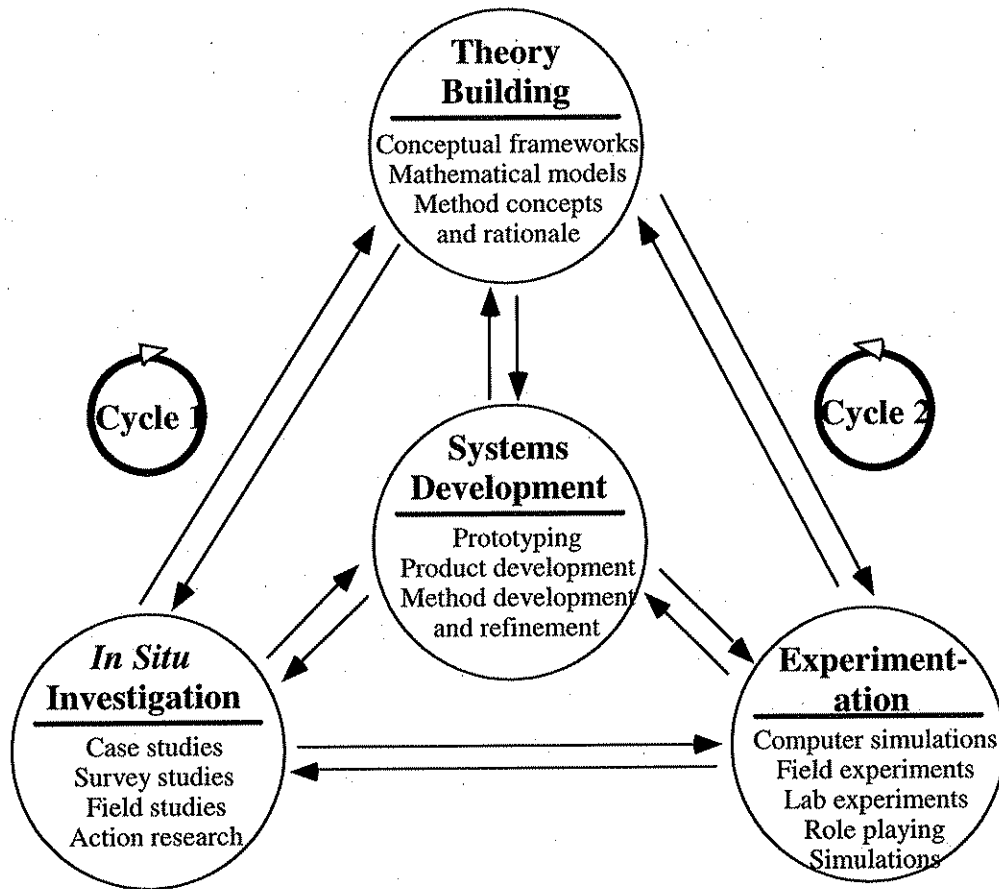


Figure 3: Systems Development in IS Research (adapted from Venable and Travis, 1999 after Nunamaker et al., 1991).

The process of the entire action research study is depicted in figure 4. The ISD activities studied are shown with ovals. The researcher activities are depicted as rectangles. Rounded rectangles are used to show the evidence collected.

Before the advent of the study, the researcher/analyst and the two users entered the situation with their individual world views. These world views included their conceptualisations and expectations about how information systems development (ISD) is

conducted, what they would be doing, and how they would interact with each other. The large circle at the top of figure 4 shows the ISD activities themselves, divided into smaller ovals, which are the individual sessions. At the conclusion of the modelling phase, as shown by the lower oval, the method users (students) were interviewed by the researcher in semi-structured mode in to gather their perceptions of the modelling experience and how the BRD method assisted (or otherwise) in reasoning about the problem situation.

Evidence was collected at each stage during the study, as shown by the rounded rectangles in figure 4. Transcripts were made of each ISD session. Copies were kept of all the training materials and all of the BRDs

generated during the ISD sessions. Following each ISD session, the analyst/researcher recorded a summary of his perceptions of the sessions in his research journal.

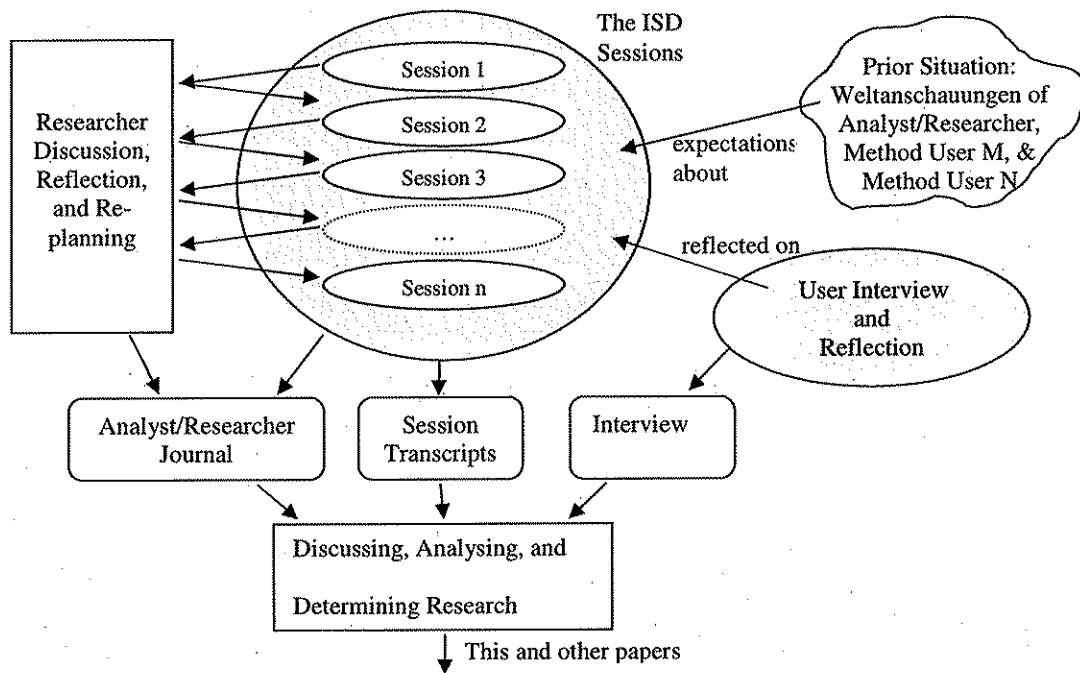


Figure 4: The Research Process.

Analysis of Evidence Method Usage

The use of a process-oriented method is interesting as the domain (digital libraries) traditionally lends itself to decomposition by data-oriented methods (such as entity-relationship models). The challenge for the BRD method in this context is to demonstrate the utility of the process viewpoint.

Figure 2 shows one of the diagrams generated from the study. It uses many of the constructs available in the BRD

design notation. In particular, the diagram demonstrates the "wrapping" of navigation logic (the constructs with dashed lines) around process logic. This separation of navigation logic from process logic is important both in terms of the comprehensibility of the diagram and also generating a clean design. In this diagram, the events are more decorated than their analysis counterparts in that they (the design events) have named parameters that match the input triggers (e.g. T69, T70 and E57). At the design level, these

parameters may be simple types (e.g. char, int, boolean, date) or user-defined types (e.g. formType ::= [text | word | pdf | postscript | rtf | latex | html | xml]), that represent the types of documents that can be stored in this implementation of a digital library.

In considering the team as interpreters of data, the theory under examination is the BRD method itself. During the interviews, the participants were queried about their use of the method in several ways in order to attempt to establish the level of learning that had taken place. Initially, both participants demonstrated shallow learning by showing that they could describe a process or complete a set of tasks. As the interview progressed, there was some evidence of deep learning when the participants were asked what problems they faced and how they chose to solve them. Deep learning was suggested by the way the participants described the alternatives available to them when faced with modelling problems. They had internalised the method and were not just repeating a process by rote as in the initial parts of the interviews. This is illustrated by the following exchange taken from an interview transcript:

[Researcher] What about detecting inaccuracies, problems that came out?

[User N] All right, the EST (note: the EST is a tabular representation of the lawful state space of the business rules diagram) helped in detecting inaccuracies...In the process of drawing the diagrams [unintelligible] group member - I don't really remember it very well, or maybe if the EST was wrong or something, like let's say we had a problem here, maybe typing errors or

something. I don't really remember one at the moment.

[Researcher] That's all right. How useful was each representation (i.e. text - business rules, graphic - BRD, and table - EST)?

[User N] I like the event specification table - it helped me figure all out my errors that are taking place at a particular place and helped me fix them. By using the EST I would be able to fix the business rules or the business rules diagram. Ok, both...both...everything helped me but I like the EST.

[Researcher] Maybe you could tell me how the EST helped you to fix the business rules?

[User N] Like the other example...ahhh...it just threw out, like...we didn't have a particular business rule when you are going through the EST, it let us know that a particular business rule is not there, or is not there in the business rules dia...or is not included in the business rule diagram.

[Researcher] Maybe we could look at how the EST helped you to change or fix the diagrams?

[User N] Like when you look at the EST, at a particular row and column, you're looking at a particular condition taking place or not taking place and when you read it you see there is no business rule for that. In the process of working out we realise there's no business rule so it helps us edit business rules, so we can consult the client [and inform him that] "this particular business rule is not there".

[Researcher] Right, and how would that affect the diagrams?

[User N] So, once you've got a new business rule, maybe you have to add an event or a state into the diagram - that will help us as the event or state.

Initially User N can't think of an example of requirements validation but, when the question is framed differently, is drawn to describing how he used an EST to validate both business rules and the corresponding diagram. At first, he demonstrates surface learning by describing the process of constructing diagrams by applying the state-event-state meta-rule. Later, he is able to discuss the effect of validation using multiple representations. Deeper learning is implied here as he uses a (correct) concrete example to elaborate his generic learning.

As the method users were able to analyse, design and construct a prototype digital library, the evidence suggests that process modelling is a useful tool for reasoning about requirements in what is normally a data-oriented domain. A potential criticism is that the success of the method for this type of implementation was not due to the method but was more closely related to the technology chosen for the implementation. The XML database used a tree structure to hold data whilst a traditional relational database uses a tabular structure. Tree traversal can be seen as a dynamic process rather than a static data representation thus it could be argued that part of the success of the study was due to the nature of the database used to implement the digital library.

Students as Substitutes for Professional Developers

A concern for this study was the use of undergraduate students as the method users and thus it is appropriate at this point to discuss the issue of using students as substitutes for professional developers.

Using third year students instead of practitioners raises an interesting question concerning the authenticity of this study with regard to action research. The fundamental question being tested in this study was to ascertain whether the extended BRD method could be used to model the domain of a digital library i.e. does it contain a necessary set of constructs to satisfactorily model the problem domain?, are the steps in the method sufficiently clear for the analysis and design to proceed in an orderly fashion? It was considered that the environment of student projects in a university with regular supervisory oversight offered advantages for the monitoring of these questions more so than the alternative. As regards the authenticity, this study was a real project for a real client.

Greenberg et al. (1987) argue that studies that employ students are useful in that they provide opportunities for comparison between student and non-student groups. More to the point, Remus (1986) studied the use of students versus professionals in managerial decision-making and found that there was no significant difference between groups of students and managers. Whilst managerial decision-making and information systems development are not necessarily identical, they do share some commonalities in terms of both being attempts to solve complex problems

using a range of formal, semi-formal and informal tools and techniques. More recently, Briggs et al. (1996) found (MBA) students to be good substitutes for professionals in studies relating to the use of technology. This suggests that the idea of using students as proxies for professional IS developers has some validity, depending on the context of the research and the problem complexity.

Conclusions

The study provided evidence that it is possible to successfully analyse, design and develop a prototype digital library system using a process-oriented method. Additionally, recalling that the system has a web-based interface, a state-based method was used to model a stateless system (the Web). A useful

extension to the BRD method would be a data-oriented perspective which may increase its utility in other domains. The results from the study suggest, however, that the existing method is powerful enough to represent complex problems, yet simple enough to teach to undergraduates.

Whilst the study considered various essential elements of a digital library system such as metadata and the storage/retrieval of disparate resources, some limitations of the study were firstly, that it is difficult (and perhaps unwise) to generalise from a single action research study and secondly, the study was not concerned with information policy or organisational issues (which certainly affect the implementation of digital library systems).

References

- Briggs, R.O., Balthazard, P.A. and Dennis, A.R. (1996). "Graduate Business Students as Surrogates for Executives in the Evaluation of Technology". *Journal of End User Computing*, 8 (4), 11-17.
- Cathro, W. (1999) 'Digital libraries: a National Library perspective' in *Strategies for the next millennium: Proceedings of the Ninth Australasian Information Online & On Disc Conference and Exhibition*, ALIA, Sydney. Retrieved on 26 September 2005 from <http://www.nla.gov.au/nla/staffpaper/cathro4.html>
- Greenberg, J., Gordon, M.E., Slade, L.A. and Schmitt, N. (1987). 'The College Sophomore as Guinea Pig: Setting the Record Straight/Student Guinea Pigs: Porcine Predictors and Particularistic Phenomena'. *Academy of Management Review*, 12(1), pp. 157-163.
- Harel, D. (1988). 'On Visual Formalisms'. *Communications of the ACM*, 31 (5), 514-30.
- Jacobson, I., Christerson, M., Jonsson, P., and Övergaard. G. (1992). *Object-Oriented Software Engineering*. Reading, MA: Addison-Wesley.
- Jayaratna, N. (1994). *Understanding and Evaluating Methodologies*. London: McGraw-Hill.
- Klein, H. K. and Myers, M. D. (1999) 'A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems', *MIS Quarterly* 23c(1), 67-93.
- McDermid, D.C. (1998). 'The Development of the Business Rules Diagram', PhD thesis, Curtin University of Technology.
- Melão, N. and Pidd, M. (2000). 'A conceptual framework for understanding business processes and business process modelling', *Information Systems Journal*, 10, 105-129.
- Nunamaker, J.F. Jr., Chen, M. and Purdin, T.D.M. (1991). 'Systems Development in Information Systems Research'. *Journal of Management Information Systems*, 7 (3), 89-106.
- Remus, W. (1986). 'Graduate Students as Surrogates for Managers in Experiments on Business Decision Making'. *Journal of Business Research*, 14, 19-25.
- Venable, J. R. (1999). 'A Proposal and Prototype for an Information Systems Digital Library'. *Pages 1118-1129 of: Hope, B. and Yoong, P. (eds), Proceedings of the 10th Australasian Conference on Information Systems*, Wellington, New Zealand.
- Venable, J.R., and Travis, J. (1999). 'Using a Group Support System for the Distributed Application of Soft Systems Methodology'. *Pages 1105-1117 of: Hope, B. and Yoong, P. (eds), Proceedings of the 10th Australasian Conference on Information Systems*, Wellington, New Zealand.