Modelling Tacit Knowledge via Questionnaire Data

Peter Busch and Debbie Richards

{busch,richards}@ics.mq.edu.au
Department of Computing
Macquarie University, Sydney, Australia

Abstract. The transfer of tacit knowledge is important in ensuring that an organisations most valuable assets do not walk out the door. While much controversy surrounds the definition of tacit knowledge and whether it can be captured, in this paper we follow a psychological approach based on the work of Sternberg at Yale that seeks to measure tacit knowledge via the capture of responses to work-place scenarios. Focusing on the information technology (IT) work-place, we have developed a tacit knowledge inventory which forms part of a questionnaire given to experts and non-experts in three separate IT organisations. In psychology, descriptive statistics are typically used to analyse the responses. We have chosen a more qualitative and visual approach and have used formal concept analysis (FCA) for data analysis that better suits our small sample size. Using FCA we were able to identify participants that responded similarly to the peer-identified experts. In this way the organisation is alerted to the important role these individuals potentially play.

1 Introduction

Tacit knowledge is that vast store of experiences, technical know - how, skill sets and wisdom that permits us to function from a basic survival level, to interacting in our complex knowledge rich societies, that over time largely becomes codified. It is "knowledge that resides in the minds of the people in an organisation but has not been put in structured, document-based form" (Davenport, De Long and Beers 1998). Tacit knowledge has also been likened to intuition, but in reality intuition is more likely 'used to access tacit knowledge' (Brockman and Anthony 1998). We believe much tacit knowledge is articulable, whereas intuition is less so, lying more within the subconscious spectrum. There is naturally much more to it than this.

Although Fleck (1997) ascribes tacit knowledge at the whole organisational sense as being a form of meta or cultural knowledge, the purpose of our work is to examine tacit knowledge in individuals and more particularly the diffusion of such knowledge through the organisation. This paper reports research that extends the techniques of Sternberg (1995) by developing and administering a tacit knowledge inventory to IT practitioners; using Formal Concept Analysis

(FCA) (Ganter and Wille 1999, Wille 1982, 1992) to model the data captured; and Social Network Analysis (SNA) (Scott 1991) to map the tacit knowledge flows between participants. In short, we elicit both the ethical and realistic responses (as a pilot study revealed that a participants' actual response often differed from what they believed to be most appropriate) from domain experts and non-experts to a number of typical IT scenarios. The responses, together with biographical data, are captured via a questionnaire. The results are visualized as a lattice permitting an alternative to the descriptive statistics normally produced as a result of questionnaire processing. As a final step, the communication patterns and relationships between respondents are modeled using SNA to determine if any bottlenecks exist. For space, we do not discuss SNA or its use in this paper.

In the next section we introduce the application domain. In section 3 we describe our use of FCA. In section 4 we show how FCA has allowed us to identify individuals behaving similar to the peer-nominated experts. Our conclusions and future directions are given in Section 5.

2 The Application Domain - IT Tacit Knowledge

A study has been conducted in three IT organisations of varying sizes and structure and covering the private and public sectors, which we refer to as organisations X, Y and Z[1]. The questionnaire included three parts: biographical data, SNA-related questions requesting the frequency and nature of interactions with others in the organisation and a final part with scenarios and various answer options to consider. We have used FCA to model the biographical data and scenario responses.

The example in Figure 1 shows one of the scenarios randomly assigned to participants. Only answer option 12 for this scenario is shown in the Figure. Unlike the biographical data, the scenario section deals with one dimensional ordinal data. For example, we ask only that respondents select a value from extremely bad through to extremely good. We see an example of a Likert scale in Figure 2. Scenarios are randomly but equally distributed to participants so that a range of scenarios are covered with each individual only needing to respond to six scenarios. Tacit knowledge is measured and identified by determining the responses of those identified as experts by their peers in the survey. The biographical data allows profiles of experts vs non-experts to be developed.

3 Modeling Questionnaire data with Formal Concept Analysis

Modeling the scenario responses using FCA permits patterns in the results to be visualized. Such an approach is considered useful particularly if questionnaire

[1] The number of respondents involved in the study were 108, 7 and 16 in organisation X, Y and Z, respectively. The data collected is confidential and thus unavailable for distribution

Scenario 3

You as a team leader are responsible for implementing a payroll system for another branch within the parent organisation. Although you are expected to do the bulk of the work (55%), you do have five other colleagues able to help as you so desire. The project should take 12 months in total to complete.

You have undertaken some of the initial systems design work largely yourself for the past couple of months, and you now require your colleagues to further help you with the next stage which is mainly that of coding.

You are comfortable with hierarchy, however some of your team members are not. You delegate some tasks to subordinates within your team. One of the team members who specialises in programming has been allocated some software specification work, but would prefer really just to be programming. This person has performed well on coding related tasks in the past, but at this point in time lacks project management skills which would prevent him from becoming an effective team leader. Nevertheless you feel that the person should at least do some of the software specification work.

Rate each of the following responses in relation to the given scenario. It is advisable to read all of the responses before replying.

Answer 12 Give him fewer but more specific tasks to do, because it is simply not worth the effort to argue with him. Besides his skills in coding mean that he will be able to effectively contribute here, but then you can be rid of him, to concentrate on testing with other team members of your choice

Fig. 1. Illustrating Sample Scenario 3 with Answer Option 12.

Choose one:	○	○	○	○	○	○	○
	Extremely Bad			Neither Good nor Bad			Extremely Good

Fig. 2. Likert scale (of ordinal type) used in the tacit knowledge testing process

sample sizes are small, as the quantitative data may not be substantial enough to permit valid statistical conclusions to be drawn.

Assuming the reader possesses at least basic familiarity with FCA, we do not introduce it here. Following Wille (1982) we use the terms: (G)egenstande for the objects, (M)erkmale for the attributes, and the relationship between them: (I)deale. We refer to the combination of (G,M,I) as a formal context. A binary K(ontext) (in German) may be formally expressed thus:

K:= (G, M, I)

A multivalued context may be expressed a quadtuple:

K := (G, M, W, I) and I $\subseteq G \times M \times$ W

where the relationship I is a subset of the combined components of Objects (G), Attributes (M) and merksmaleWerte (W) (Attribute-values). In short, "the aim of FCA is to find (formal) concepts and construct the hierarchy between them. To achieve this aim, the many valued context must be transformed into a single valued one by so - called conceptual scaling" (Ganter and Wille 1989 in Bartel and Brggemann 1998 :24). Along the lines of Kollewe (1989), the FCA approach is able to process questionnaire data. We translate the multi-valued questionnaire responses into a formal context, a binary cross table which may be constructed, where: K = Formal table with its corresponding data; G = The participant; M = The responses; W = The value of the response; I = Relationship between the responses, their values and the participants

Kollewe's (1989) use of questionnaire data used the units of questioning as the objects (G), the questions as the attributes (M) and the answers to the questions as the attribute-values (W). In a pilot study (Richards and Busch 2000), it was decided to represent the data differently as it seemed more intuitive. We regarded the participant as the object (G) that has a number of features (M) such as age and position in addition to a set of responses and their values. This approach made data entry and validation easier as there was a one-to-one correspondence between the survey returned and the participant. We begin to implement a crosstable using the Anaconda(software, where the rows are embedded Structured Query Language statements that in turn access the data from a relational database.

We can convert the Likert scale in Figure 2 into a crosstable to provide a visual representation of the scale that will be used in the display of the data. The conversion of likert-scale data to a crosstable has been previously performed by Spangenberg and Wolf (1988). That study used the repertory grid approach to elicit the responses of anorexia nervosa patients to various people in their lives based on a number of bipolar scales. For example, one scale used in the study was a six-point likert scale with open-minded at one end and reserved at the other end. For that scale, if a score of 1 or 2 was given for an individual, the open-minded attribute would be marked with a cross. If a score of 5 or 6 was given, the reserved attribute would be checked. A score of 3 or 4 received no crosses and could be interpreted from the lattice by that individual being neither open-minded or reserved. Our work differs in that we are not concerned with repertory grid data and, as shown in Table 1, we do not handle multi-valued responses in such a simplistic way.

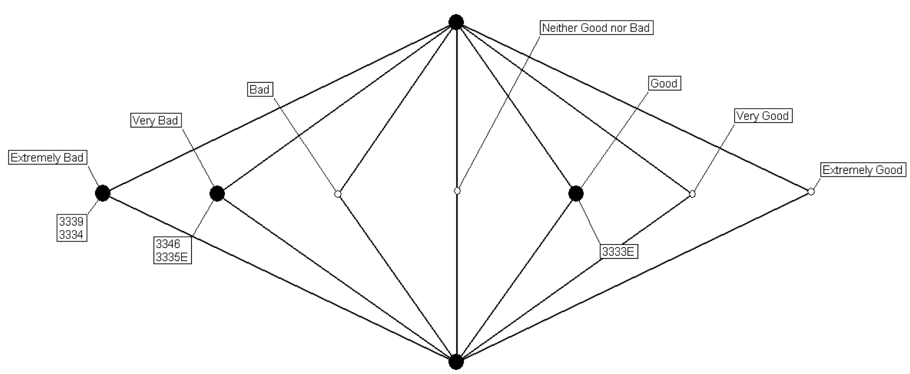

Fig. 3. illustrating the concept lattice with data included from the database table. Note how the diagram shows results for the values Extremely Bad (2 respondents), Very Bad (2 respondents) and Good (1 respondent).

Using the complementary Toscana(software, which takes the Anaconda(data, one is able to eventually construct the complete concept lattice as illustrated in Figure 3. In this instance we are able to visualise the keys (in the relational database sense) of individuals who have chosen particular values for an actual tacit knowledge answer option.

In the development of suitable scales for each of our data elements, we considered possible uses of the subsumption relation. Taking the biographical question, 'please select the highest formal qualification (or equivalent) you have obtained' as an example (Figure 4), we can create a scale which captures the subsumption between each of the qualifications. The scale developed will depend on subjective judgement and in this particular example vary across cultures and societies. For

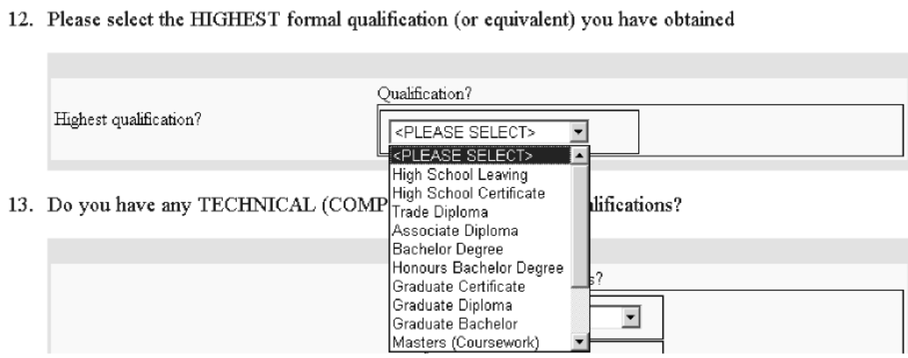

Fig. 4. illustrating a biographical question, in this instance highest formal qualification obtained

4 The Identification of Expert Non-Experts

Having collected the survey data and developed the scales to be used in Toscana we began to explore a number of questions we hoped the data would answer. Many of the questions revolved around the responses of experts vs non-experts. Following Sternberg, in our approach, experts were selected by their peers from the list of participants in the online survey. The question involving identification of experts was not specific to a particular scenario. Those not identified as experts became the control or normal group. 32, 5 and 5 individuals from organisations X, Y, Z were nominated as experts by their colleagues. One question of particular interest was whether we could identify participants that were 'hiding their light under a bushel' which would reduce the likelihood of tacit knowledge being transferred by them to other workers. To answer this question we wanted to determine if there were individuals who behaved like experts but had not been designated as an expert. This new group of people would be the expert non-experts. The ramification for identifying expert non-experts is that this sample group are likely to constitute employees who also have high levels of tacit aka 'managerial' knowledge. This is important to remember when we examine the interactions of staff when looking at the results of the Social Network Analysis.

The creation of a formal concept lattice for each scenario answer, showing the ethical and realistic response of both experts and non-experts, meant that a profile could be created of personnel who had answered close to that of experts. Let us examine this process in closer detail using the sample scenario and answer option given in Figure 3. For answer 12 to scenario 3, initial descriptive statistics revealed that the 17 experts were inclined to be a little more negative or pessimistic (ethically) (mean of 2.5) when dealing with this answer, than the 23 non-experts (mean of 3.2) on the whole. Likewise the experts were also marginally more negative realistically (mean of 3.6), than the non-experts (mean of 3.8). Examining the median values however we note that experts were even more negative ethically (2.0, or Very Bad) than the non-experts (3.0 or Bad). Whilst realistically both groups were actually fairly non-committal (4.0 or Neither Good nor Bad).

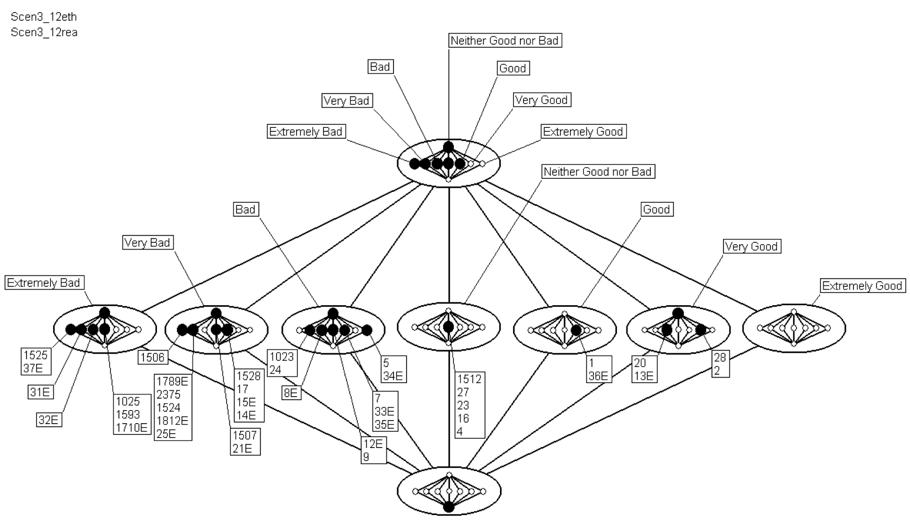

Fig. 5. illustrating formal concept lattice for Scenario 3, answer 12

While descriptive statistics allowed us to measure the sensitivity between the expert population and the non-expert population, to determine if there were any statistical differences between the responses of the two populations we performed a Wilcoxon test. Due to the sample size we could not conduct a test for each scenario and answer option but needed to use all the data. We found no statistically significant difference between the expert and normal groups. Given that via FCA analysis we identified 29 individuals in the non expert group as behaving like experts, this result is not surprising. In fact, we found that once we added the expert non-expert group to the expert group the Wilcoxon test did produce a statistically significant difference for the ethical responses. At this stage we reserve judgement on the value of statistical methods for analyzing this data until we have performed some more sophisticated techniques such as analysis of variance or cluster analysis. We show next how via FCA we were able to perform more fine-grained analysis at the individual level to identify meaningful differences between the expert and non-experts.

Lattices such as that in Figure 5 were used to identify expert non-experts. The reader is reminded that the lattice is read starting from the outer ellipses (ethical values) through to the inner lattice structures (realistic values). Thus in Figure 5 we can already visualise that answers are inclined somewhat toward the negative end of the spectrum for this answer. We can also see that some personnel (1528, 17, 15E, 14E, 7, 33E, 35E, 5, 34E) whilst ethically feeling negative about this particular answer for dealing with scenario 3, nevertheless feel positive about the answer from a realistic standpoint. Alternatively we may note that individuals 20 and 13E feel positive ethically about this answer, but negative about it realistically. Generally speaking however the ethical feeling tends to be more negative than the realistic one.

The point of this exercise is that in examining each formal concept lattice for each answer option for each scenario, we are slowly able to build up a picture for how non-experts have answered relative to that of experts. For example in Figure 5 above we can see that 1525, 1025, 1583, 2375, 1524, 1528, 17, 1507, 9, 7, 5, 1 and 20 have answered the same way as experts. By noting this similarity for all 125 answers (examining all concept lattices) non-experts can be listed in descending order of similarity with expert answer responses. For Organisation X, an extra 25 personnel were able to be identified who consistently scored close to that of experts using the aforementioned technique, without necessarily being identified by their peers as being experts. Examining the closeness of the scores in descending order, it was decided the top 32group. These were individuals 1, 2, 4, 5, 7, 9, 16, 17, 20, 23, 24, 27, 28, 1023, 1507, 1521, 1524, 1527, 1534, 1536, 1538, 1543, 1778, 1792, 1796.

When dealing with organisation Z, there were 4 other individuals who obtained scores close to that recorded by experts. Once again we note that at least two of the individuals are of non-English speaking origin. Participant 3334 decided not to provide any biographical details, 3339 declined to provide gender. Note once again bachelors degrees and high school certificate as being the highest formal qualifications. As there were only two individuals (out of 7 participants) who were not actually identified as experts in Organisation Y, the same technique was not adopted.

5 Conclusion

Whereas many researchers in the knowledge management field attempt to focus on the tacit component, few means actually exist to measure this type of knowledge. Sternberg's Yale University based approach could be said to be the most practical because of its applied nature. Thus in our approach we have developed a Sternberg-like workplace-based inventory for the IT-field where employees are asked to make decisions as to how they would handle a soft knowledge situation from an ethical and realistic perspective. Part of the Sternberg approach to tacit knowledge measurement is the identification of experts within the organisation by the participants. We have extended Sternberg's approach by incorporating techniques from FCA and SNA. As demonstrated in this paper, we use FCA to model the questionnaire responses and to further identify individuals (who became the expert non-expert group) that behaved similarly to the peer-selected experts. FCA thus provides a qualitative modeling alternative to the typical quantitative modeling conducted on questionnaire data using statistical methods. While not presented, using SNA we can determine the likelihood of tacit knowledge flowing from the experts and expert non-experts to others within the organisation to minimize the loss or wastage of knowledge resources.

References

1. Bartel, H-G., Brüggemann, R., (1998) "Applications of formal concept analysis to structure - activity relationships" Fresenius journal of analytic chemistry 361(1) May 1st. :23 - 28
2. Brockmann, E., Anthony, W., (1998) "The influence of tacit knowledge and collective mind on strategic planning" Journal of Managerial Issues Pittsburg; Summer
3. Davenport, T., De Long, D., Beers, M., (1998) "Successful knowledge management projects" Sloan Management Review Winter (from ABI Proquest)
4. Fleck, J., (1997) "Contingent knowledge and technology development" Technology Analysis and Strategic Management December (From ABI Proquest)
5. Ganter, R., Wille, R., (1999) Formal concept analysis: Mathematical foundations Springer - Verlag Berlin Germany
6. Kollewe, W., (1989) "Evaluation of a survey with methods of formal concept analysis" in Conceptual and numerical analysis of data: Proceedings of the 13th conference of the Gesellschaft fr Klassifikation e. V. University of Augsburg, April 10 - 12, 1989 Springer -Verlag Berlin Germany :123 - 134
7. Richards, D., Busch, P., (2000) "Measuring, formalising and modeling Tacit Knowledge" International Congress on Intelligent Systems and Applications (ISA2000) December 12 - 15 2000
8. Scott, J., (1991) Social Network Analysis: A handbook Sage Publications London U.K.
9. Spangenberg, N., Wolff, K.E. (1988) "Conceptual grid evaluation", In Classification and related methods of data analysis. Spangenberg, Norbert and Karl Erich Wolff, Eds., Amsterdam, North-Holland: 577-580.
10. Spillane, R., (1983) Stress at work: A review of Australian research Department of Psychology University of Stockholm Report Number 35
11. Sternberg, R., (1995) "Theory and management of tacit knowledge as a part of practical intelligence" Zeitschrift fr Psychologie 203(4) :319 - 334
12. Wille, R., (1982) "Restructuring lattice theory: an approach based on hierarchies of concepts". In Reidel, D. Ordered Sets, Dordrecht, 1982, pp. 445–470.
13. Wille, R., (1992) Concept lattices and conceptual knowledge. Computers and Mathematics with Applications **23** (1992) 493–522.