The Impact of Databases on Knowledge Transfer: Simulation Providing Theory

Craig Schreiber
Kathleen Carley
Institute for Software Research, International
Center for the Computational Analysis of Social and Organizational Systems
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

Abstract

The use of information technology to facilitate knowledge transfer is a major concern for organizations. This paper focuses on the use of databases and the role they play in transferring knowledge. In particular, two data types are defined – task data and referential data. Virtual experiments are run using two varying conditions, task complexity and group experience. The analysis looks at how the use of the data types differ in moderating knowledge transfer for task complexity and group experience. A theory of data type use is then developed.

Contact:
Craig Schreiber
Institute for Software Research, International
Center for the Computational Analysis of Social and Organizational Systems
Carnegie Mellon University
Wean Hall 1325
Pittsburgh, Pennsylvania 15213
Tel: (412) 268-5866
Fax: (412) 268-2338
Email: craigs@andrew.cmu.edu

Keywords: Knowledge Management, Knowledge Transfer, Organizational Learning, Transactive Memory, Databases, Information Technology, Computational Organization Theory, Agent-Based Simulation

Acknowledgement: This work was supported in part by the NSF KDI IIS-9980109 and the NSF IGERT in CASOS-9972762. Additional support was provided by CASOS - the center for Computational Analysis of Social and Organizational Systems at Carnegie Mellon University (http://www.casos.ece.cmu.edu).

The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the National Science Foundation or the U.S. government.

The Impact of Databases on Knowledge Transfer: Simulation Providing Theory

Craig Schreiber and Kathleen Carley

Introduction

In today's information driven environment, organizations are deriving more opportunities and value from intellectual assets as opposed to physical assets. Organizations try to manage their intellectual assets through knowledge management. Therefore, knowledge management is the process of managing intellectual assets. Because the value of knowledge is so important, organizations try to create cultures where the sharing of knowledge among members is fostered. Knowledge management is centered on the members of the organization because, after all, knowledge and especially tacit knowledge is contained within the members. Some of the benefits expected from successful knowledge sharing are increases in performance, adaptation, collaboration and innovation. The organizational concern over knowledge management is an increasingly important issue (Nonaka and Takeuchi, 1995).

This paper focuses on the sharing (transfer) of knowledge and the effect it has upon organizational performance. In particular the use of database technology is explored to see whether it supports the transfer of knowledge. Two data types used in databases are defined – task data and referential data. Virtual experiments are run using two varying group conditions, task complexity and group experience. The analysis looks at how the use of the data types (databases) differ in moderating knowledge transfer for task complexity and group experience. This work adds to the literature by first defining a new concept – data type. Second, it begins to build a theory of knowledge transfer employing data type and the contingent conditions under which it applies.

Background

Knowledge as defined in this paper is contextual information. This definition of knowledge is used because with expertise comes the ability to apply what is known. So when members learn knowledge they learn the information and how to apply it in context. This definition is also used because the computational model represents knowledge in the same way. When agents learn a piece of knowledge they can readily apply this knowledge to the task and increase performance. There is no transitional phase between learning information and acquiring or creating knowledge.

There is a plethora of studies focusing on the transfer of knowledge within organizations. These studies have shown that successful knowledge transfer internal to an organization can help solve problems (Constant, Sproull and Kiesler, 1996; Brown and Deguid, 2000), increase performance (Rulke and Galaskiewicz, 2000) and provide a competitive advantage (Argote and Ingram, 2000).

But knowledge barriers exist within organizations. Barriers such as not knowing that members with the desired knowledge are in the organization or knowing they exist but not knowing what knowledge they hold. A key aspect of a successful knowledge transfer initiative is the ability to overcome these knowledge barriers and to have members interact.

Enter information technology. Organizations have turned to the use of information technologies to overcome knowledge barriers. By overcoming knowledge barriers, organizations can bring members and knowledge together. This can happen either through interactions with the technology alone or through member interactions that are facilitated by the technology.

While information technology is no panacea, there is extensive work showing how the use of information technologies within organizations have aided the transfer of knowledge and have connected members who were not previously connected (Constant, Sproull and Kiesler, 1996; Fulk, et.al., 1996; Carley, 1995; Bikson and Evaland, 1990; Leduc, 1979). In particular, organizations have implemented knowledge repositories as a means to link members with knowledge. Examples of knowledge repositories are catalog filing systems and databases. Databases are particularly useful because they can be accessed by many members.

Data Types: Task and Referential

In their study of weak ties, Constant, Sproull and Kiesler (1996) noted that information providers helped information seekers by providing technical information or referrals through information technology. 53% of answers received by information seekers contained referral information. Also, Contractor, et.al. (forthcoming) concluded in their study of corporate intranet use that reposited data was used more as a pointer to expertise (referral). Members of this group used the database to facilitate human interactions. While these studies found interesting significant results on the use of information technologies for transferring information, the focus was not on any distinction of data type. The statements about referrals were observations that the authors felt important enough to note. So what are the data types and how are they different?

Task data are knowledge related to the performance of a task or a solution to a problem. In most database applications, the member would search for specific knowledge that is required to complete a task and receives the specific knowledge the database has on the topic. The Eureka database for Xerox service representatives is a good example of this type of data (Brown and Duguid, 2000).

Referential data 'refers' a member to an expert in the topic of interest. In application, a member would search for a topic on which they need some knowledge. The knowledge needed could be for the performance of a task, a solution to a problem or anything else. The search would give the member a list of names of experts in that area of knowledge. The member can then contact the expert(s) and find the knowledge that is needed.

From the top level it can be seen that referential data differs from task data in that it gives a contact reference rather than direct knowledge and that a member has to take an additional step in order to obtain the desired knowledge. Thinking through these steps for each data type it is clear that the data types entail two different processes for obtaining knowledge.

Task data is a purely technical process whereas a member queries the database and obtains the results. If successful, the query results in the correct knowledge being transferred. The use of task data is a fairly straight-forward and simple process but referential data is more complex.

Referential data is a social process facilitated by technology. The member uses the pointer provided by the technology and through dyadic contact with the expert obtains an answer. If successful, the dyadic contact results in the correct knowledge being transferred but there are other implications since there is actual dyadic communication. The communication resulting from the reference may have structural implications on the social network. These new communications may develop weak or strong ties. It could also result in the bridging of structual holes. In addition, referential data is a way to recognize expertise within the organization. Recognition of expertise within groups leads to the sharing of unique knowledge that is otherwise left unshared (Stasser, Stewart and Wittenbaum, 1995).

Transactive Memory

Transactive memory (Wegner, 1986) refers to the idea that people have an understanding of other's domains of expertise. This understanding provides a framework from which members may attain knowledge that is external to their own. Although the seminal work concentrated on intimate relationships, subsequent work has shown that a strong transactive memory within a group increases performance (Moreland, Argote and Krishnan, 1996). More recent work shows that transactive memory is a property of work relationships and emphasizes the importance it has to organizations (Hollingshead, 2000).

Referential data as a technological form of transactive memory can compliment the human form of transactive memory. This is true for the knowledge seeker and the knowledge provider. The technological reference adds to the transactive memory of the knowledge seeker and the dyadic communication adds to the transactive memory of both the knowledge seeker and knowledge provider.

Model

The model used was Construct-TM (Carley, 1990) (Carley and Schreiber, 2002). Construct-TM is a multi-agent model of group interaction whereas the agents communicate, learn, and make decisions in a continuous cycle. The Construct-TM model is used because it has been validated several times (Carley, 1990; Carley and Krackhardt, 1996; Carley and Hill, 2001) and because it is able to represent transactive memory and database technology.

The representation of transactive memory is meaningful because referential data is a technological form of transactive memory. It is assumed that if an agent can use their own transactive memory to seek out another agent for knowledge then they will do so and not use the technological form. Otherwise, an agent without the ability to use their own transactive memory will always turn to the database and the usage and value of referential data could potentially be overstated. The transactive memory of agents is not always accurate or complete and changes over time as the agents communicate and learn.

Each agent has both task and transactive knowledge and when interaction occurs either type of knowledge can be communicated. When task knowledge is communicated the receiving agent simply learns the task knowledge. When transactive knowledge is communicated it is a referral from the transactive memory of agent X telling agent Y to go to agent Z for the knowledge. It is then up to agent Y to follow up with agent Z.

Databases are represented as agents with distinct information processing capabilities. By representing the database as an agent it is an integral part of the group and analysis can be done on the interactions between the human agents and the technology and the impact of learning or adding knowledge to the database on the behavior of the group. Two types of databases are represented to correspond with the two data types. Task databases contain only task data and referential databases contain only referential data.

The Construct-TM model measures task performance by using the binary choice task. For the binary choice task each agent decides whether there are more 1's or 0's in a binary string based on the knowledge they have at the time of the decision. Performance, therefore, is the accuracy of these individual agent decisions as compared to the true answer.

Virtual Experiment: The Impact of Task Complexity and Experience on Performance

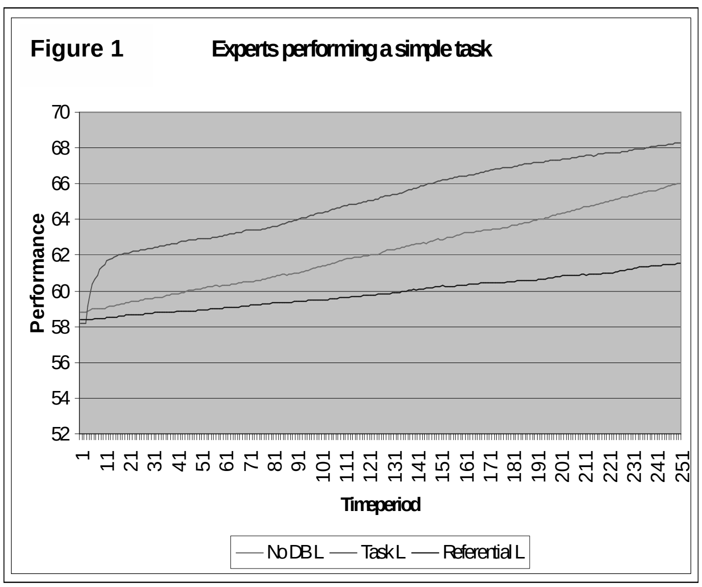
Previous work suggests that, holding the structure of the organization constant, an increase in task complexity can degrade organizational performance (Carley and Schreiber, 2002) and that both task and group experience can lead to improved performance (Littlepage, 1997). Given that task complexity and group experience has an effect upon performance the question is can the use of a database moderate these effects? If so, do the data types have differing moderation effects?

The design for this experiment is 3x3x2 with varying conditions for database technology (no database, task database, referential database), task complexity (low, medium, high) and experience (experienced, non-experienced). The no database condition is a baseline so that any moderating effects can be detected. The experiment is monte carlo with 100 runs per cell.

The dependent variables measured in this experiment are group performance, contributions to the database, retrievals from the database and referrals. Group performance is the main variable of interest. The contribution and retrieval variables are collected so that any difference in interaction with the databases across conditions can be explored. The referral variables, from agent transactive memory and from the referential database, are also collected to explore any variance across conditions.

Virtual Experiment Results

The results of the experiment reflect the conclusions of other studies on task complexity and experience. As task complexity increased group performance decreased Also, experience improved performance. What this experiment shows in addition is that the use of a database does have an effect and moderates both task complexity and experience. In addition, the data types have different moderating effects.



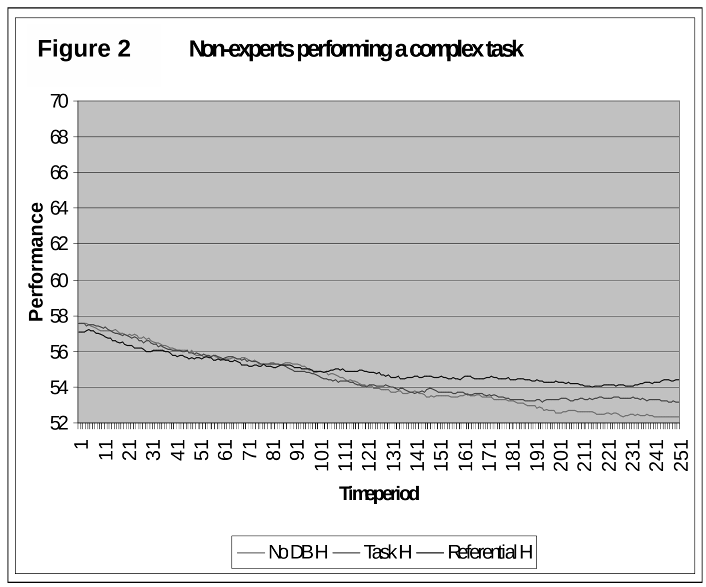


Figure 1[1] shows that experienced agents performing a simple task improve organizational performance over the baseline by using a database with task data. On the other hand for this same condition, if the agents use a database with referential data then organizational performance decreases. The same result occurs for non-experienced agents performing a simple task just the impacts on performance are not as severe do to the inexperience of the agents. The opposite moderation effect occurs for agents performing a complex task. For the experienced group, the use of a database with referential data does not decrease performance but does not increase it either. In contrast to experienced agents performing a simple task this is a notable difference. The non-experienced group benefits tremendously by using a referential database, see Figure 2[1]. Although performance decreases overall due to the inexperience and high complexity of the task, referential data helps to lessen the drop in performance. For the use of

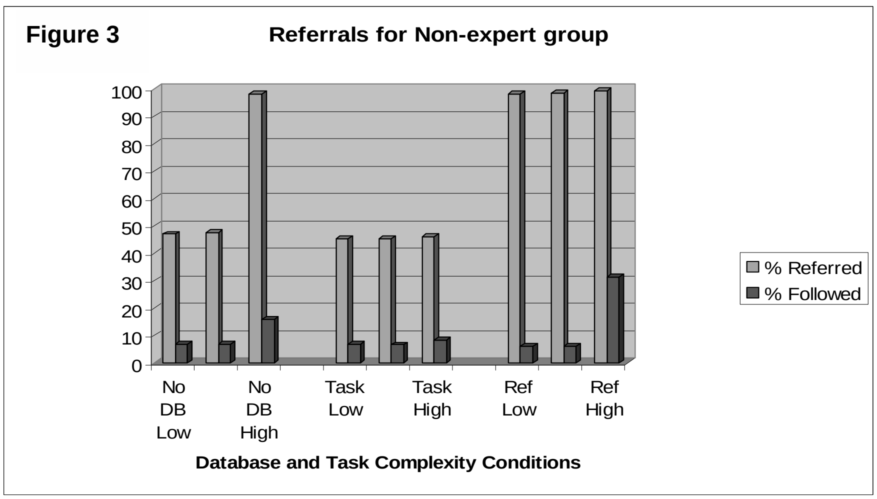
[1] The medium task complexity condition is not described in the results because the results for this condition just fall in between those of the low and high task complexity conditions.

a task database with a complex task, the experienced group performs worse than the baseline and the non-experienced group sees a slight improvement on performance over time but not much. Again, this is opposite of the moderation effect of using task data while performing a simple task

For simple tasks, task data is transferred facilely and because it is a simple task the task knowledge is easily learned by the agents. When task data has a positive impact it is usually noticeable in the short-run due to the ease of learning. Referential data, on the other hand, slows down performance because it takes time to obtain and follow up on the reference to learn the knowledge.

For complex tasks, knowledge is not so easily learned and diffused. The value of task data use declines significantly because the agents just are not learning fast enough. Referential data gets more of the needed knowledge into the heads of the agents. Figure 3 shows how the use of referrals helps knowledge transfer and improves the performance of a complex task. What is important to note is that it is not just the referral but the follow-up on the referral and the successful transfer of knowledge.

In the task database condition, the knowledge that the agents receive from the database is of little help to the performance of the complex task. The agents spend most of their time looking to the database for knowledge but the database is incomplete. The amount and complexity of knowledge needed to perform the task is much higher than in the simple condition. Therefore, agents cannot contribute to the database fast enough.



The referential database does provide help to the agents performing a complex task. Agents contribute referential data which is much less in terms of volume and complexity so the referential database is more complete. When using a referential database the agents receive more referrals under all conditions. In the low and medium conditions, the agents do not need to follow-up on the referrals as much because the task is not complex enough but receiving the referral knowledge is taking up time that would be better spent by learning the simpler task knowledge quickly. When the task is complex the agents receive referrals and follow up on these referrals as well. Even though this still takes up more time than in the task database condition it transfers the complex knowledge that is needed for the task and therefore performance is improved although it is much more of a long-run impact.

It should be noted that referrals in the baseline condition are only agent referrals but in the referential database condition are agent referrals and referential data from the database. Obviously the referential database is connecting more agents together than does the use of agent referrals only. For complex tasks, the amount of referrals for the baseline and the referential database conditions are close to identical but the follow ups are definitely in favor of the referential database, see Figure 3.

Conclusion

The results show that task complexity and experience effects are moderated by the use of data types. In short, a theory has been proposed that can now be empirically tested. Task data is helpful in transferring knowledge for simple and moderate task complexities. Referential data is helpful in transferring knowledge in complex tasks, especially for non-experts. By understanding the task complexities that an organization faces, management can use appropriate data types within database technology to enhance knowledge transfer and produce successful knowledge management.

Limitations

The computational model is limited in that all knowledge is treated as equally timely, valuable, factual and interpretable. Also, task and referential data have access to equal knowledge. It may be the case that referential data has a greater ability to tap into tacit knowledge.

Future Directions

A next step would be to look at the combined interaction effect when both data types are present. Before this can be modeled, empirical work needs to be done to determine how group members interact with the data types when the option to choose one or the other is available. It may be that one data type is generally preferable over the other. It could be that it is contingent on some context. In any case, the study of these interactions will provide valuable insight into the use of databases and for knowledge transfer.

Another avenue to explore would be a longitudinal study focusing on the change in the social network due to the use of referential data. Such a study will be useful for understanding the structural changes that occur whether these changes are intended or not. Realizing and predicting the implications of structural changes due to the use of referential data will provide guidance to strategically plan for desired results or contingencies.

References

Argote, L. and Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage of firms. *Organizational Behavior and Human Decision Processes*, 82: 150-169.

Bikson, T. and Eveland, J. D. (1990). The interplay of work group structures and computer support. In J. Galegher, R. Kraut and C. Egido (Eds.), *Intellectual teamwork: Social and technological bases of cooperative work*. (pp. 245-290). Hillsdale, NJ: Lawrence Erlbaum.

Brown, J. S. and Duguid, P. (2000). Balancing act: How to capture knowledge without killing it. *Harvard Business Review*, 78(3): 73-80.

Carley, K. M.. (1990). Group Stability: A Socio-Cognitive Approach. In Lawler E., Markovsky B., Ridgeway C., and Walker H. (Eds.) *Advances in Group Processes: Theory & Research.* Vol. VII. (pp. 1-44). Greenwhich, CN:JAI Press.

Carley, K. M.. (1995). Communication Technologies & Their Effect on Cultural Homogeneity, Consensus, & the Diffusion of New Ideas. *Sociological Perspectives*, 38(4): 547-571.

Carley, K. M. and Krackhardt, D. (1996). Cognitive Inconsistencies and Non-Symmetric Friendship, *Social Networks*, 18: 1-27.

Carley, K. M. and Hill, V. (2001). Structural Change and Learning Within Organizations. In A. Lomi and E. R. Larsen (eds), *Dynamics of Organizations: Computational Modeling and Organization Theories*. Menlo Park, CA: MIT Press/AAAI.

Carley K. M.. and Schreiber, C. (2002). *Information Technology and Knowledge Distribution in C^3I Teams*, 2002 Command and Control Research and Technology Symposium, Monterey, CA.

Constant, D., Sproull L., and Kiesler, S. (1996). The kindness of strangers: The usefulness of electronic weak ties for technical advice. *Organization Science*, 7(2): 119-135.

Contractor, N., Brandon, D., Dandi, R., Huang, M., Palazollo, E., Ruta, C., Singh, V., Su, C. (forthcoming). *Multi-theoretical multi-level model of information retrieval and allocation among human and non-human agents*.

Fulk, J., Flanagin, A., Kalman, M., Monge, P., and Ryan, T. (1996). "Connective and Communal Public Goods in Interactive Communication Systems." *Communication Theory*, 6, 60-87.

Hollingshead, A. (2000). Perceptions of expertise and transactive memory in work relationships. *Group Processes and Intergroup Relations*, 3(3): 257-267.

Leduc, N. (1979). Communicating through computers. *Telecommunications Policy*, pp. 235-244.

Littlepage, G. (1997). Effects of Task Experience and Group Experience on Group Performance, Member Ability, and Recognition of Expertise. *Organizational Behavior and Human Decision Processes*, 26(2): 133-147.

Moreland, R., Argote, L., and Krishnan, R. (1996). Socially shared cognition at work: Transactive memory and group performance. In J. Nye & A. Bower (Eds), *What's social about social cognition?* (pp. 57-84), Thousand Oaks, CA: Sage.

Nonaka, I. And Takeuchi, H. (1995). *The knowledge creating company: How Japanese companies create the dynamics of innovation*, New York, NY: Oxford University Press.

Rulke, D. L. and Galaskiewicz, J. (2000). Distribution of knowledge, group network structure, and group performance. *Management Science*, 46(5): 612-625.

Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind, In B. Mullen & G. R. Goethals (Eds), *Theories of group behavior*, (pp. 185-205), New York: Springer-Verlag.

Stasser, G., Stewart, D. D., and Wittenbaum. G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology*, 31: 244-265.