Chapter 15

Outsiders on the Inside: Sharing Know-How Across Space and Time[1]

Thomas A. Finholt, Lee Sproull, and Sara Kiesler

Summary

This study examined how employees of a global corporation sought out technical information from employees distant from them. Data were drawn from the contents and access records of two computer-based employee message archives. The "peer archive" contained questions and answers exchanged over a company-wide electronic mailing list by ordinary employees. The "expert archive" contained questions from employees and answers from designated experts. We predicted that archive accesses by employees would increase with their geographic distance from engineering headquarters, and that accesses would increase to the extent that the archive contained know-how information. Content analysis demonstrated that the peer exchanges were more informal, more personal, and were more likely to contain specific references to the company's products, tools, and customers, all indicators of know-how information. Analyses of use showed that workers accessed the peer archive more than the expert archive; this difference increased with greater geographic distance from engineering headquarters. The results suggest that know-how archives can be useful in connecting distant employees seeking and providing technical information across space and time.

Citation:
Finholt, T., Sproull, L., & Kiesler, S., (2002). Outsiders on the inside: Sharing know-how across space and time. In P. Hinds & S. Kiesler, *Distributed Work*, (1st ed., pp. 357-379). Cambridge: MIT Press.

Organizations are focused on improving access to and use of organizational knowledge--both "know-what" and "know-how" (Brown and Duguid, 1998). Know-what is explicit rules and principles. Know-how is the ability to put know-what into practice. For example, a help line operator may understand general principles of computer operating systems, but this understanding is useful only if it can be deployed to solve a caller's particular problem. Know-what is relatively easily transmitted through books, manuals, or formal instruction and computer-based documents and databases. Know-how, by contrast, is often difficult to transmit in documents and formal instruction because it is embedded in work practices

Technicians and engineers have always shared know-how in informal face-to-face conversations (e.g., Allen, 1977; Kusterer, 1978; Latour and Woolgar, 19xx, von Hippel, 1994). Face-to-face conversation ties know-how to a particular place– because the likelihood of face-to-face interaction diminishes with increased distance (see Kiesler & Cummings, this volume). Face-to-face interaction also typically ties know-how to a particular time because it is usually ephemeral, i.e., it is not recorded and archived for later use.

The dilemma for field engineering is that geographic and temporal distance impose barriers to developing and sharing know-how in face-to-face interaction. Peers or experts with relevant know-how may be located far from the field engineer's customer site. Or they may be physically close, but their potentially helpful conversations may have occurred prior to a point at which they were relevant or needed. To the extent that wider access to know-how represents a strategy for improving field engineering for customer support, organizations must find ways to develop and disseminate know-how that do not depend on geographic or temporal co-location.

Widespread connectivity to computer networking technology, in some cases, has supported the development and dissemination of know-how across distance. Constant, Sproull, and Kiesler (1996) describe how employees of a global corporation found technical information by posing know-how questions over their corporate computer network to mass mailing lists. Although these were often questions about company products, they were not questions that could be answered by the formal documentation about how things were supposed to work. On average, a question received private email replies from eight people. Only 8% of those answering knew the question asker. Yet employees found replies to be very useful and, in fact, in half the cases the replies actually solved the employee's problem.

The broadcast question with private replies can be very useful to the individual questioner, but this mechanism does not spread know-how more broadly. For this reason, organizations may create technology such as bulletin boards or electronic mailing lists ("distribution lists") for publishing both questions and answers to anyone who subscribes. Applications of this nature allow many people to read and contribute to both the questions and the replies.

One drawback of email lists for exchanging know-how is that, even though it is convenient, people tend to ignore messages that are irrelevant to their immediate concerns. The half-life of a message thread, that is a question and all replies to it, is typically four days or less (Galegher, Sproull, and Kiesler, 1998). Thus, the benefits of information sharing through email lists are most accessible to geographically distant employees who are paying attention when a conversation occurs. Temporally-distant employees, that is, those who are not interested in a topic when it is being discussed but who develop an interest in the topic later, may not benefit from electronic mailing lists.

A searchable archive of distribution list email messages or bulletin board posts can extend the benefits of electronic interaction over time and to a broader audience (e.g., Ackerman 1994; Ackerman et al. 1996). It is technically feasible to archive electronic exchanges of know-how contained in the kinds of exchanges documented by Constant, Sproull, and Kiesler (1996). Every employee who broadcasts a question and receives replies can submit both the question and its replies to an archive, or exchanges can be archived automatically.

Even though it is technically-feasible to archive spontaneous exchanges of know-how, it is not clear that reading archived reports of someone else's problem accompanied by suggested solutions will, in fact, be useful. If archived messages describing problems are couched in idiosyncratic terms, they may not substantially resemble the problem of the person searching the archive. If archived replies suggesting solutions are also idiosyncratic, they may be deemed inapplicable. The original question asker might have engaged in additional clarifying communication with the person who replied, such as a phone call or private email, but someone accessing an archive has much less opportunity to ask follow-up questions of those who provided solutions. Even if archived replies are accompanied by information about the source, such as the source's location, this information could be obsolete by the time it was accessed.

Thus, whereas contemporaneous electronic exchanges plausibly can produce and disseminate know-how across distance, archives of those exchanges might not be useful to people needing know-how across time. Perhaps more useful would be archives of questions sent to officially-designated experts, with the experts' replies. These archives would be more authoritative and more generalizable across both distance and time, to the extent that expert respondents focus on universal and enduring principles underlying questions. On the other hand, designated experts, because of their more universal perspective, may provide overly general rules and canonical information rather than pointed stories and advice. Furthermore, the average reply from a designated expert, whose job requires replying to all questions, may be less highly motivated by the wish to provide a quick, helpful answer than the average reply from a peer who chooses to reply to a particular engineer's question (and not to others). Finally, archives containing several coworker replies to each question, even if the average reply is of lower quality than that of the experts' replies, are likely to offer multiple ways of thinking about a problem rather than a single "official" answer.

These considerations led us to hypothesize that, despite the unofficial and voluntary source of its contents, an archive of informal questions and answers by peers would be preferred by field engineers over an archive of experts' answers to questions. Also, because a company's engineering headquarters tends to produce the highest concentration of engineering know-how in one place, we hypothesized that the further away field engineers are from their company's engineering headquarters, the more they will use an archive of peer exchanges, and the more they will prefer it over an archive of expert exchanges.

A Study of Know-How Archives for Field Engineers

In 1988, the first author collected nearly a year's data on field engineers' use of message archives at Tandem Computers, a Fortune 500 computer manufacturing company (since acquired by Compaq Computer) (Finholt, 19xx). One archive contained questions by field engineers and replies to them by fellow employees in the company; the other archive contained questions by field engineers and answers from designated experts. Here, we describe engineers' use of the archives to understand how people distant in space and time from the original exchange might find such information useful. We note at the outset that these are historical data; computer networks and people's use of them have developed markedly since the 1980s. However, while contemporary implementations may use more modern technology, the purpose and function of the Tandem archives are similar to mechanisms widely used today, particularly with regard to worker-to-worker and worker-to-expert exchanges. Further, for its time, the Tandem network and applications were very advanced. For example, 87% of Tandem employees across the world had access to the networked archives with an easy-to-use interface. This high proportion of employees connected to the network, while unusual in the late 1980s, mirrors current levels of connectivity and use within corporate intranets. Hence, these historical data seem well suited to testing our arguments and hypotheses.

At Tandem, as at other organizations, employees used face-to-face conversations and meetings to develop and exchange know-how. However, opportunities for conversations and meetings were limited by geographic dispersion. Sixty five percent of Tandem's 10,077-person workforce was distributed over 306 field offices on five continents. The Tandem field engineers installed and supported Tandem products at customer sites all over the world; thus most field engineers worked at a distance from their engineering colleagues.

Tandem supported a corporate computer network that reached virtually all employees. Further, management encouraged heavy use of the corporate network for both work-related and personal communication. Electronic mail and other computer-mediated communication were expected to span barriers of distance and time to improve contact among distributed employees and to speed the flow of information among remote sites. Tandem employees used a "class" email structure to organize the flow of electronic mail. They exchanged electronic messages with each other and with work groups via "first class mail." They broadcast work-related electronic messages to the entire company via "second class mail." They broadcast non-work messages to their local sites or regions via "third class mail." Each class of mail was delivered into a separate in-box on the employee's computer display, thus providing a simple way of organizing incoming mail.

At Tandem, field engineers used the corporate computer email system to seek and provide work-related information. They could query a broad audience through use of second class mail or target queries to specific respondents through use of first class mail. Tandem employees' broadcast question and private reply system described above and by Constant, Sproull, and Kiesler (1996) was viewed as so useful that in the early 1980s, the field engineering organization began to save these questions and answers into message archives accessible throughout the Tandem corporate computer network. Tandem employees accessed the archives through an archive server with a simple, common interface. The archive server allowed employees to specify a variety of complex searches by combining keywords and data fields. One archive, which we will call the peer archive, contained the technical information exchanges that occurred via second class electronic mail. Replies in these exchanges came from ordinary members of Tandem's field engineer workforce who voluntarily responded to a broadcast second-class question with information and advice. The peer archive resulted from the practice of saving questions, and the replies to these questions, as public "reply files." Reply files were incorporated into the peer archive when a member of the field engineering headquarters staff, seeing an announcement of a public reply file, would retrieve the file via the corporate network and add it to the peer archive. Figure 15-1 shows an example of a reply file stored in the peer archive.

The second archive, which we will call the expert archive, consisted of questions sent in first class mail from ordinary employees to designated experts at Tandem engineering headquarters. In these exchanges, questioners addressed requests for help to specific computer mailboxes publicized within the organization and corresponding to various Tandem products;

designated engineers with special expertise with these products read and replied to these requests. Typically, there was a one-to-one correspondence between experts and mailboxes, although some boxes were read by groups of experts. Each expert filed a copy of the question and his/her reply into the expert archive. About half of the exchanges via second-class mail were archived (Constant, Sproull, and Kiesler ,1996); it is unknown what fraction of queries to the designated expert mailboxes were filed in the expert archive. Figure 15-2 shows an example of an exchange stored in the expert archive.

Figures 15-1 and 15-2 about here

Method

We examined employee login statistics over forty-five weeks in 1988 to the peer and expert archives. Employees logged into the archives to search for or read archive content, but not to contribute new content (see the comments in the previous section on archive creation). For each login, the archive server wrote a record indicating when a login occurred, the site from which a login was initiated, the commands issued during a login session, and the identification numbers of archived messages returned during a login session. Archive use was captured at the site level because Tandem's archive logs did not definitively map archive logins to identifiable individual users. The risk of ecological fallacy from this approach, however, was mitigated by the finding that at the subset of sites where individual use could be verified there was low variance across employees in message archive use.

We performed content analyses on a sample of the peer and expert archived reply files stored from 1983 to 1988, by randomly selecting 10% of the files stored each month during this period. In months where only one to nine files were stored into an archive one file was selected at random to ensure a contiguous sample across time. As a result of this strategy 78 (12.0%) of the expert files were sampled compared to 465 (10.1%) of the peer files. The discrepancy between the expert and peer samples reflected the high number of months when only one to nine files were stored into expert. We operationalized know-how through the frequency of three types of critical keywords indicative of know-how: personal experience keywords, referral keywords, and practical knowledge keywords. A text scanning program was written in SNOBOL4+ (Emmer, 1985) to automatically process the full text of each sampled reply file to extract content. The program matched content in the reply file with lists of keywords selected as markers of the different content types. The keywords for personal experiences were first person pronouns and contractions, which indicate an informal style. The keywords for referrals were pointing verbs, such as "call" and "contact." Keywords for practical knowledge were names of products and references to "customer." The program tabulated the occurrence of keywords per reply file which were then expressed as rates per thousand words to allow comparison between the peer and expert files, which typically differed dramatically in length.

We used corporate data bases to ascertain or calculate the values of control variables used in regression analyses. These included distance in miles of a site from engineering headquarters; median tenure of employees at a site; number of employees at a site; and presence of sales and customer support at a site.

Results

Table 15-1 contains descriptive data showing the nature of the two archives. In December, 1988, the peer archive contained 4582 information exchanges, each with a question and its associated answers. The expert archive contained 651 information exchanges. Consistent with observations of the creation of second class mail reply files, the average reply file in the peer archive had more replies, more unique repliers, and more unique replier sites than the average reply file in the expert archive.

Tables 15-1 and 15-2 about here

Preference for Peer Information. In support of our hypothesis that archives of coworker exchanges would be preferred to archives of “official” exchanges with experts, the peer archive was accessed much more frequently than the expert archive. Moreover, the comparisons summarized in Table 15-2 show that questions and replies in the average reply file in the peer archive had more personal experience content, more specific referral content, and more practical knowledge content than the average reply file did in the expert archive.

Figures 15-1 and 15-2 illustrate how technical knowledge was conveyed in the peer and expert archives. The peer reply file in Figure 15-1 shows first-person description of events: Repliers 3, 5, and 6 each mentioned their own experience. By contrast, the expert reply file in Figure 15-2 mentioned the experience of a formal sub-unit, “engineering,” rather than referring to the replier’s own experience. In the peer reply file, repliers used more informal contracted verb forms (e.g., it’s; don’t) rather than more formal full verb forms (e.g., it is; do not). By contrast, in the expert reply file, the replier used more formal full verb forms. Figures 15-1 and 15-2 also illustrate how referral content is conveyed. In the peer reply file, repliers provided several pointers to other employees or to information generated by other employees. By contrast, the expert offered no such pointers. Finally, in the peer file, Replier 6 noted the need for close customer consultation, and Replier 3 noted the possibility of variance across customer sites in lines 42 to 46. By contrast, the expert provided no information about customer experience.

Effects of Distance. We postulated that as the distance between field sites and Tandem engineering headquarters increased, physical proximity to the heaviest concentration of knowledgeable employees decreased, reducing the opportunity for face-to-face exchange of know-how. Therefore, distance from engineering headquarters should predict accesses to the networked archives, especially to the peer archive. Geographic distance was measured as the distance in statute miles between a given site and Tandem engineering headquarters in Cupertino, California.

Table 15-3 summarizes the descriptive statistics for the variables used in our analyses. These analyses included only data from those sites with a documented connection to the Tandem computer network. Of the 306 Tandem sites in 1988, 151 had documented connections to the computer network. These 151 sites represented 9060 employees, or 87 percent of the Tandem workforce. 141 of the 151 sites generated complete data.

Table 15-3 about here

The average Tandem network site was more than 2,000 miles away from engineering headquarters, and only 36 out of the 141 Tandem network sites were located within 500 miles of it. These data indicate the high degree of geographic dispersion among Tandem network sites.

Data on logins per employee suggest that the computer archives were not frequently used at any sites, but the variance and range of the data suggest that the computer archives played a larger role at some sites than others. Specifically, 74% of the 141 sites had sales affiliations, and these sites accounted for 85% of all logins to the peer archive and 83% of all logins to the expert archive during the 45 weeks of this study. This pattern reflected the close association between sales and field engineering workers when closing deals and supporting customers.

We examined message archive use in a hierarchical regression against: a) control variables correlated with geographic distance and the dependent variables, such as seniority of employees at a site, size of a site, whether a site was part of the sales organization, and the availability of the computer network at a site; b) geographic distance; c) type of archive accessed, that is, peer versus expert; and d) the interaction of distance and archive type.

A significant proportion of sites recorded no archive use during the period of the study. Therefore, archive use was analyzed using the standard zero, lower-bound Tobit model (Tobin, 1958). The Tobit model is an alternative to ordinary least squares (OLS) regression designed for situations where the dependent variable is limited (in this case, it couldn't assume values less than zero) and a large number of observations are clustered at zero. The concentration at zero violates assumptions for OLS analysis. Simply modeling the probability of a limit or non-limit value, as a probit model would do, throws away useful information (i.e., collapses all non-limit cases into a single class). The Tobit model, then, is a hybrid of the regression and probit approaches. Under Tobit, models are tested that account for both the effect of independent variables on whether a dependent variable is at limit or non-limit, and the effect of independent variables on the dependent variable when above the limit value.

The Tobit regressions tested hypotheses about computer archive use through hierarchical comparison of four theoretical models starting with the control variables alone, then adding distance from engineering headquarters, then adding archive type, and culminating with a term for the interaction between distance and archive type. The full model had the following form:

$$\mathit{LOGINS\ PER\ EMPLOYEE} = b_0 + b_1\mathit{SENIORITY} + b_2\mathit{SIZE} + b_3\mathit{SALES\ FUNCTION} + b_4\mathit{REGISTERED\ ACCOUNTS} + b_5\mathit{DISTANCE} + b_6\mathit{ARCHIVE} + b_7\mathit{DISTANCE{*}ARCHIVE} + e\ ,$$

where a) *LOGINS PER EMPLOYEE* represented the number of logins to the computer archives per employee at a site; a) *SENIORITY* represented the median organizational tenure in years of employees at a site; b) *SIZE* represented the number of employees at a site; c) *SALES FUNCTION* indicated affiliation with the sales organization as determined from the corporate directory (1 = employees at the site reported to the sales organization; 0 = no employees at the site reported to the sales organization); d) *REGISTERED ACCOUNTS* represented the percentage of employees with registered computer accounts at a site; e) *DISTANCE* represented the distance in miles to Tandem engineering headquarters; f) *ARCHIVE* indicated the archive type (1 = peer; 0 = expert); and g) *DISTANCE*ARCHIVE* represented the interaction of distance with archive type.

Table 15-4 summarizes the results of the hierarchical comparison of the theoretical models. The Moderated Model was the best fit to the data (Pseudo $R^2 = .29$, $c^2(7) = 112.74$, $p < .01$). Two control variables, seniority and sales function, were significant: $b_{seniority} = -.22$, $t(274) = -3.91$, $p < .01$; $b_{sales\ function} = 1.18$, $t(274) = 4.65$, $p < .01$, indicating that lower seniority and affiliation with the sales organization led to greater archive use. The coefficient for archive type was significant: $b_{archive\ type} = .1.08$, $t(274) = 7.91$, $p<.01$, confirming that, with controls, use of the peer archive exceeded use of the expert archive. Finally, the overall effect of distance lost significance when the interaction was included in the model, but the coefficient for the interaction term was significant: $b_{interaction} = .18$, $t(274) = 2.29$, $p<.05$. This finding supports our hypothesis that accesses to the peer archive would be higher with more distance than accesses to the expert archive. The significant change in the value of the log-likelihood score, represented in the last row of Table 15-4 as "*G* of improvement," also supports the conclusion that there was a differential effect of distance on peer use compared to the expert archive, although the change in the magnitude of the Pseudo R^2 was not great. Because it is plausible to expect that the effect of distance is not linear, we reran the analyses using log of distance and using a dichotomous variable for distance (0=at engineering headquarters; 1=elsewhere). The pattern of results was the same across all treatments of distance.

Table 15-4 about here

Analyses of Alternative Hypotheses

Since the peer files contained more information archived each month, they would also be likely to contain more recent information on any particular topic. It is therefore possible that higher use of the peer archive stemmed from a need for recent information and not a need for contact with distant knowledgeable peers. A preliminary goodness-of-fit test was conducted to assess the effect of timeliness on access to stored reply files in the peer and expert archives. There was disproportionate access to new peer archive reply files relative to old ($c^2_{(5,.01)} = 68.4$, $p<.01$), but no similar pattern for access to new and old reply files in the expert archive ($c^2_{(4,.05)} = 3.6$, NS). Based on this finding, the age of reply files read at a site was regressed in an OLS model on the key independent variables (remoteness, archive type) and the control variables (seniority, sales function, size, and network access). The sample for this analysis was restricted to sites that used the archive during the study period, therefore n = 84. As with the Tobit regressions, above, the OLS regressions tested four hierarchical models, culminating in the following full model:

$$AGE\ OF\ REPLY\ FILES = b_0 + b_1 SENIORITY + b_2 SIZE + b_3 SALES\ FUNCTION + b_4 REGISTERED\ ACCOUNTS + b_5 DISTANCE + b_6 ARCHIVE + b_7 DISTANCE*ARCHIVE + e$$

where *AGE OF REPLY FILES* represented the median age in years of reply files accessed at a site.

The full model was the best fit to the data, but the model did not perform well (Adjusted $R^2_{\text{full model}} = .03$, $F(8, 159) = 1.81$, $p<.10$). The inability to detect consistent and strong relationships between the age of files read, site distance, and archive type suggested that timeliness did not explain the effect of distance in archive use.

A second alternative explanation is that higher use of the peer archive reflected normative use patterns. According to social influence theory, individual choices about use of communication technology reflect patterns of use demonstrated by peers and fellow group members (e.g., Fulk & Boyd, 1991). Analysts impute social influence through examining variation in use patterns across offices or sub-units of organizations (e.g., Rice & Aydin, 1991). Social influence theory would predict that message archive use will increase as the proportion of employees at a site using archives increases. While the field engineers did spend considerable time alone or in small teams at field sites, they also had assigned home offices where they spent significant time together. Therefore we explored whether the behavior of others within these offices influenced use of the message archives. We regressed individual archive use, operationalized as the median archive logins per archive user at a site, on the independent variables (distance, archive type, and a social influence measure--number of archive users per site) and on the control variables (seniority, sales function, size, and network access). Again, because of censoring on the dependent variable, five theoretical Tobit models were hierarchically compared, culminating in the following full model:

$INDIVIDUAL\ ARCHIVE\ USAGE = b_0 + b_1 SENIORITY + b_2 SIZE + b_3 SALES\ FUNCTION + b_4 REGISTERED\ ACCOUNTS + b_5 ARCHIVE\ USERS + b_6 DISTANCE + b_7 ARCHIVE + b_8 DISTANCE*ARCHIVE + e$ Where *INDIVIDUAL ARCHIVE USAGE* represented the median archive logins per archive user at a site.

The main effects model (i.e., everything except the interaction term) was the best fit to the data (Pseudo R^2 = .40, $c^2(7)$ = 87.06 p < .01). In this model, the number of archive users at a site was significant ($b_{archive\ users}$= .05 $t(274)$ = 5.05, p<.01) and important (B = .35). But distance also was significant ($b_{distance}$ = .18, $t(274)$ = 3.02, p<.01) and important (B = .29). Therefore, although the number of archive users at a site influenced individual archive use, it did not cancel effects due to distance or archive type.

Discussion

The results from this study support our argument that computer message archives can facilitate dissemination of know-how for a geographically dispersed workforce. Our analyses showed a pronounced preference by engineers for consulting the archive containing informal advice from peers as compared with the archive containing advice from designated experts. Moreover, this preference increased with greater geographic distance from engineering headquarters, suggesting that the peer archive performed better than the expert archive in offering remote workers an alternative to physical proximity for acquiring know-how. The preference for the peer archive did not seem to be due to a preference for more recent information. It did seem to result, partly, from a social learning process in which a person was more likely to use an(y) archive if more people in his/her immediate environment also used it.

This study does not tell us exactly what it is about informal email exchanges of questions and replies that are useful when they are archived. It may be partly the specific contents of personal experience, referrals, and practical knowledge. It may be partly the diversity and multiplicity of replies to each question. Note, for example, in Figure 15-1 the questioner asked about the feasibility of converting a program to a more recent version. The repliers were employees in the Netherlands, Germany, France, Missouri, and California, suggesting broad experience with the program within Tandem. Further, the repliers agreed on several points. For instance, Repliers 1, 2, and 6 all confirmed the existence of on-line help for the conversion process and that this help was valuable. Repliers 3, 5, and 6 all confirmed that the type of conversion referred to in the question was normal and that others within Tandem had performed such conversions. In this instance, consistency among the replies crudely indexed reliability and the number of replies roughly validated the original question's significance. The existence of multiple replies did not guarantee that the questioner would find the best solution to the conversion problem, but the multiple replies improved the probability that the questioner would be able to discern the best solution of those shared by the repliers.

Other unmeasured differences might have existed between the two archives we studied, and also could account for differential access. For example, engineers might have sent questions to experts mainly when their questions were of a general nature, thinking the experts would not want to be bothered with idiosyncratic local problems. That would lead the experts to reply in kind, leaving the expert archive full of general advice. Another possibility is that the experts archived less diligently than peers did, leaving their archive with advice less valuable to the average engineer. From our previous study of the original exchanges (Constant, Sproull, and Kiesler, 1996), we know that employees who used the second class system announced a reply file in only about half of all cases. Presumably, they announced only the reply files they thought their fellow engineers might find of use.

None of the processes described above undermines our conclusions about the preference for archives of exchanges of know-how among peers. However, other processes might change preferences for archives of know-how exchanges. For example, if all employee questions and answers were automatically archived by machine instead of by humans judging their quality and usefulness, an archive might contain so much useless information that nobody would use it.

Implications

This study did not explore the impact of message archive use on Tandem's productivity or profitability. However, the results do have some bearing on the relationship between information system design and organizational performance. Hutchins (1995) hypothesizes that the efficiency of key organizational cognitive processes, such as information distribution, can affect an organization's effectiveness. The results of our study suggest that the efficient distribution of know-how can be supported by computer-based information systems. In geographically distributed settings where access to know-how is critical, organizations with information systems that resemble Tandem's employee message archive may outperform organizations with more traditional information systems. A key direction for future work, then, will be a test of Hutchins' hypothesis through comparison of performance gains produced by creating greater access to know-how balanced against the cost of creating this greater access.

Historically, know-how has been viewed as local knowledge about work practice that emerges from specific experiences in particular settings. An unspoken theme of this paper has been the function of message archives in transforming know-how from information held and circulated among local workers into information held and circulated globally. Transforming local knowledge into global knowledge is probably beneficial under certain circumstances but not others. For example, when workers share a common technology, useful insights about that technology that emerge in a single locality can be generalized to a wider audience. Thus an engineering workaround developed at one site for a standard product is likely to apply to that product wherever it is used. However, local insights that are tied to cultural and environmental idiosyncrasies may not generalize well. For instance, know-how about a sales approach successful with Tandem's North American customers may be completely inappropriate for Asian customers.

Assuming that an organization can benefit from transforming local knowledge into global knowledge, producing this transformation through the creation of message archives may depend as much on organizational policies regarding information sharing as it does on the design of information systems. If access to mailing lists or the distribution of mailing lists is restricted, workers lose opportunities to request or share help. Similarly, lack of organizational incentives for information sharing may reduce contributions to "public" data resources (Orlikowski, 1992; Markus & Connolly, 1990). As a result, potential content for message archives may be diminished. Another important direction for future research, then, is an analysis of how organizational information handling policies interact with information systems to shape what information is retained by organizations.

In the years since Tandem employees began using archives of technical message exchanges, many useful technical archives have appeared in other settings. Many of them are not located on corporate intranets but rather are found on the Internet. The open source movement supports many of these archives which are used by programmers and developers seeking to understand and extend open source code or features. These archives more resemble Tandem's peer archive than the expert archive in that most questions have multiple replies from volunteers and there are no organizationally-designated formal experts, although some people are acknowledged as more expert than others by the community (Moon and Sproull, this volume).

Within organizations the "knowledge management" community has invested substantial resources in building "knowledge repositories," archives of expertise that employees can draw upon when they have technical questions (e.g., Hansen, Nohria and Tierney, 1999; Henderson, xx). These typically do not reflect the question/answer(s) structure of the Tandem or open-source archives. They do, though, resemble Tandem's expert archive in that "answers" are typically provided by or validated by a designated expert rather than coming from multiple volunteers. A big challenge for "knowledge managers" is getting people to use their repositories. The evidence suggests that employees still prefer asking for help from people, especially when the knowledge to be transferred is complex (Hansen, 1999). To the best of our knowledge, none of these endeavors also provides a repository analogous to the Tandem peer archives to give employees the direct choice of accessing know-what organized by experts or know-how voluntarily offered by peers.

An interesting hybrid is the Xerox Eureka system , which is a repository of "tips," which describe how machine service problems have been solved in the field. Tips do not come from designated experts; nor do they report official engineering documentation. They are voluntarily submitted by field service technicians who describe how they actually solved thorny problems in the field (rather than how the official documentation said they should be solved). Whereas it does not display the question/answers structure of the Tandem employee and open-source archives, the Eureka repository does record know-how voluntarily offered by ordinary employees. Corporate evaluations suggested that Eureka was extensively and productively used by field service personnel in an effort to improve field service distributed over time and space. The system was so heavily accessed that the official corporate product documentation group decided to download all of its official documentation into the system so that technicians would have easy access to it too. Ironically, according to a supervisor of the original (de Kleer, 2000), the official know-what was so voluminous that it slowed down the system and made it difficult for technicians to find the know-how that had been the system's original rationale. Submission of new tips declined as did use of the system by field service personnel.

We do not advocate eliminating official documentation. However, a growing body of evidence suggests that corporate officials in charge of officially-sanctioned know-what should respect and support systems that encourage peers to share know-how over space and time.

References

Ackerman, M.S. (1994). Augmenting the organizational memory: A field study of Answer Garden. Proceedings of the ACM conference on computer-supported cooperative work 1994 (pp. 243-252). New York: ACM Press.

Ackerman, M.S., & McDonald, D.W. (1996). Answer Garden 2: Merging organizational memory with collaborative help. Proceedings of the ACM conference on computer-supported cooperative work 1996 (pp. 97-105). New York: ACM Press.

Allen, T.J. (1977). Managing the flow of technology, MIT Press, Cambridge, MA.

Bell, D. G. Bobrow, D.G., Rainamn, O. & Shirley, M.H. (1997). Dynamic documents and situated processes: Building on local knowlede field service. In T. Wakayama, S. Kannapan, C. M. Khoong, S. Navathe, & J Yates, (eds.), Information and process integration in enterprises:Rethinking documents (pp 1-15). Norwell, MA: Kluwer Academic Publishers.

de Kleer, J. (2000). Informal remarks presented at the 3rd Intangibles Conference - Knowledge: Management, Measurement and Organization, May 18-19, 2000, Stern School of Business, New York University.

Brown, J.S., & Duguid, P. (1998). Organizing knowledge. California Management Review, 40, 90-111.

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

Emmer, M.B. (1985). SNOBOL4+: The SNOBOL4 language for the personal computer. Prentice-Hall, Englewood Cliffs, NJ.

Finholt, T.A. (1993). Outsiders on the inside: Sharing information through a computer archive (Doctoral dissertation, Carnegie Mellon University, 1993). Dissertation Abstracts International Section A: Humanities and Social Sciences, 54, 3912.

Fulk, J., and Boyd, B. (1991). Emerging theories of communication, Journal of Management, 17, 407-446.

Galegher, J., Sproull, L., Kiesler, S. (1998). Legitimacy, authority and community in electronic support groups. Written Communication, 15, 493-530.

Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. Administrative Science Quarterly, 44, 82-111.

Hansen, M.T., Nohria, N., Tierney, T. (1999, March-April). What's your strategy for managing knowledge? Harvard Business Review, 77, 106-116.

Henderson, J.C., Sussman, S.W., & Thomas, J.B. (1998.) Creating and exploiting knowledge for fast-cycle organizational response. In D. J. Ketchen, Jr., Turnaround research: Past accomplishments and future challenges. Advances in applied business strategy (Vol 5, pp 103-128). Greenwich, CT:JAI Press.

Hutchins, E. (1995). Cognition in the wild. Cambridge, MA: MIT Press

Kusterer, K.C. (1978). Know-how on the job: The important working knowledge of unskilled workers. Boulder, CO: Westview Press.

Latour, B. & Woolgar, S. (1986). Laboratory life: The construction of scientific facts. Princeton: Princeton University Press.

Markus, M.L. and Connolly, T. (1990). Why CSCW applications fail: Problems in the adoption of interdependent work tools. Proceedings of the Conference on Computer Supported Cooperative Work (pp. 371-380). NY: ACM Press.

Orlikowski, W.J. (1992). Learning from Notes: Organizational issues in groupware implementation. Proceedings of the ACM conference on computer supported cooperative work (pp. 362-369). New York: ACM Press.

Orr, J.E. (1989). Sharing knowledge, celebrating identity: War stories and community memory in a service culture. In D.S. Middleton & D. Edwards (Eds.). Collective remembering: Memory in society (pp. 169-189). Beveraly Hills, CA: Sage.

Orr, J.E. (1996). Talking about machines: An ethnography of a modern job. Ithaca, New York: ILR Press.

Rice, R.E., and Aydin, C. (1991). Attitudes toward new organizational technology: Network proximity as a mechanism for social information processing, Administrative Science Quarterly, 36, 219-244.

Tobin, J. (1958). Estimation of relationships for limited dependent variables, Econometrica, 26, 24-36.

von Hippel,E. (1994). Sticky information' and the locus of problem solving: Implications for innovation. Management Science, 40, 429-439.

Endnotes

```
     Document #: 3538 - Document length:  120 lines - [August,
19xx]
[program v1] to [program v2] CONVERSION
```

QUESTION

```
My customer, [university name], is planning to convert
approximately 200 application programs from [program v1] to
[program v2]. They would like to know of other customers'
experience in doing this. Any story, with details, would be
welcome.
                         Many thanks,
                         [Questioner's full name]
---------------------------------------------------------

*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*
*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*
```

```
=  COMMENTS from the "TNT" group!      August, 19xx =
=  Please note that the contents of this file have not  =
=  yet been reviewed for technical validity.       =
=                    Tools-N-Technology =
=      ( FOR INTERNAL USE ONLY)          =
*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*
*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*=*
```

REPLIES

```
FROM: [Replier1's ID]@HAMBG
Hi [Questioner's first name]!
There exists a utility to convert from [program v1] to [program
v2]. I don't know exactly where it resides but you should check
QUEST-database.
Hope that helps.
Regards
[Replier1's first name]
---------------------------------------------------------
     FROM: [Replier2's ID]@PARIS
[Questioner's first name],
On node \PRUNE you can find files containing usefull
informations for doing convertion from [program v1] to [program
v2] and how to code with [program v2]..
They are on : \PRUNE.$ [program] L. [program] L85.CONVGD
\PRUNE.$ [program].TOMC85. [program] GUIDE
[Replier2's first name]
---------------------------------------------------------
     FROM: [Replier3's ID]@HOLLAND
<Questioner's first name>,
In Holland we started two conversions: <customer's name> had no
real problems - contact <person's ID> <another customer's name>
were advised to stop the conversion, because the compiler had
too many problems, it interfered with the project contact
<person's ID>
success,
<Replier3's full name>
---------------------------------------------------------
     FROM: [Replier4's ID]@CSO
[Questioner's nickname]:
There is a good article on [program v2], written by one of the
developers ([author's name]) in the February, 19xx, Systems
Review (Volume 2, Number 1). One section of the article
highlights the differences between the two languages.
[Replier4's first name]
---------------------------------------------------------
     FROM: [Replier5's ID]@SILICON
I talked to one of my accounts who has converted to [program
v2]. and about the only change they have made from their
[program v1] programs is to replace computer statements that
used division with explicit divide statements to insure correct
precision in the results.
---------------------------------------------------------
```

```
     FROM: [Replier6's ID]@STLOUIS
[Questioner's first name],
I'm going thru my second account conversion right now. Here are
some quick notes on the experiences.
1) All of [previously referenced author]'s commentary on the
conversion process is valuable and should be reviewed with the
customer. I believe the location of this info has already been
supplied to you.
2) There has never been a requirement to use the reserved word
conversion program at my accounts. It's comforting to know its
available but I suspect that most customers will not run afoul
of the new reserved words.
3) [program v2] is much more sensitive to numeric field
manipulation. The standard is very specific and rigidly adhered
to in the Tandem implementation. This means that many programs
that ran under [program v1] will trap under [program v2],
usually with numeric overflow. Example: My customer had a
program that moved PIC 9(5) to PIC 9(4). Obviously this is bad
practice, but it happens frequently. Under [program v1] the
program continued execution and the invalid results of the move
were later discarded by other program logic. Under [program v2]
a trap #02 reslted. Customers also report occasional divide-by-
zero traps.
4) The Standard specifically states that any key-field
referenced in a program be defined alphanumerically. Consider
the following example: : : SELECT  A-file   ASSIGN TO <disc-
file> ORGANIZATION INDEXED RECORD KEY IS A-File-Key : : FD A-
file    01 A-Record. 02 A-File-Key. 03 Acct-No    PIC 9(6). 03
Jrnl-No PIC 9(6). : : START A-File KEY = Acct-No, GENERIC ....

Under [program v1] this compiles and executes as expected. Under
[program v2] the 'START' statement is rejected by the compiler.
The Standard requires that Acct-No be alphanumeric. According to
at least 1 TPR I've read the compiler will relax this
restriction in a future release, probably C00. You should check
this situation as it may already be dealt with by the recent
EBF.
Hoping this helps ... [Replier6's full name]
```

Figure 15-1. Question and replies in the Peer archive.

```
     Document #:3 - Document length: 36 lines -26 SEP xx
TRANSFER OF ACU UNITS TO THE SAGE
```

QUESTION

```
I have received a question wondering if asynchronous ACU units
which now use 2 async patch panel ports can be transferred to
the [hardware system]. These are implemented with [communication
system] on the async board, and use a special 'Y' cable. Will
the [hardware system] support this?
```

[Questioner’s full name]\world.support [Questioner’s first name]

REPLY

Following comments by \CASG.CASG.[Expert’s ID] on 27 SEP 19xx, 16:21:05

THERE IS NOT, AT PRESENT, ANY SUPPORT FOR ACU’s ON THE [hardware system] (A NEW LIM WITH THE RS-366 INTERFACE) HAS BEEN DESIGNED AND BUILT BY ENGINEERING, BUT IT HAS NOT BEEN PUT INTO PRODUCTION, AND NO SOFTWARE HAS EVER BEEN WRITTEN TO MAKE USE OF IT, NOR IS ANY SCHEDULED.

AUTO-ANSWER IS SUPPORTED BY SOME TYPES OF LINES ON THE [hardware system], E.G. <[line type], BUT NOT ALL.

ONE POSSIBILITY FOR YOUR AUTO-DIAL CAPABILITY MIGHT BE TO USE A “SMART MODEM” WHICH GETS ITS DIAL DIGITS FROM THE DATA INTERFACE. THIS WOULD REQUIRE SOME PROGRAMMING EFFORT, AND MIGHT BE LIMITED TO [line type] AND POSSIBLY [line type]. TO MY KNOWLEDGE NO ONE HAS DONE THIS TO DATE, THEREFORE I CANNOT GUARANTEE SUCCESS.

THERE IS NO [communication system] EQUIVALENT SUPPORT USING [line type] WITH ASYNC POINT TO POINT LINES. THEREFORE, A PROGRAM WRITTEN WITH [communication system] USING 2 ASYNC PORTS WILL NOT WORK AND COULD NOT EVEN BE CONVERTED TO CP6100.
[Expert’s initials]

Figure 15-2. Question and reply in the Expert archive.

Table 15-1. Archive Attributes and Use Over 45 Weeks.

	Archive Type	
Archive Attributes	Peer (n = 428)	Expert (n = 73)
Mean replies per question	7.17(7.02)***	1.05(.23)
Mean unique repliers per question	6.57(6.19***	1.05(.23)
Mean unique replier sites per question	5.81(4.84) ***	1.04(.20)

Archive Use		
Number of unique users	1708	638
Number of archive logins	8450	1537
Number unique sites from which archives were accessed (n = 151)	108	96

***p < .01.

Table 15-2. Mean Occurrences (per 1000 words) of Keywords Used in Content Analysis of Questions And Replies In Archives.

	Questions		Replies	
Content	Peer (n = 465)	Expert (n = 78)	Peer (n = 465)	Expert (n = 78)
Personal experience (# first person pronouns, e.g., I, me)	14.9(17.0)	14.5(18.2)	14.9(10.7)***	6.5(9.3)
Informal language (# contractions, e.g., can't)	3.5(7.6)	3.2(12.3)	5.1(5.2)***	1.1(3.7)
Pointers (# pointing verbs, e.g., call, contact)	0.7(4.2)***	0.2(1.0)	2.8(4.6)***	1.0(2.8)
References to Tandem products	10.9(14.3)***	2.4(10.6)	4.7(6.3)***	1.9(4.8)
References to customers (# "customer [s]")	6.7(10.6)***	3.4(6.9)	2.5(3.4)***	1.2(3.3)

Table 15-3. Descriptive Statistics, Network Sites, n = 141.

Variable	Statistics			
	M	SD	Min	Max
Distance (miles from engineering headquarters)	2636.9	2328.7	0.0	9182.4
Seniority (median tenure in years)	3.9	1.4	0.8	7.7
Size (number of employees)	63.2	89.2	1.0	553.0
Registered computer accounts (percent of employees)	93%	19%	0%	100%
Logins per employee, peer archive, 45 weeks (1988)	1.6	1.8	0.0	9.0
Logins per employee, expert, 45 weeks (1988)	0.4	0.3	0.0	1.7

Table 15-4. Regression Coefficients (Tobit Analysis) Predicting Mean Archive Logins per Employee, n = 141.

Variables	Control	Distance Only	Distance and Archive Model	Moderated Model
Intercept	-.66	-1.74	-2.32**	-2.01
Registered computer accounts[a]	.13	.15	.14	.14
Seniority	-.33***	-.22**	-.22***	-.22***
Size[b]	.02	.04	.03	.03
Sales function	1.44***	1.25***	1.21***	1.18***
Distance[c]		.16***	.16***	.06
Archive type			1.58***	1.08***
Distance X Archive type[d]				.18**
Pseudo R^2	.11***	.14***	.28***	.29**
G of improvement		9.2**	63.1***	5.2***

Footnotes to Table 15-4.

[a] Registered computer accounts expressed in units of 10^1 percent for interpretative clarity

[b] Size measured in units of $x^{0.5}$ after recommendation of Weisberg (1985)

[c] Distance expressed in units of 10^3 miles for interpretative clarity

[d] Distance X Archive expressed in units of 10^3 miles for interpretative clarity

*** = $p < .01$, ** = $p < .05$, * = $p < .10$

[1] This chapter is based on a doctoral dissertation by the first author (Finholt, 1993). The research was supported by a National Science Foundation Graduate Research Fellowship, the System Development Foundation, the Xerox Palo Alto Research Center, and Bellcore. We gratefully acknowledge the cooperation and help of Jimmy Treybig and the rest of the Tandem organization in providing access to their corporate network and archives. In addition, we thank Michael Cohen, Baruch Fischhoff, Mark Kamlet, Gary Olson, Michael O'Leary, Stephanie Teasley, Steve Whittaker, Susan Fussell, and Pamela Hinds for comments on earlier drafts.