

ORGANIC KNOWLEDGE MANAGEMENT FOR WEB-BASED CUSTOMER SERVICE

Stephen D. Durbin, Doug Warner, J. Neal Richter, and Zuzana Gedeon
RightNow Technologies

ABSTRACT

This chapter introduces practical issues of information navigation and organizational knowledge management involved in delivering customer service via the Internet. An adaptive, organic approach is presented that addresses these issues. This approach relies on both a system architecture that embodies effective knowledge processes, and a knowledge base that is supplemented with meta-information acquired automatically through various data mining and artificial intelligence techniques. An application implementing this approach, RightNow eService Center, and the algorithms supporting it are described. Case studies of the use of eService Center by commercial, governmental and other types of organizations are presented and discussed. It is suggested that the organic approach is effective in a variety of information-providing settings beyond conventional customer service.

INTRODUCTION

The phrase "organizational data mining" in the title of this book suggests the importance of tapping all sources of information within an organization. The bare term "data mining" is most often applied to the extraction of patterns and relationships from databases or other structured data stores, enabling the productive use of information otherwise buried in overwhelming quantities of raw data. More recently, methods have been developed to extract information from relatively unstructured text documents, or at least to render that information more available via techniques of information retrieval, categorization, and extraction. But in spite of such progress, one major source of organizational knowledge often remains inadequately managed.

It is widely recognized that much of the knowledge of any organization resides in its people. A major difficulty in tapping this key resource is that much of this knowledge is not "explicit," but rather "tacit." For our present purposes, we call "explicit" the sort of knowledge which could be captured relatively easily in a document such as a memorandum, a manual or a white paper. In contrast, "tacit" knowledge is generally not committed to any permanent, structured form, because it tends to be strongly dependent on context or other variables that cannot be described easily. Because of its difficult nature, as well as its importance, the concept of tacit knowledge has received much attention in the recent literature (see, e.g. Nonaka & Takeuchi, 1995; Stenmark, 2000; Richards & Busch, 2000), though its roots

go back at least to Polanyi (1966). It has become clear that the obstacles to capturing such knowledge are not merely technical, but psychological, sociological, and even philosophical. No simple solution can be anticipated to this inherently difficult problem. Nevertheless, one can hope to identify certain features of the problem that are likely to be important in designing systems to deal with it.

In the following, we shall present our view of some key aspects of human-centered knowledge acquisition and dissemination. We do this within the context of a specific software application, RightNow eService Center (RNeSC), which was originally developed and is primarily used for Web-based customer service. This is not the limited domain it might at first appear, for the basic paradigm of knowledge exchange between producers (e.g. customer service representatives, university staff, or government agencies) and consumers (e.g. customers, students, or citizens) can be applied very generally. To cover this broad spectrum using a common terminology, we shall refer to the producers as “experts” and the consumers as “novices” or “end-users,” while the general term “users” will encompass both groups.

Focusing on the knowledge management aspects of our application, the fundamental goal is to facilitate information-finding by end-users and information-providing by experts. We recognize that the information transfer, though asymmetric, occurs in both directions. Indeed, one of our main points is that for end-users to learn effectively, the experts must also learn about the end-users and their information needs. Furthermore, we note that the same basic paradigm can also apply to the situation where experts and end-users are the same population: our software is actually used in that way within a number of organizations, including our own.

Data mining is key to the function of RNeSC in more than the metaphorical sense of eliciting knowledge from experts or the conventional sense of extracting information to generate various reports on the system status, history, and use. Beyond these, the continuous analysis of text exchanges and the mining of user interaction logs represent embedded data mining functions that are crucial to the performance of RNeSC. Their main purpose is to extract what could be considered tacit knowledge of both experts and end-users about the relationships among knowledge items in the knowledge base. This meta-knowledge would be both tedious and overly demanding for users to provide directly, but greatly improves the operation of the system in terms of user experience, as we shall describe.

Our aim in this chapter is to present our approaches to knowledge acquisition and access, and show how they are implemented in the RightNow eService Center application. We outline various statistical and artificial intelligence techniques that are used in the process. Based on extensive usage information provided by companies and educational

and governmental institutions that have used RNeSC, we describe some practical aspects of deploying and using the application. Finally, we discuss future trends and draw several conclusions.

KNOWLEDGE MANAGEMENT FOR CUSTOMER SERVICE

Knowledge Management Issues

We begin with a few general observations relating to the tasks of collecting or acquiring knowledge from people and providing it to others. We make no attempt to survey the vast literature on knowledge management, but simply note that a great deal of effort has gone into analyzing the nature of knowledge in its various forms, and in particular the feasibility of capturing it for re-use or training. As mentioned, much discussion has centered around the distinction between explicit and tacit knowledge (Nonaka & Takeuchi, 1995; Stenmark, 2000; Sternberg, 1999). Though not clearly separable, these two types of knowledge are equally significant. Because tacit knowledge is often unique to an organization, it is considered a major source of competitive advantage, distinguishing that organization from others. Furthermore, tacit knowledge presents special management problems as personnel changes. For some (Polanyi, 1966; Cook and Brown, 1999), tacit knowledge is by definition that which cannot be expressed, while others (Nonaka and Takeuchi, 1995; Stenmark, 2000; Richards & Busch, 2000) consider “externalization” of tacit knowledge to be possible in appropriate settings. We believe the latter view is more appropriate to the domain of generalized customer service.

The process of conveying knowledge -- both explicit and tacit -- from expert to novice can be divided into stages, each associated with certain artifacts. This division is not unique. One traditional approach, exemplified by the creation of product documentation, models the process in two stages: 1) the expert writes the documentation, and 2) the novice reads it. While straightforward and familiar, this approach places a heavy burden on the expert to anticipate all the knowledge that could be required, and present it in a way that can serve all those who might need it. An equally heavy burden is placed on the novice, who must extract from the resulting large body of knowledge just that part corresponding to his or her need. Naturally, it often happens that what the novice needs is not fully provided, or the context is different enough that the novice fails to find the separate “nuggets” that, combined, would meet the need.

For example, it is difficult, if not impossible, to write a service manual for some piece of equipment that covers all repair situations. Yet an experienced repairperson can usually figure out what is needed on a particular job. If that specific repair procedure is described, elements of tacit knowledge are implicitly captured. In most organizations, as described in Brown and Duguid (1991), such “stories” are circulated informally within a community of practice. If

they can be recorded and made more widely available, as was done with the well-known Eureka system at Xerox, the resulting knowledge base can be of extraordinary value to the organization (Powers, 1999; Fischer and Ostwald, 2001).

A second traditional approach is simple dialog between expert and novice, in which the expert can both assist in the expression of the novice's needs and convey the knowledge in the most effective way for the particular novice. Such a model is the ideal of the conventional help desk. It typically results in the greatest benefit to the novice, but, depending on the setting, it may be highly burdensome and expensive to have an expert always available for each novice.

The model of the knowledge transfer process embodied in the architecture of eService Center comprises elements of both the traditional approaches described above. Our interactive approach starts with the novice's specific information need. This is not necessarily clearly formulated, so we provide various means for the novice to satisfy the need via a self-service knowledge base. If that effort is unsuccessful, the novice must express the need in the form of a text message, which is sent to an expert. The expert then responds, drawing on accumulated experience, including appropriate elements of tacit knowledge. The response is thus much more limited in scope and tailored to the immediate need. In this setting, tacit knowledge which might not have found its way into a manual, is "activated;" the expert realizes "intuitively" what will work best in the case at hand. This leads to what we consider an important aspect of any approach that aims to capture such knowledge: it is easiest to do so at the point of application, that is, in the consideration of a particular situation that calls for such knowledge. In a final stage of the process, the expert can choose to add the newly articulated knowledge to the knowledge base, thus enhancing self-service capability on the part of other novices.

Knowledge management in our model contrasts with that in the traditional model in several key regards. First, knowledge creation by the expert occurs not in a relative vacuum, but in a specific, situated context. This facilitates application and capture of tacit knowledge, which is stored in the knowledge base along with the context. The knowledge transfer is not limited to transmission via a static artifact such as a manual, but either by direct, personal response of expert to novice, or via the novice's ability to locate the knowledge on his own. Since the latter is preferred, it is important to provide tools that assist the novice to navigate the knowledge base. Finally, utilization of the knowledge tends to be more effective under our model because of increased relevance to the particular situation of the novice.

Note that we have so far taken the expert to be omniscient. In most real cases, the expert also might need to refer to the knowledge base in the course of responding to a novice, especially if the “expert” is really an expert-in-training.

The Customer Service Domain

Customer service represents a quintessential knowledge management problem. Answers, i.e. information or knowledge, must be identified, transcribed, or acquired by or from experts (e.g. customer service representatives) and then provided to novices (end-users) in response to their questions. Because of the economic importance of customer satisfaction, significant resources may be devoted to this function. In recent years, many companies and other entities have found it necessary to maintain a presence on the World Wide Web, and customer service is naturally one of the functions that can be provided by this means. However, the journey has not always been easy or successful.

Historically, the first step toward Web-based service was that of simply listing contact phone numbers and e-mail addresses on a Web page; end-user inquiries were then handled through these more traditional channels. This approach has the advantage of using existing infrastructure, but is typically very expensive per transaction, especially as there is now a general expectation of rapid response, even 24 hours a day. The majority of organizations are still at this level.

A second generation of Web-based service provides a set of answers to frequently asked questions (FAQs) on a support Web page. The composition of such a FAQ list is based on the accumulated experience of customer service representatives (CSRs). If well written and organized, this can significantly reduce the number of repeated inquiries received by CSRs, reducing their overload and increasing their productivity. However, unless the common inquiries are quite stable over time, this method requires a significant maintenance effort to keep the FAQ list organized and up to date. In many cases, depending on the organization, change can be relatively great on a weekly or monthly time scale. This change can also be unpredictable: although it may be easy to see that introduction of a new product will lead to inquiries related to that product, it is not so easy to foretell what external events, such as a new law or regulation, or new products offered by competing companies, will cause a shift in end-user information needs. A further problem with this type of service is that as the number of FAQs grows larger, it becomes increasingly difficult for users to navigate to find answers to their questions.

A third level of Web-based service involves the provision of search capability over a set of indexed documents that constitute the online knowledge base. With such a system, answer-containing documents can be added independently of each other, and the structure is essentially the invisible one provided by the search facility. Related documents are,

by definition, those returned together in response to a specific search query. Depending on the design of the search engine, it may or may not provide additional features such as natural language input or matching to related words such as wordform variants (drive, driver, driving, ...) or synonyms (car, automobile, ...). These well-known search engine problems have led some companies to deploy conversational question-answering systems, or “chatbots.”

Unfortunately, at their current stage of development, such systems require extensive knowledge engineering in the form of identifying input question patterns that should be recognized and the links to the corresponding answers. Beyond pre-scripted patterns the performance degrades rapidly. Furthermore, this type of system either does not support overviews and browsing of the knowledge base, or again requires knowledge engineering to create and maintain a taxonomy and relate it to the collection of knowledge base documents.

At present, there is a wide range of levels of customer service available on the Internet. Some organizations have managed a good fit between what they need and what they provide. But many are still struggling with expensive, cumbersome systems that do not serve them well. In some cases, organizations simply don't have a good understanding of what state-of-the-art customer service can or should be. But probably most often it is a lack of resources -- both human and financial -- that limits the quality of service. For this reason, a constant aim in the development of eService Center was to minimize the effort necessary to establish and maintain the system. This entailed an architectural design in accordance with the above and other considerations, as well as integration of data mining and artificial intelligence techniques to reduce the burden on users.

AN ORGANIC KNOWLEDGE BASE

The Organic Approach

To find a way to meet the sometimes conflicting needs of experts and end-users, we believe that attention must first be focused on the core of the system where knowledge is stored, namely the knowledge base. In the type of system we envision, this is a publicly available (via the World Wide Web), dynamic collection of documents that we shall refer to as Answers. We assume here that it is created and maintained by the experts (e.g. CSRs); in actuality, there may be distinct managers who perform various important functions, but whose role is outside the scope of this paper.

The knowledge base must first of all contain the knowledge sought by end-users. How does one know what this is? Within our organic approach, the reply is simple: let the end-users identify what is needed by the questions they

ask. This implies that the knowledge base not exist in isolation; it must be closely coupled to channels through which end-users ask questions.

As a concrete example, take the case of a computer services help desk that receives several complaints about problems with a new software version. Though one might expect certain problems to arise with such upgrades, it would be extremely difficult to attempt to forestall complaints by preparing a comprehensive troubleshooting guide. In contrast, it is relatively easy for a technician to answer a specific question in a given context, and easier for the end-user to understand that answer than the full troubleshooting procedure. If the Answer is published in the knowledge base, incoming questions on the topic may be reduced, either because only that single case led to problems, or because other users could resolve their problems also by following the approximate answer. If not, new support requests will come in, and then either the first answer could be modified or expanded, or a new answer added. This adaptive, “just in time” approach is very efficient in terms of experts’ effort.

Thus, along with the knowledge base itself, end-users must have direct access to experts when they don’t find the information they need in the knowledge base. It is their needs which drive knowledge creation, while the experts’ effort is conserved. A similar concept of a experts backing up a knowledge base in a system for organizational memory, an application close in spirit to customer service, has been described and studied by Ackerman (1998).

According to the organic growth scenario, the knowledge base is initially seeded with a fairly small set of Answers to known or anticipated FAQs. After that, Answers are added as needed to respond to newly incoming questions frequent enough to merit creation of a public Answer. (Of course, Answers to predictable concerns can also be added even before questions arise.) Depending on the organization, the threshold could range from one to perhaps hundreds.

This approach has a number of advantages for experts:

- tacit as well as explicit knowledge can be brought to bear on the specific questions;
- Answers can be based on existing private responses (made before reaching the threshold);
- no time and effort are spent creating unneeded Answers.

Furthermore, experts have a natural motivation to upgrade private responses to publicly available Answers: a single Answer can eliminate the need for many private responses. As in any knowledge management endeavor, it is also important for the organization to create a culture in which such contributions are recognized or rewarded in one way or another.

End-users also benefit. Unlike a traditional call center or contact center, this approach develops an authoritative, self-service knowledge base, accessible 24 hours a day, in which an end-user can generally locate information faster than she can compose, say, an e-mail describing her question or problem. This is especially true if the initial problem description is unclear and a series of back-and-forth communications would be necessary for clarification. End-users with novel or inherently personal questions can still receive service through the traditional channels, now improved because of the reduced load on the experts.

Critical to success of this approach is the ease with which end-users can actually navigate the knowledge base to find information. Frustration leads not only to negative attitudes towards the organization, but if anything increases the burden on experts. In contrast, a positive experience for the end-user enhances trust and loyalty, key assets for non-commercial as well as business entities. Ensuring such a positive interaction requires attention to the psychology as well as the statistics of searching. To support each end-user's quest, the interface to the knowledge base must be as intelligent as possible, and be adaptable to the range of search skills that different users may have. This entails use of natural language and artificial intelligence (AI) methods. Appropriately integrated, these techniques can also be applied to improving performance of the system from the experts' point of view.

Feedback from end-users to experts, in addition to that implicit in the asking of questions, should be such as to facilitate various forms of optimization, as well as provide understanding of end-user behavior that may be significant to the organization. This can be accomplished by mining records of user interactions with the knowledge base.

In the following, we detail how this organic approach is embodied in RightNow eService Center. After briefly introducing the overall system, we focus on those aspects related to the knowledge base, as this is where most of the artificial intelligence (AI) and data mining techniques come into play. We describe the use of the system in practice, as well as the algorithmic techniques employed and their multiple roles.

System Architecture

RightNow eService Center is an integrated application that combines e-mail management, Web self-service, live collaborative chat, and knowledge management. The core of the application, from our present perspective, is the publicly accessible Answer knowledge base and the tools by which it is created, maintained, and accessed. In addition, there is a database of customer service Incidents, i.e. messages from end-users, which are fully tracked from initial

creation--via e-mail, Web form, or live chat--through resolution and archiving. Figure 1 illustrates these key components and how they are involved in end-user and CSR (or other expert) knowledge-related transactions.

As indicated in the previous section and in Figure 1, the architecture of RNeSC provides a strong interaction between question and answer channels. As CSRs respond to the questions submitted by end-users, they naturally become aware of trends and commonalities among them. At any time, a private CSR reply can be proposed as a potential public knowledge base item, or a new Answer can be composed on the basis of previous replies or predicted information needs. Depending on organizational practices, the item might be reviewed or edited by collaborators or managers before being made a publicly available Answer.

In typical operation, the main knowledge flow (in terms of volume) is from the knowledge base to end-users who are successful in their searching. But even if they are unsuccessful, or if they make no attempt at self-service, the contents of their question may suggest that one or more relevant Answers actually exist. In that case, Answers can be suggested automatically, based on search technology described later. These Suggested Answers can either be routed directly to the end-user as an auto-reply, or to the CSR engaged in formulating a personal reply. Naturally, CSRs also make direct use of the knowledge base for their own information, especially novice CSRs.

In this paper we leave aside the multiple administrative functions of RNeSC, though these are vital to its overall ease of use (especially from a CSR's point of view). Some of these use AI techniques also employed in the central knowledge management functions. For example, one of the criteria that can be used in routing incoming questions to individual CSRs is an emotive index that estimates the degree to which the tone of a message is angry, neutral, or happy. This determination uses the same natural language processing algorithms described later, in combination with wordlists and grammar rules. As eService Center is available in about 15 languages and dialects, implementing this feature takes a significant effort. Also not indicated in Figure 1 are a module that generates a wide variety of reports to aid in evaluating transaction statistics, CSR performance, and Web site usage. These are developed through both batch and incremental analysis of system interaction records. Finally, except to mention a notification function that allows a user to be informed of any changes in a selected Answer, we won't detail here the many customization options that users can set.

Using the Knowledge Base

It is widely appreciated that knowledge comprises not only facts, but relationships among these, as well as perspective on their importance, relevance, etc. A knowledge base organized to incorporate or reflect such meta-knowledge provides a much better match to user habits and expectations, and is consequently easier to use. In RNeSC this meta-knowledge is acquired through several techniques. In addition to intelligent searching, these include adaptive clustering and classification of text documents (the knowledge base Answers), and collaborative filtering techniques that mine usage patterns to extract implicit user feedback on importance, timeliness, and relatedness of knowledge base items. We will describe these techniques as they might come into play while a user navigates a knowledge base.

An illustration of a simple end-user view of a knowledge base is shown in Figure 2. This page is reached after first selecting the “Answers” link on the support home page, before any search has been made. The Answers shown are listed in order of historical usefulness--called solved count--which measures how helpful an answer is likely to be based on the experience of previous users. If the knowledge base is not too large and the end-user is looking for information that is commonly sought, there is a fair probability that the appropriate Answer will be listed in the first set. If the solved count of answers happens to follow a Zipf distribution, then even with 500 items in the knowledge base, there is nearly a 50% chance that the appropriate Answer will be within the top ten.

The solved count is obtained from a combination of explicit and implicit user feedback. If enabled, each Answer page carries evaluation buttons (e.g. labeled 0%, 25%, 50%, 75%, 100%) that the user can select to indicate the degree to which his or her question was answered; these contribute proportionately to the solved count. Since relatively few users make the effort to provide explicit feedback, we also derive an implicit evaluation from the user’s actions. Simply choosing to view a particular Answer is taken as a partial vote for its usefulness. If the Answer is the last one viewed, it is assumed that it provided the information sought, and the vote is given a higher weight (though still less than an explicitly approved Answer).

An Answer that appeared promising from its title might prove insufficient. If so, to the extent the title represented the content, an Answer with similar or related content might help the user. Each Answer page can be provided with links to a variable number of the most closely related Answers. The relatedness ranking, like the solved count, has explicit and implicit components. The explicit relatedness is derived from text similarity, currently based on the vector model common in information retrieval (see e.g. Manning & Schütze, 1999, p. 296), with stopword removal and conflation of words having the same stem. To obtain an implicit relatedness score, the application maintains a link matrix, the corresponding element of which is incremented each time an end-user navigates from one Answer to

another, presumably related one. The increment is larger if the second Answer is the final one viewed or is given a high explicit rating.

The methods just mentioned for capturing user perceptions of usefulness and relatedness are inspired by both collaborative filtering (Levy & Weld, 2000) and swarm intelligence (Dorigo, Di Caro & Gambardella, 1999) approaches. In our application, rather than software agents traversing a network as in the usual form of swarm intelligence, it is human users whose paths leave a trace as a pheromone-like record. The resulting link matrix certainly contains noise in the sense that not every item-to-item transition is made by users only on the basis of perceived relatedness. Nonetheless, averaged over many users who each tend to be searching for information related to a specific need, we have found that the strong links indicate useful relationships. By means of the accumulated links, the application learns which other items in the knowledge base are most closely related to a given one.

The algorithm as described so far would be appropriate for a static knowledge base, but not for a changing one. Just as an insect pheromone trail evaporates with time, so we perform an “aging” process by which both solved count and link values are periodically reduced in strength when not reinforced. This aging keeps the knowledge base responsive by enforcing the primacy of recent usage patterns. By this means the most useful Answers “float” to the top of the list and appear on the very first page. For a fuller discussion of these collaborative and swarm intelligence methods, see Warner et al (2001).

Both the solved count and the link matrix represent a form of knowledge acquired from users about items in the knowledge base. From a knowledge management point of view, the role of this meta-knowledge is to aid in the principal knowledge transfer by making it easier for end-users to find the Answers they need.

To find Answers to less frequently asked questions, end-users may need to perform a search of the knowledge base. Intelligent search is a prerequisite for easy access to information. In RNeSC, queries entered into the search box (Figure 2) can be processed according to a variety of search modes, including natural language input and similar phrase searching (which carries out spelling correction and synonym expansion). Searches can also be restricted to pre-defined products and categories, if such taxonomies have been established. The results of a search can be displayed in order of relevance or solved count.

End-users may or may not come to a support Web site with specific questions, but in either case they may find it convenient to browse the knowledge base from a more distant perspective, gaining an overview of the available information. Our system offers a browse mode of access in which categories of documents are displayed as folders,

labeled with the key terms most descriptive of their contents (see Figure 3). Clicking on a folder opens it to display documents and sub-folders corresponding to more specific categories. The automatically determined labels on the folders give a summary of the contents. Because the user can navigate by selecting subfolders and individual documents without needing to type search terms, the browse mode is especially helpful when the user is unfamiliar with the terminology used in the Answers, and hence might have difficulty formulating a productive search query. If desired, it is also possible to search within a browse category. In a sense, the ease of browsing is related to the tacit knowledge of a user about the subject area. Most people are able to recognize what they're looking for much more easily than they can articulate it.

The browse function is made possible by a hierarchical categorization of the text items in the knowledge base. For this we employ a modification of the fast, hierarchical clustering algorithm BIRCH (Zhang, Ramakrishnan, & Livny, 1996), the result of which is used to learn RIPPER-style classification rules (Cohen, 1995). The final topic hierarchy is determined by classifying all knowledge base items according to the learned rules. Because of the inherent multiplicity and subjectivity of similarity relationships, we allow single items to be classified in multiple places where they fit well. This makes using the browse interface much more convenient, as the user can locate an item along various paths, and does not have to guess what rigid classification might control the listing. New Answers are, on creation, simply inserted into the hierarchy according to the classification rules. After a predetermined amount of change in the knowledge base, due to either modification or addition, a re-clustering is performed so that the browse hierarchy reflects the current state of the contents, rather than a fixed hierarchy.

The features on the basis of which the clustering is performed are obtained from the document texts by shallow parsing. The natural language processing starts with part of speech tagging via a transformation-based tagger (Brill, 1994). Rules are used to identify noun phrases, which receive the highest weight as features, though selected other words are also used. In addition, customer-supplied keywords and product or category names provide highly weighted features. The weights of feature words are additionally adjusted on the basis of the frequency with which users have searched for them, as reflected in a table maintained with the knowledge base. The clustering procedure is actually carried out several times with different sets of parameters, and the best clustering according to a composite figure of merit is chosen.

To assist CSRs in composing responses, as well as to optionally supply automated responses to end-users submitting questions, RNeSC can be configured to automatically suggest Answers. This is done by first processing the

text of the question as if it were a search query. Simply taking the top-ranked Answers returned can result in spurious matches. Hence, they are filtered by checking whether they would appear in the same cluster as would the question text, now treated as an Answer for categorization. If this feature is used by a CSR, the suggested Answers are directly pasted into a reply form, where they can be edited by the human expert.

As with the solved count and the link matrix, the clustering represents automatically generated meta-knowledge that serves to aid knowledge acquisition by end-users. To evaluate scientifically the utility of such aid would require extensive human testing, which we have not carried out. However, both our own observations and, more importantly, the experience of RNeSC users, as described in the next section, indicate that the benefits can be significant.

USER EXPERIENCE WITH eService Center

The system we describe has been used, through several versions, by a wide variety of commercial, educational, and governmental organizations. Drawing from their accumulated experience, we present both aggregate statistics and case studies illustrating the dramatic reduction of time and effort for knowledge base creation and maintenance, and the increase in satisfaction of knowledge base users. This holds across the spectrum of organizations and applications, including those outside the area of conventional customer service.

Different organizations use the system in a variety of ways. The Rotherham, England Metropolitan Borough Council uses it as a community clearinghouse where answers are provided to all kinds of questions one might contact a city office about. As of this writing, it contains 476 answers to questions ranging from regularly recurring ones such as “Can I report a pothole in the road?” to more timely ones such as “Do you have any information regarding the Queen’s Golden Jubilee?” Statements by the Council make it clear that they view this information service for citizens, part of an eGovernment initiative, as very analogous to a business’s support for customers. Although the majority of the 16,000 daily hits on the site are from the UK, there are also high numbers from the US, Taiwan, Germany, France, Sweden and Denmark, some of which, it is hoped, may represent people looking to invest in the UK and attracted by Rotherham’s assets.

Within our own company (RightNow Technologies), independent instances of RNeSC are used for external customer support and for internal company information. More interesting is its use as a resource for developers, who answer each other’s questions -- a case of experts and end-users being the same population. It also provides a defect posting and tracking system shared by the development and quality assurance departments. The resulting history of

bug fixes, with each incident often carrying contributions from several developers and testers, is a heavily used company resource. In terms of knowledge management theory (see, e.g. Brown and Duguid, 2000), each bug history document constitutes a “boundary object” collaboratively produced by two groups within the organization, serving to facilitate communication between them.

Due to the high degree of automation of RNeSC, the ease of installation is such that it has been accomplished in as little as a few days or even one day. Once set up, the knowledge base can grow rapidly. For example, the United States Social Security Administration started with 284 items in their initial knowledge base, and over 200 new items based on user-submitted questions were added within two weeks. After two years, the number has stabilized at about 600, though the composition continues to change. Due to the public availability of the knowledge base, the number of telephone calls has dropped by 50%, from 50,000 to 25,000 daily. Similar experiences are common.¹

The ability of a Web self-service system to handle dynamic fluctuations in usage can be very important. As one example, an announcement of a rate hike by the U.S. Postal Service led to a short term increase in visitors to the support site of Pitney-Bowes, which provides mailing services, of nearly 1000% over that for the previous rate hike. Attempting to handle such volume via telephone or e-mail would have resulted in huge backlogs.

One quantitative measure of end-user success in finding information is the self-service index, defined as the percentage of end-users who are able to find Answers online, rather than sending a message to a CSR. Table 1 is excerpted from a Doculabs study (Watson et al, 2001) in which it was found that the self-service index for organizations using RNeSC ranged from 75% to almost 99%, averaging 85-90%. The lower values for some categories of organization, such as telecommunications or travel services companies, may be due to a greater number of end-user-specific questions in these areas. Nonetheless, given typical costs of \$30 per telephone transaction, \$10 per e-mail exchange, and \$1 per Web interaction, such high self-service rates can lead to dramatic savings. According to anecdotal reports from users, the benefits described are largely attributable to the features of RNeSC described in this paper.

DISCUSSION AND FUTURE TRENDS

We believe that the performance of the RNeSC application in a range of settings is evidence that the underlying principles have a sound practical basis. Nevertheless, there is certainly room to do better. Some improvements are incremental, such as making the clustering algorithm more adaptive to knowledge bases that may differ significantly in

the nature and length of the documents they contain, and in the granularity of the product and category divisions they use, if any. More difficult is the issue of descriptive labels for the clusters; the area of multi-document summarization is one of active current research (see e.g. Mani & Maybury, 1999).

More qualitative enhancements can be obtained from applying AI techniques to a greater number of functions. Advanced machine learning techniques can potentially be employed wherever rules are used, including incident routing, text categorization, and natural language processing. In the latter area, sophisticated question-answering systems will probably soon reach the point of being commercially viable, at least within restricted subjects. A fluent conversational interface to a knowledge base would fulfill many developers' dreams. Until that is available, the art is to provide some approximation with capabilities that outweigh the disappointments.

Another trend is toward greater personalization of user interfaces. Care must be exercised to ensure such customization facilitates rather than constrains. The extent to which significant personalization is feasible for frequent and for one-time users remains to be investigated.

Along other lines, the pursuit of applications in different sectors of knowledge management could suggest a new mix of features. RNeSC is already quite flexible and user-configurable, and could evolve in many different directions. We believe that many of its advantages as a customer service application could be realized in related areas as well.

CONCLUSION

We have presented an organic approach to knowledge creation and delivery that emphasizes rapid response for dynamic information environments. The user-driven architecture helps mobilize tacit knowledge and dramatically reduces the time and expense of creating a knowledge base. Facilitated and cooperative creation of knowledge base documents takes place as an extension of the normal activities of experts. Continuous mining of implicit end-user recognition of the importance and relationships of information items enables the system to adapt quickly, while remaining easy to use through automated re-organization. As embodied in the Web-based customer service application RightNow eService Center, the system uses a number of AI techniques to facilitate construction, maintenance, and navigation of a knowledge base of answers to frequently asked questions. These techniques include collaborative filtering, swarm intelligence, natural language processing, text clustering, and classification rule learning. Many of these individual techniques have been similarly employed in other commercial applications, but we know of no other

system that combines all of them. Customers using eService Center report dramatic decreases in support costs and increases in customer satisfaction due to the ease of use provided by the “self-learning” features of the knowledge base.

We have argued that the principles and methods of our approach are also applicable in other settings, for example, government agencies reaching out to concerned citizens. In fact, organizations and associated constituencies with information needs are ubiquitous in modern society. Ubiquitous also is the need for software tools to assist them. “Since it is the value added by people--context, experience, and interpretation--that transforms data and information into knowledge, it is the ability to capture and manage those human additions that make information technologies particularly suited to dealing with knowledge.” (Davenport & Prusak, p. 129).

NOTES

1. See further case studies at <http://www.rightnow.com/resource/casestudies.php>.

REFERENCES

Ackerman, M. S. (1998). Augmenting organizational memory: a field study of Answer Garden. ACM Transactions on Information Systems, 16 (3), 203-224.

Brill, E. (1994). Some advances in transformation-based part of speech tagging. In Proceedings of the twelfth national conference on artificial intelligence. AAAI Press.

Brown, J. S., & Duguid, P. (1991). Organizational learning and communities-of-practice: toward a unified view of working, learning, and innovation. Organization Science, 2, 40-57.

Brown, J. S., & Duguid, P. (2000). The Social Life of Information. Boston: Harvard Business School Press.

Cohen, W. H. (1995). Fast effective rule induction. In Machine learning: proceedings of the twelfth international conference. Lake Tahoe, California: Morgan Kaufmann.

Cook, D., & Brown, J. S. (1999). Bridging epistemologies: the generative dance between organizational knowledge and organizational knowing. Organization Science, 10, 381-400.

Davenport, T. H., & Prusak, L. (1998). Working Knowledge. Boston: Harvard Business School Press.

Dorigo, M., Di Caro, G. & Gambardella, L. M. (1999). Ant algorithms for discrete optimization. Artificial Life, 5 (2), 137-172.

Fischer, G., & Ostwald, J. (2001). Knowledge management: problems, promises, realities, and challenges. IEEE Intelligent Systems, 16, 60-72.

Levy, A. Y., & Weld, D. S. (2000). Intelligent Internet systems. Artificial Intelligence, 118, 1-14.

Mani, I., & Maybury, M. T. (Eds.). (1999). Advances in Automatic Text Summarization. Cambridge: MIT Press.

Manning, C. D., & Schutze, H. (1999). Foundations of Natural Language Processing. Cambridge: MIT Press.

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

Polanyi, M. (1966) The tacit dimension. London: Routedledge & Kegan Paul.

Powers, V. J. (1999). Xerox creates a knowledge-sharing culture through grassroots efforts. Knowledge Management in Practice (18), 1-4. American Productivity & Quality Center.

Richards, D., & Busch, P. (2000). Measuring, formalising, and modeling tacit knowledge. In Proceedings of the ICSC Symposium on Intelligent Systems and Applications. NAISO Academic Press.

Stenmark, D. (2000). Leveraging tacit organizational knowledge. Journal of Management Information Systems, 17, 9-24.

Sternberg, R. J. (1999). What do we know about tacit knowledge? Making the tacit become explicit. In R. J. Sternberg and J. A. Horvath (Eds.), Tacit knowledge in professional practice (pp. 231-6). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Warner, D., Richter J. N., Durbin, S. D., & Banerjee, B. (2001). Mining user session data to facilitate user interaction with a customer service knowledge base in RightNow Web. In Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining (pp. 467-72). New York: Association for Computing Machinery.

Watson, J., Donnelly, G., & Shehab, J. (2001). The Self-Service Index Report: Why Web-Based Self-Service is the ROI Sweet-Spot of CRM. <http://www.doculabs.com>

Zhang, T., Ramakrishnan, R. & Livny, M. (1996). BIRCH: an efficient data clustering method for very large databases. In Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, pp. 103-114. New York: Association for Computing Machinery.

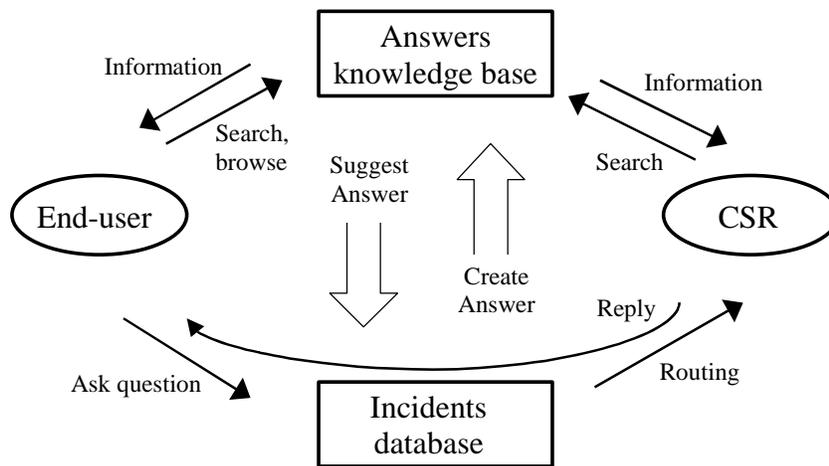


Figure 1. Principal knowledge-related transactions in RightNow eService Center. End-users search the Answer knowledge base for information; if they cannot find what they need, they submit a question, which is stored and tracked in an Incidents database, and replied to by a CSR. CSRs also use the knowledge base, and add to it by creating

new Answers, typically suggested by frequently asked questions. Answers to questions can be suggested from the knowledge base either to assist CSRs in forming replies or as auto-replies to end-users. See text for a fuller description.

USF University of South Florida
Information Technologies
 A Division of Budgets, Human Resources and Information Technology
 BHRIT Divisions > | BPA | HR | IT | ODT | VAP | VP OFFICE | CONTACT IT | STAFF

Support Home | **Answers** | Ask Questions/Report a Problem | My Stuff | Login | Help

Search [] Browse []

Category: Info Tech. []
 All Categories []

Search Text (optional) [] Search []

Search by: Phrases [] Sort by: Default Sort []

Powered by Rightflow

86 Answers Found Page: 1 of 5 Go []

Select Category	Subject	Solved Count
1 WWW Related	Updated - How do I install Verizon (formerly GTE) White Pages?	420
2 Telecom - Voicemail	Setup new student voicemail account	170
3 Network Issues - Gen	Banner/Oracle connection over RoadRunner	164
4 OASIS/Banner - Main	Minimum PC requirements for student OASIS	157
5 Outlook/E-Mail	Deleted E-Mail needs to be recovered	149
6 Virus Activity	W32/Nimda@mm virus	86
7 OASIS/Banner - Main	Student OASIS on-line evaluations	83
8 Telecom - Features/Q	Phone headsets	81
9 Telecom - Problems	How do students report phone jack problems?	80
10 Listserv	Creating a new Listserv List	75
11 Dial-Up Connectivity	Establishing Dial-up access to USF	60
12 multiple	Updated - Obtaining a new login/e-mail account	56
13 Listserv	Unsubscribe for USFTalk listserv	51
14 GEMS - Appt. Status	Unable to Generate, Unable to initiate process, FMHI	48

Figure 2. Portion of the Web browser display from the support page of the University of South Florida Information Technologies division. The page is configured to list by default the historically most useful Answer s (i.e. highest solved count). Users may search in various modes by entering search text, and they may contact a CSR via the “Ask a Question” tab.

USF University of South Florida
Information Technologies
 A Division of Budgets, Human Resources and Information Technology

BHRIT Divisions > | BPA | HR | IT | ODT | VAP | VP OFFICE | CONTACT IT | STAFF

Support Home | **Answers** | Ask Questions/Report a Problem | My Stuff | Login | Help

Search [] Browse []

Search Text (optional): [] Search []

Groupings	Answers
<ul style="list-style-type: none"> Listserv, Telecom - Changes/Requests, OASIS/Banner - Web, GEMS - On-Line Certs (Payroll) (82) Telecom - Features/Questions, code, Outlook/E-Mail, Software - Microsoft Office (37) Outlook/E-Mail, Outlook, emails, Telecom - Features/Questions, Software - Internet Explorer, ABS-FIN (22) Software - Microsoft Office, Word, GEMS - Appt. Status Form, Software - General, Telecom - Problems (11) Windows (ALL VERS), script, SuperServer - Setup (15) PPRD, thor, trusty (1) 	<ul style="list-style-type: none"> Programming buttons on a 8102 model telephone Score: 9 Long Distance Authorization Codes Score: 6 Pone Authorization Codes Score: 3 How to create broadcast distribution list Score: 0

Figure 3. Web browser display from the support page of the University of South Florida Information Technology division. This page displays a hierarchical set of folders and subfolders, where a given folder (like a typical computer file system) may contain both subfolders and Answer documents.

Industry	Visits	Escalations	Self-Service Index
General Equipment	342,728	4,144	98.79%
Manufacturing	22,784	489	97.85%
Education	8,400	317	96.23%
Entertainment/Media	113,047	4,622	95.91%
Financial Services	40,574	1,972	95.14%
Contract Manufacturers	77,838	4,203	94.60%
Utility/Energy	19,035	1,122	94.11%
ISP/Hosting	147,671	8,771	94.06%
IT Solution Providers	53,804	3,277	93.91%
Computer Software	449,402	27,412	93.90%
Dot Coms	267,346	20,309	92.40%
Medical Products/Resources	17,892	1,451	91.89%
Professional Services	24,862	2,142	91.38%
Insurance	40,921	3,537	91.36%
Automotive	3,801	373	90.19%
Retail/Catalog	44,145	6,150	86.07%
Consumer Products	1,044,199	162,219	84.46%
Computer Hardware	101,209	15,759	84.43%
Government	108,955	17,347	84.08%
Travel/Hospitality	27,099	4,610	82.99%
Association/Nonprofit	14,620	2,772	81.04%
Telecommunications	809,320	202,158	75.02%
Overall Total	3,779,652	495,156	86.90%

Table 1. Self-service index for various types of organizations using RightNow eService Center. The self-service index is the fraction of end-users that find needed information in the Answer knowledge base, rather than initiating contact (escalating) with a support person via e-mail or online chat.