

Information Self-Service with a Knowledge Base That Learns

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

■ Delivering effective customer service over the internet requires attention to many aspects of knowledge management if it is to be both satisfying for customers and economical for the company or other organization. In RightNow eSERVICE CENTER, such management is built into the architecture and supported by automatically gathering meta-information about the documents held in the core knowledge base. A variety of AI techniques are used to facilitate the construction, maintenance, and navigation of the knowledge base. These techniques include collaborative filtering, swarm intelligence, fuzzy logic, natural language processing, text clustering, and classification rule learning. Customers using eSERVICE CENTER report dramatic decreases in support costs and increases in customer satisfaction because of the ease of use provided by the self-learning features of the knowledge base.

Many companies small and large, as well as educational, government, and other types of noncorporate organizations, now find it imperative to maintain a significant presence on the World Wide Web. One of the major organizational functions that is still in the early stages of being delivered by the internet is customer service, that is, remedying complaints or providing answers to a particular audience. This task involves many aspects of knowledge management, at least if it is to be convenient and satisfying for customers as well as efficient and inexpensive for the company or organization. On a basic level, it is essential (but not sufficient) to handle the administrative overhead of tracking incoming questions and complaints, together with outgoing responses, over different channels such as e-mail,

web forms, and live chat. Beyond this, to support customer service representatives (CSRs), and to assist customers seeking help at peak load times or after hours, it is necessary to provide both a knowledge base containing needed information and a convenient, intuitive means of accessing this knowledge base. Even were it not for the expense of maintaining a large staff of CSRs always available, it is found that many people prefer to find answers to their questions directly on the internet rather than take the time to compose a sufficiently detailed e-mail message or wait in a telephone queue, possibly playing tag with a CSR for days before resolving their concerns. Furthermore, CSRs can experience boredom and burnout from constantly handling similar questions; in many cases, they are not using their skills most efficiently.

The most common and straightforward response to this situation is to write and make available on a web page or pages a set of answers to frequently asked questions (FAQs). Such a web page provides a basic solution to the problems mentioned earlier, but except in the simplest and most static cases, it requires continued expert maintenance to keep the FAQ list current and organized. In addition, if the number of FAQs surpasses a few dozen, it becomes difficult for users to navigate the FAQ pages to find the answers they seek.

At the opposite end of the sophistication scale, a number of conversational interfaces to knowledge bases have recently appeared, which can be personified as human or character "chatbots" or represented more soberly as simple input-and-response text fields. A user is invited to enter natural language questions and will receive replies, the quality of which de-

depends on the level of natural language understanding the system has of both queries and items in the knowledge base. Although continuing progress is being made in the question-answering field (Voorhees and Harman 2001), the commercially available chatbots are based mainly on pattern recognition and prewritten scripts, which require a sizable knowledge engineering effort to create and maintain and, therefore, are most feasible for support pages that change slowly. We believe that some sort of metaknowledge (as is represented by the patterns and scripts) is indeed an essential element in facilitating access to knowledge. However, it is also one of our goals to minimize the level of human effort necessary to construct and maintain the knowledge base. We therefore tend to prefer approaches in which the AI is behind the scenes, invisibly supporting users in their interactions with the software.

Our approach centers around a dynamic database of FAQ documents, which we call Answers. Metaknowledge relating to the usefulness of, and relationships among, Answers is acquired automatically as the knowledge base is used. This metaknowledge is utilized to spare the experts from most organizational upkeep and also to make it easier for users to find Answers. By means of the architectural design, with its close coupling of end user questions and CSR answers, the creators of the knowledge base tend to be kept up to date on the information needs of end users, closing a feedback loop that optimizes operation of the system. Because of the general design, and the broad configurability of the system, it can be adapted to many different situations involving knowledge producers and consumers.

In this article, we describe how this approach is embodied in RightNow eSERVICE CENTER (eSC). After briefly introducing the overall system, we describe in greater detail those aspects of the application related to the knowledge base because this is where most of the AI techniques come into play. We also discuss the experiences of customers using eSC.

The RightNow eSERVICE CENTER Application

RightNow eSC is an integrated application that combines e-mail management, web self-service, collaborative live chat, and knowledge management. Most customers choose to deploy it in a hosted environment, but it is also available for individual installations on multiple platforms. It consists of more than 500,000 lines of code, primarily in C but also in C++, JAVA, and PHP as well as HTML. The first prototype was construct-

ed four years ago; the most recent significant upgrade, to version 5.0, involved about 11 months of effort by approximately 16 full-time developers and 7 quality-assurance testers.

The core of the application, from an AI perspective, is the publicly visible Answer knowledge base and the tools by which it is created, maintained, and accessed, which are discussed more fully in the following section. In addition, there is a roughly parallel set of private customer service incidents that are fully tracked from initial creation (by electronic mail, web form, or live chat) through resolution and archiving. Some of our customers use eSC only for the functions associated with the incidents, and others use it exclusively for self-service web support pages. Most use both aspects of the software, and many have an Answer knowledge base that is quite dynamic and comprises 100's or even 1000's of documents, but numbers of non-public incidents are typically much larger.

Other important features of eSC not discussed in this article include extensive administrative functions; customization options; and a wide variety of reports to aid in analysis of transaction statistics, CSR performance, and web site use. One AI-related feature that we allude to briefly here is an emotional index that is determined for incoming messages as well as in real time for agents involved in live chat sessions with end users. The emotional index rates a message on a scale from negative (upset, angry) through neutral to positive (happy) and is derived using a lexicon of rated words and a set of grammatical rules applied to the part-of-speech tagged text, produced as described in a later section. This index can be used in rules for routing incoming messages, for example, sending angry messages to a veteran CSR while perhaps providing an automated response to positive ones meeting some other criteria.

Constructing an Organic Knowledge Base

In traditional practice, knowledge bases have been constructed by domain experts, who do their best to record, in some form of document, what they know and believe to be necessary for a given task or situation. This paradigm might work reasonably well in capturing knowledge for narrow, static subject areas, but in the case where information needs are constantly changing, the burden of frequently adding new knowledge items can become significant. Although it might be easy to predict that the introduction of a new product will lead to inquiries related to the product, it is not as easy to foretell what external events, such as

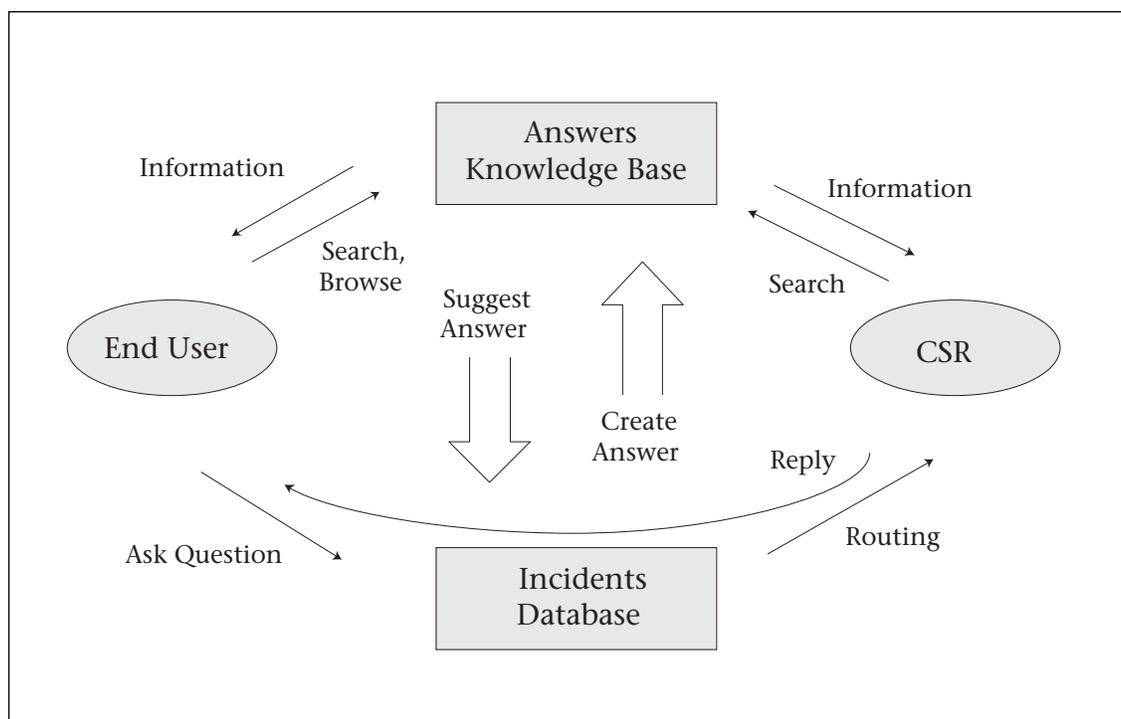


Figure 1. Principal Knowledge-Related Transactions in RIGHTNOW Web eSERVICE CENTER.

End users search the Answer knowledge base and may submit questions. Customer service representatives both use the knowledge base and add to it by creating new Answers.

a new law or regulation or new products offered by competing companies, will cause a shift of end user information needs. In the absence of human maintenance, conventional Answer lists are brittle in the sense that they break as information becomes out of date or irrelevant. Our aim has been to construct a more robust framework that would use AI methods to do as much as possible and, thus, require minimal human resources.

The eSC knowledge base is termed *organic* because of the natural way it is seeded and evolves over time. A key element of our system is that both growth and organization are responsive to end users' shifting demands; thus, eSC integrates question-and-answer channels and works in the following way (figure 1). The knowledge base is first seeded with a relatively small set of Answers to the most predictable or FAQs. Many end users coming to the support web site will find their answers among these, but if not, they are encouraged to submit their questions by e-mail or the web-based form provided on the support home page. As CSRs respond to these, they naturally tend to become aware of trends and commonalities among incidents. At any time, a CSR response, or an edited and extended version of one, can be proposed as a potential Answer. Depending on organizational practices, the item could be re-

viewed or edited by collaborators or managers before being made publicly available. The general availability of the answer will then result in a reduction of incoming queries on the topic. Even if such queries continue, there is now an item in the knowledge base available to end users and CSRs as a Suggested Answer. Answers are suggested by treating the end user's message as a search query, then filtering the returned set of Answers by requiring them to be in the same topic cluster as the query. They can be provided automatically to end users who submit questions or to CSRs responding to questions.

Just-in-time knowledge delivery was described as an important concept in knowledge management by Smith and Farquhar (2000). We extend this just-in-time paradigm to apply to knowledge creation, which is driven by end users and their unmet needs, and the CSRs' or other experts' time and effort are conserved. Thus, users are more likely to find relevant and satisfactory Answers, and the CSRs will have more time to focus on the usually smaller fraction of nonrepetitive questions.

Navigating a Self-Learning Knowledge Base

It is widely understood that knowledge comprises not only isolated facts or data but also re-

The screenshot shows a web browser interface for the University of South Florida Information Technology Division. At the top, there are navigation links for BHRIT Divisions (BPA, HR, IT, ODT, VAP, VP OFFICE) and CONTACT IT STAFF. Below this is a support menu with tabs for Support Home, Answers, Ask Questions/Report a Problem, and My Stuff, along with Login and Help buttons. A search bar is present with a 'Search' button and a 'Browse' button. The search interface includes a 'Category' dropdown set to 'Info Tech', a 'Search Text (optional)' input field, and 'Search by' and 'Sort by' options. Below the search area, it indicates '86 Answers Found' and 'Page: 1 of 5'. A table lists the answers with columns for 'Select Category', 'Subject', and 'Solved Count'.

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 ESC Support Page of the University of South Florida Information Technology Division.

The page is configured to list by default the historically most useful Answers (highest solved count). As a result, there is a high probability that a relevant Answer can be viewed with a single click. The ask-a-question tab provides a form by which questions can be submitted to support personnel (for example, customer service representatives).

relationships among them as well as perspectives on their importance, relevance, and so on. A knowledge base that encompasses such meta-knowledge provides a much better match to human user habits and expectations and is consequently easier to use. We call the ESC knowledge base *self-learning* because it acquires this metaknowledge through a number of AI-based techniques rather than through human-constructed ontologies, templates, or other forms of knowledge engineering. The techniques we use include natural language tools for feature selection, adaptive clustering and classification of text documents, and collaborative filtering and swarm intelligence methods to extract implicit user feedback. We discuss these techniques as they might come into play during an interaction with the knowledge base.

An illustration of the first end user view of a typical knowledge base is shown in figure 2. By default, this page lists a configurable number of knowledge base Answers, sorted in order of their solved count (rightmost column of the display). The latter quantity is a measure of how helpful the answer is likely to be, based on the analysis of previous user activity, as we describe shortly. Because the solved count distribution tends to be Zipf-like; that is, with a few Answers of highest current interest and many

in lower demand, there is a fair probability (dependent on the total number of Answers) that the information needed is available within a single click of this first page.

If the title of an Answer looks promising to an end user, a click on it brings up the full text (along with graphics or any other additional information that can be provided on an HTML page). If the information there does not completely answer the user's question, he/she might return to the original list or might elect to follow one of a ranked set of Related Answer links attached to the Answer page. The relatedness ranking is derived from two sources: (1) a simple document similarity measure based on word cooccurrence (with stop-word removal and stemming) and (2) accumulated implicit recommendations of previous users.

To capture user perceptions of usefulness and relatedness of Answers, we use both explicit and implicit feedback in a manner inspired by collaborative filtering (Levy and Weld 2000) and swarm intelligence (Dorigo, Di Caro, and Gambardella 1999) algorithms. Associated with each Answer is a usefulness counter (solved count) that is increased each time the Answer is viewed and can also be increased (or decreased) by an explicit rating that the user submits by clicking one of a set of rating buttons displayed with the Answer. In addition, a sparse link matrix structure is maintained, the corresponding element of which is incremented each time an end user navigates from one Answer to another, presumably related, one. Because a new knowledge base has no user-derived links, these links are initially supplied according to statistical text similarity alone. In a way analogous to pheromone evaporation in social insect navigation, both usefulness and link values are periodically reduced in strength when not reinforced. This "aging" keeps the knowledge base responsive by emphasizing recent usage patterns.

Of course, this links matrix contains noise in the sense that not every transition is necessarily made by users only on the basis of perceived relatedness. Nonetheless, when 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. The potential tendency for highly ranked Answers to be overly reinforced because of their position in the list is mitigated by several factors. A user is unlikely to select an Answer if it does not appear related to his/her information need (as with any information or web page design, titles are important). If a selection turns out to be mistaken, its usefulness can be downgraded directly by the explicit rat-

ing mechanism and indirectly relative to later Answers that satisfy the user's needs by way of the implicit mechanism. Also, the aging process decreases each Answer's usefulness (solved count) by a constant multiplicative factor, which reduces higher solved counts by greater amounts. For a fuller discussion of these collaborative and swarm intelligence methods, see Warner et al. (2001).

Users with specific information demands, especially if they are less common, can locate information most quickly by searching the knowledge base. Queries entered in the search box allow for a variety of search modes, including natural language input and similar phrase searching (which carries out spelling correction and synonym expansion). A search can be restricted to a given product or category, and returned Answers can be ordered by match weight or historical usefulness. The frequency with which terms are searched for constitutes one report that is useful to system managers. If some commonly entered search terms happen not to appear in the Answer documents, these terms can be added either to Answer-specific lists of keywords or a general synonym list, along with corresponding terms that appear in the knowledge base documents.

Searching in a document collection is a long-studied problem in information retrieval (Baeza-Yates and Ribeiro-Neto 1999). Our basic search algorithm uses a conventional vector representation of queries and documents, with removal of stop words, stemming to recognize related word forms, and phrase identification to increase precision. Beyond the synonym expansion we offer, some other systems can utilize either general (Burke et al. 1997) or domain-specific (Everett et al. 2002) ontologies to further improve searching. We hope to add such a feature in a future release.

End users might or might not come to a support web site seeking specific information, but in either case, they might find it convenient to browse the knowledge base from a higher-level point of view, gaining a broad perspective on the available information. As shown in figure 3, our system offers a browse mode of access where categories of documents are displayed as folders, which are labeled with the key terms most descriptive of their contents. Clicking on a folder opens it to display documents and subfolders corresponding to more specific categories. Merely glancing at the labels on the folders at the highest level gives an outline summary of the contents of the knowledge base. Because the user can navigate by selecting subfolders and individual documents without needing to type search terms, this browse

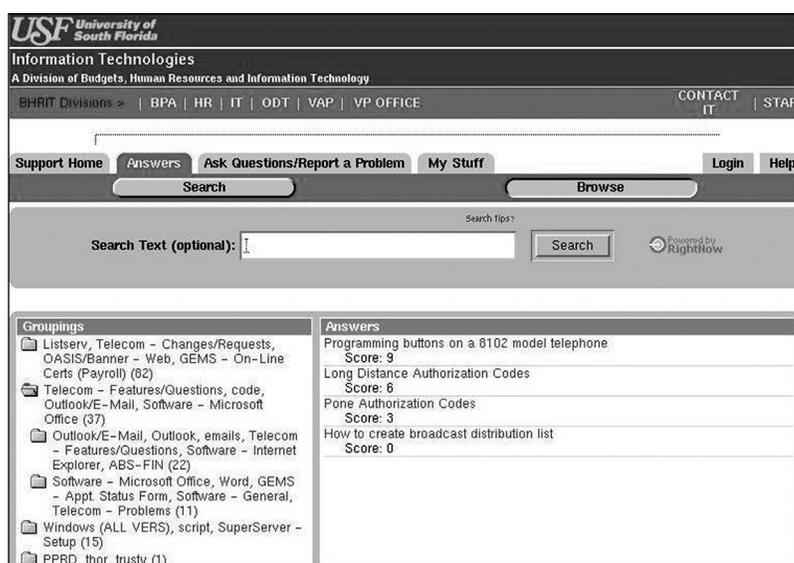


Figure 3. Web Browser Display from the Browse View of the eSC Support Page of the University of South Florida Information Technology Division.

This page displays a hierarchical set of folders and subfolders, where a given folder (like a typical computer file system) can contain both subfolders and Answer documents. The selected folder is the second one at the highest level of the hierarchy (leftmost in the display). Searching can be carried out within a selected browse folder.

mode is especially helpful when the user is unfamiliar with the terminology used in the Answers and, hence, would have difficulty forming a productive search query. Thus, we enlist the user's tacit knowledge, his/her ability to recognize more easily than articulate.

Supporting a browse function without a human-defined ontology requires a taxonomic organization of the text items in the knowledge base. Our method for doing this is illustrated in figure 4. We use a heavily modified version of the fast, hierarchical clustering algorithm BIRCH (Zhang, Ramakrishnan, and Livny 1996), which is run repeatedly while the threshold parameter is varied under fuzzy adaptive control. The best result, according to a clustering figure of merit-incorporating cluster size, cluster cohesion, and branching factor, is used as a basis for learning RIPPER-style classification rules (Cohen 1995). The final topic hierarchy is created by classifying knowledge base items according to the rules, allowing each item to potentially be classified in multiple places. Multiple classification recognizes the inherent multiplicity and subjectivity of similarity relationships. It makes searching by the browse interface much more convenient because the end user can locate an item along various paths without backtracking and does

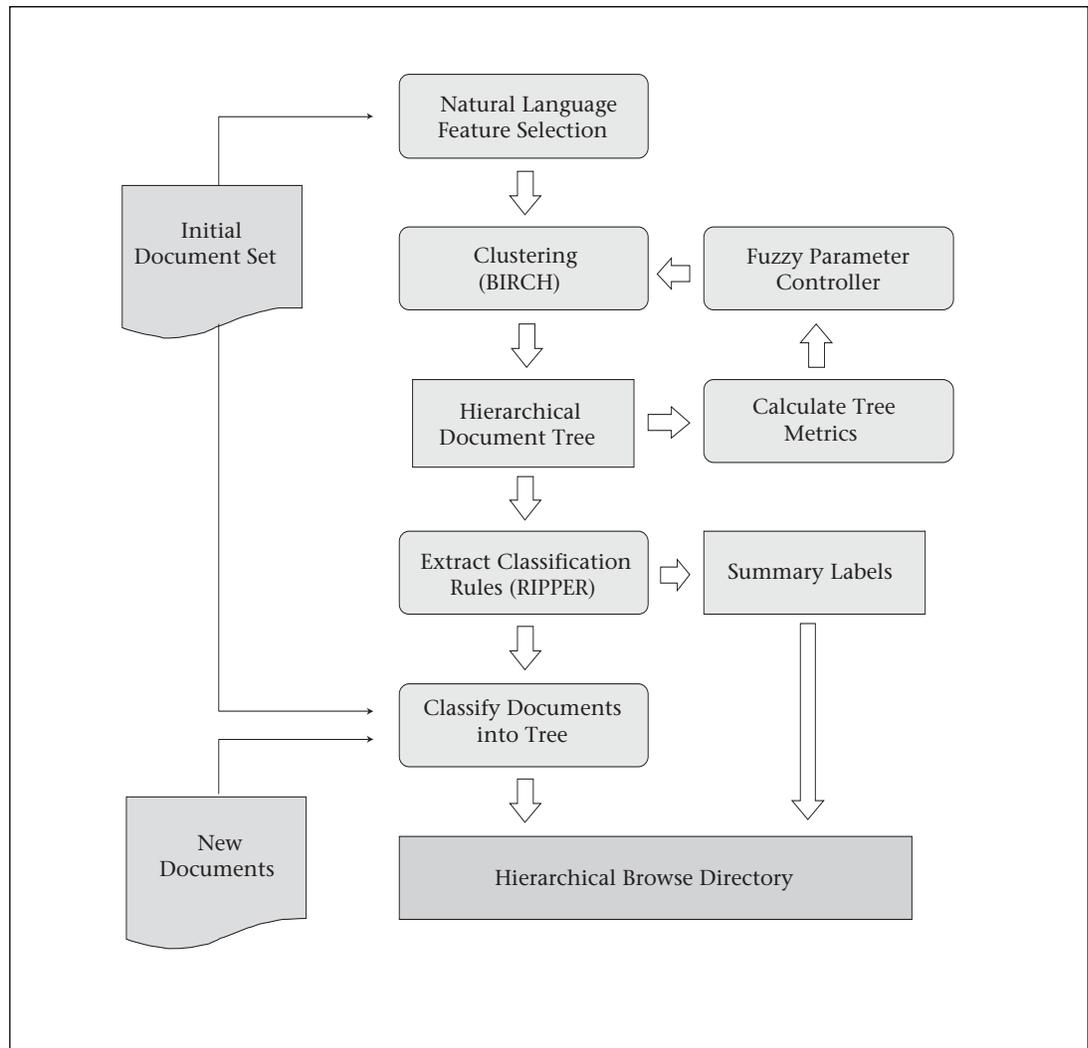


Figure 4. Flowchart Showing the Clustering and Classification Algorithm Used to Create the Browse Display.

Documents are clustered with parameters determined by a fuzzy adaptive controller; classification rules are then derived that are used to place both initial and new documents in the hierarchy.

not have to guess what rigid classification might control the listing.

The features on which the clustering is based are obtained from the document texts by shallow natural language processing involving part-of-speech tagging with a transformation-based tagger (Brill 1994). Noun phrases are identified and receive the highest weight as features, but selected other words are also used. In addition, customer-supplied keywords and product or category names provide highly weighted features. These features are increased in weight if they are frequently searched for by users.

Extraction of the classification rules allows new knowledge base items to simply be inserted into the hierarchy as they are created, in the same way as previous Answers. However, after

a predetermined amount of change in the knowledge base because of modification, addition, or removal of documents, a reclustering is performed so that the browse hierarchy reflects the current state of the knowledge base rather than a fixed hierarchy.

User Experience with E-SERVICE 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 have gathered both aggregate statistics and numerous case studies demonstrating the dramatic reduction of time and effort for knowledge base creation and mainte-

nance and the increase in satisfaction of knowledge base users. Such results are obtained across the spectrum of organizations and applications, including those outside the area of conventional customer service.

The ease of installation is such that it has been accomplished in as little as a day, if initial seed Answers are available and major customization is not needed. As a demonstration that is part of our sales process, companies can set up pilot installations in two to five days. 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 2 weeks. Now, after 2 years, the number has stabilized at about 600. The volume of telephone calls handled daily has dropped 50 percent, from about 25,000 to 12,500, leading to an estimated daily savings of \$62,500.

The ability of a web self-service system to handle dynamic fluctuations in usage can be very important. As one example, the January 2001 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 percent over that for the previous rate hike. Attempting to handle such volume by telephone or e-mail would have resulted in huge backlogs, but with eSC, the load was managed smoothly.

A quantitative measure of end user success in finding information, as well as cost reductions to a company, is the *self-service index*, defined as the percentage of end users who are able to find their own answers online rather than send a message to a CSR. Table 1 is excerpted from a Doculabs study in which it was found that depending on the type of organization,¹ the self-service index using eSC ranged from 75 percent to almost 99 percent, averaging 87 percent. According to anecdotal statements from customers, these benefits are largely attributable to the key elements of the self-learning knowledge base, as described earlier.² Not evident in this metric is the cost saving resulting from the knowledge-acquisition processes facilitated by eSC: The experts' time and energy is used much more effectively (Durbin et al. 2002).

In addition to standard customer service, eSC is flexible enough to be used in other knowledge management settings. A number of organizations use it internally to provide information to their members, from general interest news to specific areas such as personnel forms

and procedures. Within our company, Right-Now Technologies, it is also used as a shared information resource between quality assurance and development teams. In this use, quality-assurance testers submit bug reports (analogous to customer questions), and developers respond to them. A single bug history can contain a number of transactions involving several people on each team. This system not only facilitates the communication between the two work groups but provides a valuable organizational memory for future reference.

Discussion

Despite the current level of success of eSC, there is certainly room to do better. Some improvements, such as making clustering more adaptive to differing knowledge bases, are well under way. More difficult is the problem of automatically producing good summary labels for the clusters; our current heuristics work well in some cases and less well in others. The area of multidocument summarization is one of active current research (see, for example, Mani and Maybury [1999]), and one of our priorities is to improve this aspect of eSC in future releases.

More qualitative enhancements can be obtained from applying AI techniques to a greater number of functions. Incident routing, text categorization, and natural language processing are all areas we are working on. We are also adding more intelligence to diagnostic reports, for example, to automatically identify trends in end user questions and recommend additions to the knowledge base. This type of report is especially important in organizations with a large number of generalist experts or CSRs who only have an incomplete view of incoming questions.

As knowledge bases inevitably become larger and more complex, the need for a system such as eSC increases. The knowledge bases that eSC has been used with to date have not been extremely large, very seldom reaching more than a few thousand documents (although many more items are normally in the incidents database). Algorithmic changes might become necessary to scale some of the behavior to much larger databases, especially for processing that is done while an end user is waiting.

Another trend affecting many internet-based applications is that toward greater personalization of user interfaces. Care must be exercised to ensure such customization facilitates and enhances rather than constrains and obscures. In an information-finding task, one doesn't want to miss something because of an

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 initiate contact with a support person (escalating) by e-mail or online chat.

agent's faulty assumption. The extent to which significant personalization is feasible and desirable for frequent or one-time users is still being investigated.

In our experience developing functions for ESC, we have learned several lessons that apply perhaps especially to commercial software. For one, scalability is always a potential issue, and one should assume that some customers in the near future will use the application on a larger scale than any previous one. In the same vein, each new customer is a potential outlier in terms of knowledge base characteristics, so it is important to test algorithms on the widest possible range of real and simulated data. To ease the development task, an effort should be made to identify complementary algorithms that can serve a number of purposes in different combinations. Finally, in language-dependent natural language-processing functions, it is best to use generally applicable statistical methods wherever these can meet the need, thus minimizing the amount of language-specific engineering required.

It's important to realize that the developer's job is not done when an application is delivered. In addition to potential bug fixes, there is

constant pressure from customers to add features or modify an existing feature to fit particular needs. It is, of course, essential to respond in some way to such requests, if only to schedule an enhancement for a future release. In general, many functions can be added without increasing complexity for other users (ease of set up and use is always a goal) and without creating maintenance nightmares, by using a system of configuration options with default values such that most customers need never be aware of them.

Implementation and support issues are critical for commercial software. In the case of ESC, about 10 implementation engineers perform installation, upgrade, and various maintenance tasks for 1100+ customers. Meanwhile, approximately 20 technical support personnel respond to customer questions, and 10 provide professional services such as integration with other applications or HTML customization.

Conclusions

We have described the web-based customer service application RightNow ESC, which relies on a number of AI techniques to facilitate construction, maintenance, and navigation of a

knowledge base of answers to FAQs. These techniques include collaborative filtering, swarm intelligence, fuzzy logic, shallow natural language processing, text clustering, and classification rule learning. Many of these individual techniques have been used for similar purposes in other commercial applications, but we know of no other system that combines all of them. Customers using ESC report dramatic decreases in support costs and increases in customer satisfaction because of the ease of use provided by the self-learning features of the knowledge base.

The principles and methods embodied in ESC are also applicable in other settings. For example, the relationship between a government agency and concerned citizens is closely analogous to that between a business and its customers. In fact, organizations and associated constituencies with information needs are ubiquitous in our modern society. In the conception and development of ESC, we have emphasized the generalizable features of dynamic focus on current information needs, ease of updating and maintenance, and facilitated access to the knowledge base.

Notes

1. J. Watson, G. Donnelly, and J. Sheha. 2001. The Self-Service Index Report: Why Web-Based Self-Service Is the ROI Sweet-Spot of CRM. www.doculabs.com.
2. Case studies are available on the following web page: www.rightnow.com/resource/casestudies.php.

References

- Baeza-Yates, R., and Ribeiro-Neto, B. 1999. *Modern Information Retrieval*. New York: Association of Computing Machinery.
- Brill, E. 1994. Some Advances in Transformation-Based Part-of-Speech Tagging. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, 722–727. Menlo Park, Calif.: American Association for Artificial Intelligence.
- Burke, R. D.; Hammond, K. J.; Kulykin, V. A.; Lytinen, S. L.; Tomuro, N.; and Schoenberg, S. 1997. Question Answering from Frequently Asked Question Files: Experiences with the FAQ FINDER System. *AI Magazine* 18(2): 5–66.
- Cohen, W. H. 1995. Fast Effective Rule Induction. In *Machine Learning: Proceedings of the Twelfth International Conference*. San Francisco, Calif.: Morgan Kaufmann.

Dorigo, M.; Di Caro, G.; and Gambardella, L. M. 1999. ANT Algorithms for Discrete Optimization. *Artificial Life* 5(2): 137–172.

Durbin, S. D.; Warner, D.; Richter, J. N.; and Gedeon, Z. 2002. Organic Knowledge Management for Web-Based Customer Service. In *Organizational Data Mining: Leveraging Enterprise Data Resources for Optimal Performance*, eds. H. Nemati and C. Barko. Hershey, Pa.: Idea Group. Forthcoming.

Everett, J.; Bobrow, D. G.; Stolle, R.; Crouch, R.; de Paiva, V.; Condoravdi, C.; van den Berg, M.; and Polanyi, L. 2002. Making Ontologies Work for Resolving Redundancies across Documents. *Communications of the ACM* 45(2): 55–60.

Levy, A. Y., and Weld, D. S. 2000. Intelligent Internet Systems. *Artificial Intelligence* 118(1–2): 1–14.

Mani, I., and Maybury, M. T., eds. 1999. *Advances in Automatic Text Summarization*. Cambridge, Mass.: MIT Press.

Smith, R. G., and Farquhar, A. 2000. The Road Ahead for Knowledge Management: An AI Perspective. *AI Magazine* 21(4): 17–40.

Voorhees, E. M., and Harman, D. K., eds. 2001. NIST Special Publication 500-249: The Ninth Text Retrieval Conference. U.S. Department of Commerce, National Institute of Standards and Technology.

Warner, D.; Richter, J. N.; Durbin, S. D.; and 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*, 467–472. New York: Association for Computing Machinery.

Zhang, T.; Ramakrishnan, R.; and 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*, 103–114. New York: Association for Computing Machinery.

Stephen D. Durbin received a Ph.D. in physics in 1984 from the University of California at Berkeley and spent his formative years as a biophysicist and materials scientist at the University of California at San Diego; Carleton College; and the Materials Research Institute of Tohoku University, Japan. The allure of AI converted him to computer science, and he recently joined the Applied Research Group at RightNow Technologies. His particular interests include natural language processing, neural networks, information visualization, and knowledge management. His e-mail address is sdurbin@rightnow.com.



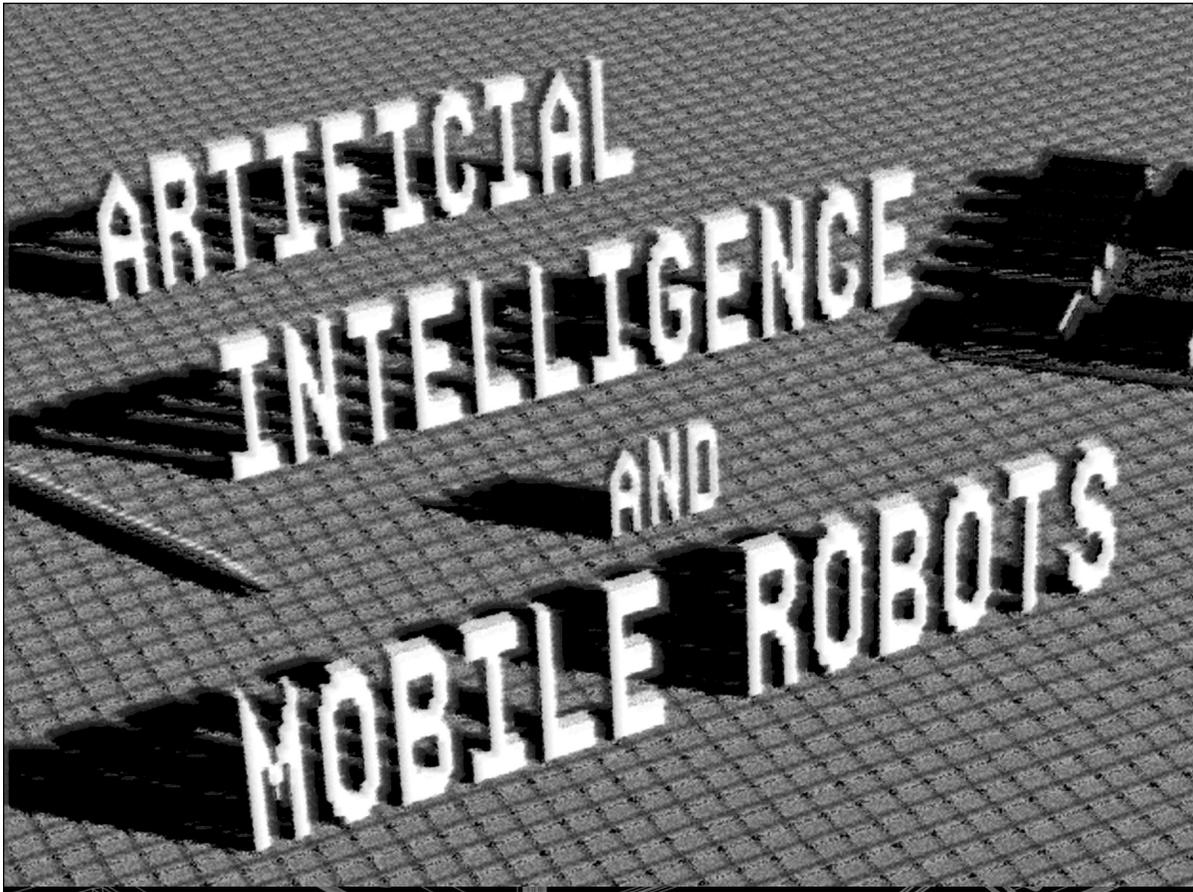
Doug Warner is head of applied research at RightNow Technologies, where he has worked since 1999. Before joining RightNow, he was a tenured Ph.D. student in psychology at the University of New Mexico, where he received his M.S. in 1994. He also has a B.S. (1991) in computer science and psychology from the New Mexico Institute of Mining and Technology. His current research interests include collaborative filtering and visualization of complex information. His e-mail address is doug@rightnow.com.



J. Neal Richter is an applied AI researcher at RightNow Technologies. He has an M.S. in computer science from Utah State University and is a Ph.D. student at Montana State University. Current interests include machine learning, natural language processing, fuzzy logic, dynamic systems, cellular automata, evolutionary computation, and theoretical evolutionary computation models. His e-mail address is nealr@rightnow.com.



Zuzana Gedeon is a member of the Applied Research Group at RightNow Technologies. She received her M.S. in applied mathematics from Georgia Institute of Technology and is currently working on her Ph.D. in computer science at Montana State University. Her professional interests include clustering and data-mining techniques. Her e-mail address is zgedeon@rightnow.com.



Artificial Intelligence and Mobile Robots

Case Studies of Successful Robot Systems

Edited by David Kortenkamp, R. Peter Bonasso, and Robin Murphy

The mobile robot systems described in this book were selected from among the best available implementations by leading universities and research laboratories. These are robots that have left the lab and been tested in natural and unknown environments. They perform many different tasks, from giving tours to collecting trash. Many have distinguished themselves (usually with first- or second-place finishes) at various indoor and outdoor mobile robot competitions.

Each case study is self-contained and includes detailed descriptions of important algorithms, including pseudocode. Thus, this volume serves as a recipe book for the design of successful mobile robot applications. Common themes include navigation and mapping, computer vision, and architecture.

6 x 9, 400 pp., ISBN 0-262-61137-6

**To order, call 800-405-1619 or go to <http://mitpress.mit.edu>.
Distributed by The MIT Press, 55 Hayward, Cambridge, MA 02142**