Evaluating a Voice Recognition System: Finding the Right Product for Your Department

Mary Freeh, Matt Dewey, and LuAnn Brigham

The Department of Radiology at the University of Utah Health Sciences Center has been in the process of transitioning from the traditional film-based department to a digital imaging department for the past 2 years. The department is now transitioning from the traditional method of dictating reports (dictation by radiologist to transcription to review and signing by radiologist) to a voice recognition system. The transition to digital operations will not be complete until we have the ability to directly interface the dictation process with the image review process. Voice recognition technology has advanced to the level where it can and should be an integral part of the new way of working in radiology and is an integral part of an efficient digital imaging department. The transition to voice recognition requires the task of identifying the product and the company that will best meet a department's needs. This report introduces the methods we used to evaluate the vendors and the products available as we made our purchasing decision. We discuss our evaluation method and provide a checklist that can be used by other departments to assist with their evaluation process. The criteria used in the evaluation process fall into the following major categories: user operations, technical infrastructure, medical dictionary, system interfaces, service support, cost, and company strength. Conclusions drawn from our evaluation process will be detailed, with the intention being to shorten the process for others as they embark on a similar venture. As more and more organizations investigate the many products and services that are now being offered to enhance the operations of a radiology department, it becomes increasingly important that solid methods are used to most effectively evaluate the new products. This report should help others complete the task of evaluating a voice recognition system and may be adaptable to other products as well.

Copyright © 2001 by W.B. Saunders Company

THE DEPARTMENT OF RADIOLOGY at the University of Utah is designed such that groups of radiologists specialize in an area and interpret primarily examinations from their specialty. They work in reading rooms dedicated to

that specialty. In addition, the department supports 16 residents who rotate between specialties, usually on a monthly basis. Nights and weekends are covered by on-call radiologists backing up the residents who are assigned rotational shifts. This configuration results in both challenges and opportunities in our transition to digital voice dictation.

We had originally hoped to introduce digital voice dictation on a limited scale, setting one station up in each reading room, to allow the radiologists a chance to test it out before it was implemented in the entire department. We soon realized that such a plan would cause problems for the residents and the on-call radiologists. All of the vendors we spoke with indicated that it is very difficult to move back and forth between analog and digital dictation methods, resulting in much frustration for the radiologists and residents. Since one of our objectives was to reduce the frustrations caused by change, we were hesitant to phase the implementation.

Our second idea was to implement in only one reading room first, then roll out the new process to the entire department. This had its limitations in that only one group of radiologists would get a chance to test the system. However, it had its advantages in that it would only affect one group, limiting the impact to one area of the department. The residents would still have to transition in and out of the system, but if we could keep our testing time short we hoped to affect only a small portion of the residents who were rotating during the time of the test. We chose the second approach.

Once we decided how we wanted to roll out the new process, we needed to identify the product that would best meet our needs. The method of operating in our department, described above, meant we had some special user operations requirements. Since we were already using a picture archiving and communications system (PACS), reading nearly all of our studies off soft copy, we knew we had special system interface needs. We also knew other departments had experienced problems implementing voice dictation so we were especially interested in the following: the quality and flexibility of the speech engine, the training and service support the vendor would provide, and the overall

From the Department of Radiology, University of Utah Medical Center, Salt Lake City, UT.

Address reprint requests to Mary Freeh, Department of Radiology, University of Utah Medical Center, 50 N Medical Dr, Salt Lake City, UT 84132. E-mail: Mary.freeh@hsc.utah.edu.

Copyright © 2001 by W.B. Saunders Company

0897-1889/01/1402-1002\$35.00/0

doi:10.1053/jdim.2001.23802

Speech Engine	Score (1-5)	Score (1-5)
Ability to recognize large number of radiology specific and medical specific terms		
Ability to easily add words to the dictionary and update the dictionary as improvements are made industry wide		
Ability to distinguish between various word uses: colon, to, two		
Ability to handle accents from foreign speakers		
Ability to train on speaker's voice in less than half an hour		
Ability to continue to improve training of speech engine as radiologist dictates		
Subtotal		
Vendor Support	Score (1-5)	Score (1-5)
Training of radiologists		
- individual sessions		
- group sessions		
- number of days on site for training support		
- willing to train the trainer		
On site installation support		
- number of personnel		
- number of days on site for problem resolution		
Training for onsite system administrator		
- number of days of training		
Service support		
- Contact method (800 #)		
- Turnaround time for response/ass stance		
- Equipment replacement time frame		
System downtime plans		
- backup of voice files		
- recovering incomplete reports		
Subtotal		
The Company	Score (1-5)	Score (1-5)
Length of time in business		
Current financial position		
Number of current customers		
Feedback from current customers		
% of sales committed to R & D		
Subtotal		
TOTAL		
Cost Comparison		
- Hardware Costs		
- Software Costs		
- Service and Support		

Fig 1. Vendor scorecard.

distribution process to referring physicians, the ability for our radiologists to interact with the system off site, the flexibility of the system to accommodate residents, and the efficiencies gained from the enhanced microphone capabilities.

SYSTEM INTERFACE

Again, to minimize the impact of change for our users, we hoped to obtain a system that could be interfaced with our existing radiology information system (RIS) and PACS. We have the IDX RIS so it was critical for our voice recognition vendor to have experience with an IDX interface. We were looking for the capability to interface bidirectionally between the speech engine and IDXRad (IDX Systems Corp, Burlington, VT). This would allow our radiologists to receive lists of examinations in the voice system that had been completed in the RIS and if necessary they could edit or sign in the RIS. We use the Radworks PACS system (Marconi Medical Systems, Cleveland, OH) and wanted to be sure the voice product could be interfaced to that system. We hoped the speech engine could be used to initiate the retrieval of an image and the

Our primary concern was to keep the impact to the users to a minimum. Knowing that changing from an analog dictation system would take many adjustments for all, we hoped to minimize those adjustments by making sure the new process would be an overall improvement to current methods. Therefore, we paid special attention to the following: the sign-on process, the editing process, the

microphone used for dictation could be used to drive most of the image manipulation tasks the radiologist needed to perform.

SPEECH ENGINE

Our biggest area of interest was the size and flexibility of the speech engine. We wanted to be sure we were not locked into a speech engine that would not be able to be updated as technology improved. Training the speech engine for a particular user, especially if a user was foreign-speaking, was also important. And we wanted to be sure the speaker could be recognized even if there were changes to his or her voice quality, as might happen when a radiologist has a cold.

VENDOR SUPPORT

Like with any new piece of equipment, it is critical to have good vendor support during the install phase and beyond. We were most concerned with training support, installation support, ongoing service support, and downtime processes.

THE COMPANY

The University of Utah Hospitals and Clinics has in the past been negatively affected when a decision was made to obtain a product from a company that was not able to continue support of that product. Therefore it was essential that we evaluate the strength of the company and provide an educated guess as to the longevity of the company.

SUMMARY

Acquiring a new product as extensive as a digital voice dictation system is a major investment that requires careful reengineering. The chart we designed helped us determine if a vendor was available that could provide what we wanted and needed. Using the chart as a guide while we evaluated the product helped us to ask the right questions and make our decision faster.

BIBLIOGRAPHY

1. Houston J, Rupp F: Experience with implementation of a radiology speech recognition system. *J Digit Imaging* 13:124-128, 2000
2. Couris J: Road to voice recognition includes planning, training. *Diagnostic Imaging* 35-39, 1999 (suppl)