

# Challenges and Opportunities in Implementing Speech Recognition in a Legal Enterprise

The integration of speech recognition with word processing and other computer applications and systems can significantly enhance the efficiency of document generation and improve overall productivity in the legal environment.

## CONTENTS

---

I. Introduction: What Is Speech Recognition and How Is It Used?.....	2
II. Summary of Benefits of Speech Recognition for Legal Professionals.....	3
III. Challenges of Implementation and Use.....	5
IV. Pilot Trial Rationale and Methodology.....	8
V. Determining Return on Investment from Speech Recognition.....	10
VI. Sample Worksheet: Return on Investment for Implementing Speech Recognition in a Legal Enterprise.....	12

---

## I. Introduction: What Is Speech Recognition and How Is It Used?

---

### SPEECH RECOGNITION—FUNDAMENTALS

Speech recognition technology enables users to communicate with computers and other speech-enabled devices by voice.<sup>1</sup> Typical products that use speech recognition range from software that produces formatted transcriptions of dictated text to devices that respond to spoken commands.

In office environments, transcription software applications provide a practical alternative for keyboarding text and data into a computer, which can be a challenge to people who cannot type well or at all, or prefer not to type. For some devices and in some environments, speech may be the only practical input method, or may provide benefits that are superior to alternatives.

### TYPICAL USERS

Potential users of speech recognition include virtually anyone who is able to speak and who uses computers to generate documents and communications. Speech recognition products specifically designed for use by legal professionals can immediately produce transcriptions, and eliminate the need for transcription services. Network capability allows convenient access throughout a legal enterprise. Typical user profiles include:

**Partners, associates, private practice attorneys, researchers, law students, paralegals, and other legal professionals.** A growing number of legal professionals work with minimal administrative staff. Many have never learned to type or are uncomfortable with a keyboard, and can create text at higher rates of speed by dictating than by keyboarding.

**Small practices.** Typically with minimal staff, legal professionals in small practices may find that dictating documents and accessing information by voice reduces the time needed for paperwork and research. The resulting free time can be used for additional client services, education, and other primary responsibilities.

---

### <sup>1</sup>HOW DOES IT WORK?

The basis of practical speech recognition technology is statistical, and generally involves the use and interrelationship of three fundamental types of information:

- A *lexicon*—a group of words and their pronunciations
- A *language model*—which specifies the relative likelihood of a sequence of words
- An *acoustic model*—the sound-related variables of a given pronunciation

Automatic speech recognition algorithms relate acoustic data (user speech) to the intended linguistic equivalent (transcription or control action). This task can range from identifying a few, readily distinguishable commands from a finite grammar—a relatively "easy" speech recognition task—to the far more computationally intensive challenge of accurately discriminating and recognizing long sequences of words, numbers, and system commands in continuous-speech, large-vocabulary dictation. The larger and more varied the lexicon, the greater the number of hypotheses the computer must evaluate to make the most plausible assessment of a given utterance— and the greater the demand placed on computer processor and memory resources. With some software, if a "misrecognition" is made, users can correct by voice command, and "teach" the application the intended transcription of the spoken words.

**Mobile professionals.** People who are frequently away from their desks can benefit from the convenience of recording dictation on a portable recording device for subsequent automatic transcription on a computer.

**People with disabilities.** Speech recognition technology is a practical and viable alternative for personnel who find it difficult or impossible to keyboard due to disability or injury.

**Senior citizens.** Keyboarding may pose special difficulty for older users who have never learned to type, suffer from arthritis, or tire easily. Speech recognition technology enables them to dictate letters and e-mail, or Web addresses and hyperlinks, and eliminate the physical work of keyboarding.

#### FUNDAMENTALS SUMMARY

Speech recognition provides a practical alternative for keyboarding text and data into a computer, which can be a challenge to people who cannot type well or at all, or prefer not to type. In some environments and situations, speech may be the only practical input method, or may provide benefits that are superior to alternatives.

## II. Summary of Benefits of Speech Recognition for Legal Professionals

**1. Increased productivity**—Speech recognition technology allows users to get more done in less time, freeing them up to focus on the primary tasks at hand. This time is gained by dictating into computers, which produces text faster than by traditional keyboarding. Users can create, edit, and format documents, send e-mail, access and update records, and navigate their desktop and the Web by speaking at a normal pace.

**2. Economy**—Eliminate the need for transcription services. Electronic and hardcopy transcriptions of dictated reports, letters, and other documents can be produced rapidly by a computer. In comparison, manual transcription incurs additional costs and turnaround time, requires skilled transcriptionists, and risks misreading of handwritten materials as well as the potential loss of the original recording or handwritten document in transit.<sup>2</sup>

**3. Speed**—Speech recognition generates text faster than keyboarding. The best typists average 50 to 90 words per minute. In contrast, text can be dictated successfully for automatic transcription at 100 to 180 words per minute—twice as fast as typing, on average, and with little or no physical effort.

In addition, a string of computer commands can be combined for activation by a single word or brief phrase called a "macro" command to perform complex tasks. Using speech-enabled macro commands and application templates, the time required to perform usually lengthy tasks can be reduced to seconds. Applications can be launched, for example, by saying, "*start Word*." Boilerplate text of virtually any length can be inserted by saying, for example, "*standard tort outline*." The combination of using templates to create,

<sup>2</sup> Witt, D.J.; Transcription service in the ED. *Amer. J. Emerg. Med.* 1995; 13:34-36.

navigate, and complete forms and calling them up by saying, "*client record ABC Corporation*" or "*invoice Jane Jones Enterprises*" can dramatically reduce the time it takes to produce and revise such documents.

**4. Accuracy**—Handwritten notes can be minimized, along with the concomitant possibility of misinterpretation and potentially serious consequences. Transcribed text can be reviewed on screen as an accuracy check. Transcription accuracy with leading speech recognition products can be very high. Users can train software in minutes to accurately recognize their voice. Overall accuracy improves with software that "learns" by accumulating corrections in "user" files. To facilitate accuracy, some products include extensive vocabularies that can be customized by users, and include specialized terms, phrases, abbreviations, and acronyms. Typically, when software "misrecognizes" dictation, the result is syntactically obvious, and the misrecognition can be easily detected and corrected. Recognition accuracy of 99 percent has been documented in third-party evaluations of over-the-counter products.<sup>3</sup>

**5. Mobility**—Documents can be created virtually anywhere by means of dictation into a high-quality recording device and subsequent transcription on a computer. This capability provides substantial productivity benefits for users who need to generate text when away from their computer, for example during breaks in client meetings or on the road. In situations requiring uninterrupted vision and physical involvement, such as in a car, or on a plane or train, dictation into a lightweight mobile recorder can be more convenient and faster than using a laptop or handheld computer.

**6. Uniformity**—When used across an enterprise, system administrators can ensure that standards and generally accepted professional and institutional practices are supported throughout a workgroup by distributing only standardized macros and vocabularies over the network. Standardized macros can help ensure adherence to the latest protocols and standards throughout the enterprise. With such centralized administrative control, speech recognition can help organizations improve accuracy, consistency, and legibility and thereby reduce the liability and risks associated with nonconformance.

**7. ADA compliance**—In the U.S.A., speech recognition technology can be a "reasonable accommodation" by employers for disabled but qualified personnel—as stipulated in the *Americans with Disabilities Act of 1990*. Such users may be sight-challenged, have limited dexterity due to physical disability, arthritis, or Repetitive Stress Injury (RSI) such as carpal tunnel syndrome or tendinitis. Speech recognition can enable users impaired with disabilities to dictate into their computers hands-free and eyes-free, and also eliminate the possibility of RSI discomfort or injury from repetitive keyboard and mouse manipulation.

**8. Expedience.** Speech recognition can be a valuable expedient in the billable-hour environment of the law office. Information can be accessed and documents, including invoices, can be created immediately by voice, hands-free and eyes-free.

**9. Reduced stress**—Speech recognition can provide significant physical and psychological benefits. Documents can be produced in less time and with significantly less physical effort and stress than by keyboarding.

---

<sup>3</sup> *PC Magazine*, October 22, 1998, Dragon NaturallySpeaking Preferred.

**10. Convenience**—Speech recognition installed on a network allows users to dictate into any configured computer on the network, creating documents in real time virtually wherever they are in an enterprise. After transcription by the computer, an electronic record is available that can be immediately printed out, e-mailed, or linked to a Web page. Users can also play back their digitized dictation in their own voice or save it as a .wav file.

### **SPEECH-ENABLED APPLICATIONS**

Many existing software applications can be speech-enabled with software development kits (SDKs), substantially enhancing their power and value. Even sophisticated speech features can be readily incorporated into existing applications with state-of-the-art software developer toolkits.

#### **BENEFITS SUMMARY**

Speech recognition technology can provide benefits including: rapid generation of documents, productivity enhancements that free up more time for primary tasks at hand, reduction of transcription cost and delay, accuracy, mobility, uniformity, ADA compliance, an expedient means of computer input in billable-hour environments, reduction of stress, and convenience.

## **III. Challenges of Implementation and Use**

---

### **ENVIRONMENTAL CONSTRAINTS**

Speech recognition applications require end-users to speak audibly, possibly for extended periods. In general, environments that are appropriate for telephone calls are likely to be appropriate for speech recognition, but some issues arise that are unique to speech recognition:

- Poor acoustic conditions, such as high ambient noise levels or high echo levels, can compromise recognition accuracy. Although such problems may be mitigated by the use of microphones, microphone headsets, or recorders that are directional and/or noise canceling, results will vary on a case-by-case basis.
- When ongoing dictation is suddenly interrupted by telephone calls or unexpected visitors, the speaker must become accustomed to turning off the input microphone to prevent irrelevant content entering the active document.
- Depending on workplace layout, audible dictation could be a simple nuisance to those nearby, or convey information not meant to be overheard.

## **USER SKILL REQUIREMENTS**

End-users must be familiar with the correct use of the particular software and headset, microphone, or recorder. They must also be able to speak thoughts clearly and fluently in a language supported by the speech recognition application. For some people, expressing themselves extemporaneously will be more difficult than for others, and a period of time will be required to make the shift from the keyboard. Occasionally, the acoustic properties of a particular voice may not yield the desired recognition accuracy.

## **HARDWARE SPECIFICATIONS**

Computer systems (including network components, wiring or optical fiber cabling) must be configured to meet or exceed the requirements specified by the application manufacturer. The quality of input speech is critical to good recognition accuracy so that microphones or recorders should be approved by the speech recognition software manufacturer.

Because speech recognition software places high demands on computer systems and networks, minimum specifications may be insufficient if additional applications or network traffic is to run simultaneously with speech recognition. For example, networked speech recognition applications can require downloading 10–20 MB speech files to client PCs and then uploading them to a speech server. As a result, the introduction of speech recognition to a network that is already experiencing high traffic flow may require upgrades in wiring, cabling, hardware, and/or software.

## **TRAINING**

Training is very helpful for expediting optimal success with speech recognition. While end-users can train themselves to some level of proficiency by using product documentation and tutorials, instruction from trainers certified by the speech recognition software manufacturer is more likely to develop mastery of the product. Leading speech recognition software manufacturers have training professionals on staff who may instruct end-users and also certify trainers in customer, Value-Added Reseller (VAR), and independent training provider organizations. Additional support may include product documentation, manufacturers' scheduled training classes and published curriculum materials, telephone and online support, e-mail support, newsletters, and online forums.

There are three general types of training:

**1. Training end-users.** Users are trained to use the software and headset, microphone, or recorder correctly. This requires learning some basic spoken commands and becoming familiar with the speech user interface. Training by instructors certified by the manufacturer is advisable.

**2. Training the software.** Large vocabulary, continuous speech recognition software is "user dependent." Each end-user must train the software by creating a "speech" file that the computer associates with his/her dictation to produce the intended words and phrases. Creating custom vocabularies requires access to source documents. End-user expertise in the application to be speech-enabled is required and assumed.

**3. Training the trainers.** Leading speech recognition product manufacturers maintain a staff of highly skilled professional trainers to train and certify new trainers in VAR or customer organizations.

### **CUSTOMIZATION**

Some speech recognition applications include features that allow end-users, consultants, or IT departments to create custom voice macro commands and custom vocabularies, and to personalize the standard application commands that came with the software. It is prudent to have such customization performed or overseen by consultants or trainers from, or certified by, the application manufacturer and at least prototyped before the software is deployed to end-users.

### **STANDALONE OR NETWORK SOLUTION?**

Standalone installations are adequate for individual end-users or small offices. However, the most efficient way to use speech across a workgroup is by using a network solution. Users train the software and create their own speech files on any properly configured PC on the network. Their speech files are maintained on a shared network server and can be downloaded as needed at any client node. Users need only train once and can then move from machine to machine, including creating records or documents on one and editing them on another. Attorneys and staff can quickly access client information at any properly configured computer on the network by speaking a simple command such as *“case record John Smith Associates.”*

Networked speech recognition applications can do more than provide the convenience of ubiquity and legibility. Applications that enable system/network administrators to download user voice files, custom vocabularies, and standardized macros and templates to clients from a central server can ensure that generally accepted professional, corporate, or institutional practices and industry standards are supported throughout the organization. Such centralized administrative control can help organizations improve consistency and thereby substantively help avoid liability and risks associated with professional, industry, or statutory nonconformance.

### **MAINTENANCE AND UPGRADES**

Provision for user and instructor training, technical support, and product upgrades should be included with a purchase contract. Some manufacturers or VARs provide upgrade assurance options, so customers can contract in advance for future software upgrades and interim releases. This provides a way to budget for anticipated technology advances.

### **OPEN LICENSE PROGRAMS FOR LARGE USER BASES**

An open license program allows multi-user organizations to streamline the purchase and deployment of software. In one typical open license program, a point value or "unit of measure" (UOM) is assigned to each product in a product line. A series of pricing thresholds or levels are established by the manufacturer as a function of product UOM values multiplied by the quantity of each product purchased on the initial purchase order.

On the basis of this initial purchase, customers qualify their own pricing level, which applies to all orders and reorders for a year. Pricing levels are reviewed annually. Savings can be substantial and renewal contracts are only required annually.

#### **CHALLENGES SUMMARY**

Implementing speech recognition can involve a number of challenges and decisions, including environmental constraints, user skill requirements, hardware specifications, training, customization, maintenance and upgrades, and questions related to the purchase decision itself.

### **IV. Pilot Trial Rationale and Methodology**

---

#### **PURPOSE OF A PILOT TRIAL**

A small-scale evaluation effort called a "pilot" trial allows an organization to test the performance and quantify the value of speech recognition in a real setting. The results of a pilot can be used to confirm expectations before making a broader commitment to speech recognition. Pilot trials typically last for 4 to 8 weeks.

#### **PHASE 1: NEEDS ASSESSMENT/WORKFLOW**

1. A core *pilot implementation team* is established, consisting of representatives from the VAR organization, the manufacturer (sales representative, field engineer), and motivated personnel within the client organization.
2. The scope and structure of the pilot trial are developed:
  - a. All areas of the organization within which speech recognition may be used are identified. Noncritical areas suitable for pilot testing are identified and prioritized.
  - b. The company and team establish the goals that will be used to evaluate the results of the speech recognition pilot.
  - c. Weighted criteria for evaluating the success of the pilot in achieving each goal and providing the basis for a final decision are established. Specific, quantifiable parameters are agreed upon to evaluate the performance and value of the pilot in terms of calculating the likely return on investment (ROI) of deploying speech recognition across the entire organization.
  - d. A timeline is created showing the milestones and responsibilities.
  - e. A commitment is made to hold weekly meetings or conference calls with the client to review status and plans for the next week. A *pilot progress form* is revised weekly with previous week achievements and next-week goals.
  - f. Participants are selected for the pilot who would also be its internal evangelists.

- g. Pilot implementation team member(s) meet with each pilot participant and plot workflow. Each participant's skill levels are assessed and the training curriculum is adjusted appropriately.
- h. A *technical support plan* is established to assure the availability of support for the users.
- i. The pilot implementation team decides which applications, templates, and macros should be speech-enabled and which vocabularies should be built. Key parameters may include potential productivity improvement, frequency of use throughout the organization, the value of enhanced consistency in the context of adherence to practice/corporate guidelines, and the basis for a pilot implementation timeline.
- j. A final timeline for pilot implementation with specific milestones for development, training, deployment *in a noncritical area*, testing, and evaluation is established.

## PHASE 2: PILOT TESTING

1. The selected applications, templates, and macros are speech-enabled and custom vocabularies are built.
2. The speech-enabled applications, templates, and macros and custom vocabularies are deployed to pilot end-user participants.
3. Training of participants is implemented.
4. Pilot testing is launched in the designated noncritical areas. Direct involvement by a manufacturer's representative in conjunction with the VAR throughout this phase is desirable to ensure success. End-users involved in the pilot should be observed during basic software training, while using voice for their applications. They should be given tips on improving speed and accuracy as well as other help as needed. Frustrations and questions should be addressed promptly. Customizations should be refined as necessary. Milestones should be monitored, documented, and modified as appropriate.

## PHASE 3: POST-PILOT EVALUATION AND DECISION

1. Pilot performance is evaluated, the qualitative benefits of full-scale deployment are determined, and the ROI is calculated.
2. The purchase decision is made.
3. If the decision is to move forward with speech recognition, full-scale deployment, training, maintenance, support, and upgrade programs are implemented.

### PILOT SUMMARY

A limited-scale, pilot trial of speech recognition enables a selected group of users in a prospective customer organization to test the performance and quantify the value of speech recognition in a real setting. The results are then used as the basis for larger scale implementation.

## V. Determining Return on Investment from Speech Recognition

---

The value of speech recognition to an organization may be considered in several strategic contexts, within which the determination of ROI is subject to specific workplace and industry factors. Some of these contexts and factors include:

- **Transcription applications**—The calculation of the value of a transcription application, whether off-the-shelf or developed by speech-enabling an existing application, as an alternative to manually transcribing or keyboarding documents is reasonably straightforward. However, beyond savings from the reduction of transcription service costs and delays, there are a number of other, nonintuitive factors to consider in evaluating ROI. These factors include the freeing up of time to perform primary tasks, enhanced conformance with workplace-related statutes and guidelines, and gains in overall productivity.
- **Human Resources-related benefits**—As noted earlier, in the U.S.A., speech recognition is a means of providing a "reasonable accommodation" by employers for disabled but qualified personnel to work with computers—a stipulation of the broadly applicable *Americans with Disabilities Act of 1990*. It can also help to eliminate the risk of Repetitive Stress Injury (RSI), such as carpal tunnel syndrome and tendonitis, caused by keyboard manipulation. As a result, the calculation of ROI should consider two items as downside costs of noncompliance to the ADA: (1) the indirect costs of RSI-related productivity loss, and (2) the direct litigation costs of ADA and RSI issues. An organization's own guidelines and records and/or those of the cogent profession or industry addressing these items can be used as a rational basis for determining costs.
- **Accuracy**—The transcription accuracy of leading speech recognition products can be very high, particularly when transcribed text is printed out or reviewed on screen as an accuracy check. To the extent that machine transcription can eliminate handwritten notes and misspellings, the potential consequences of misinterpreting those notes is also reduced.
- **The value of standardization**—Speech recognition can reduce risk by making generally accepted protocols and standards available throughout an organization with networked speech macros and templates. The calculation of the value of preventing errors of nonconformance can be based on actual cost records and/or the reasonable and expectable costs within the industry, based on actuarial data.
- **Rapid electronic document generation— Document generation**—Speech recognition is an ideal means of producing electronic documents and eliminating handwritten information. In legal practices, the ability to produce standard documents and text fragments immediately suggests enhancement of productivity—particularly with the use of mobile recorders—on an intuitive or anecdotal basis. However, quantification of value in dollar-and-cent terms requires establishing a rational basis for the purchase decision.

A basis can be established by combining (1) the value of productivity gains and (2) the potential cost of noncompliance with precedents and other standards, which could be preempted by the use of speech recognition. In (1), productivity gains are based on a theory of eliminating redundancy and/or assuring the best use of personnel. A rational approach is to determine the dollar value of the time currently spent by attorneys, skilled paralegals, and other support personnel now keyboarding their own or others' documents. With speech recognition, this time could be available instead for additional client meetings, research, administrative work, and education. In (2), noncompliance may be quantified as the short-term and long-term financial costs, for example, issuing from the loss of a case. This total can be quite substantial.

- **The limits of pilot trial data**—While a pilot trial can yield useful scaleable data with which to project the expenses and benefits of larger-scale installations, additional expense and savings data that might not be apparent during a pilot trial should be included in ROI calculations. Examples of such expenses include: capital and maintenance/operations costs accrued by enterprise-wide network improvements, including: server/router hardware, software, cable/connector, and architectural planning and construction required for recabling, loss of plant/office productivity during recabling, network configuration, and debugging. Examples of savings, as noted previously, include increases in overall productivity and avoidance of the costs of noncompliance with industry standards and with workplace-related statutes and guidelines.

**SAMPLE WORKSHEET:  
Return on Investment for Implementing Speech Recognition in a  
Legal Enterprise**

Note: This sample worksheet may not contain all line items applying to specific installations. Contingencies should be addressed according to generally accepted practices of individual organizations.

**I. CURRENT ANNUAL COSTS**

**A. Outside transcription costs**

- 1. **TOTAL Annual cost of outside transcription services** (include incidental costs—messenger, postage, etc.).....
- 2. **Total pages transcribed by outside service per year**..... \_\_\_\_\_
- 3. **Outside transcription service cost per page**  
(Line I-A-1 ÷ I-A-2)..... \_\_\_\_\_

**B. Direct costs of in-house transcription**

- 1. **Total number of pages transcribed in-house per year**..... \_\_\_\_\_
- 2. **Total time (hours) per year spent on in-house transcription**.... \_\_\_\_\_
- 3. **Average hourly salary of transcriptionist(s)**  
Note that transcriptionists may be billable partners, other attorneys, and paralegals who are keyboarding their own documents, as well as salaried support personnel transcribing live or recorded dictation or handwritten text..... \_\_\_\_\_
- 4. **TOTAL annual in-house transcription cost** (Line I-B-2 x I-B-3).....

**C. Other costs of in-house transcription**

- 1. **Intrinsic loss of personnel working on transcription tasks instead of on primary tasks, based on average hourly salary**  
(Enter the value from Line I-B-4)..... \_\_\_\_\_
- 2. **Intrinsic loss of net billings from lost client services**..... \_\_\_\_\_
- 3. **TOTAL other costs of in-house transcription**  
(Total of Lines I-C-1 and I-C-2).....

**D. Additional personnel-related costs**

- 1. **Proofreading and correction costs**..... \_\_\_\_\_
- 2. **Estimated annual cost of noncompliance with American Disabilities Act in computer operations** (legal fees, lawsuit awards/settlements, lost business opportunities) ..... \_\_\_\_\_
- 3. **Cost of computer-related RSI and similar claims**..... \_\_\_\_\_
- 4. **Estimated annual loss of personnel productivity from RSI**..... \_\_\_\_\_
- 5. **TOTAL Annual personnel-related costs**  
(Total of Lines I-D-1 through 4).....

**E. Opportunity costs of failure to use enterprise-wide standards and protocols**

Estimated annual cost of noncompliance with standards/protocols,  
such as business losses and actual or potential legal fees and/or  
lawsuit awards or settlements.....

**F. TOTAL Current annual costs**

(Enter total of all boxed items in Section I).....

**II. SPEECH RECOGNITION—STARTUP COSTS FOR FIRST YEAR**

**A. Capital costs**

**1. Desktop**

**a. Hardware costs for startup**

- 1. For Pilot Trial.....
- 2. For balance of first-year implementation  
(not including Pilot Trial).....
- 3. TOTAL Desktop Hardware (Total of Lines **II-A-1-a-1** and 2).....

**b. Software costs for startup**

- 1. For Pilot Trial.....
- 2. For balance of first-year implementation  
(not including Pilot Trial).....
- 3. TOTAL Desktop Software (Total of Lines **II-A-1-b-1** and 2).....

**c. TOTAL Desktop capital costs for startup** (Total of Lines **II-A-1-a-3**  
and **II-A-1-b-3**).....

**2. Network costs for startup**

**a. Hardware**

- 1. For Pilot Trial (PCs, servers, routers, drives, etc.).....
- 2. For balance of full implementation  
(not including Pilot Trial).....
- 3. TOTAL Network Hardware (Total of Lines **II-A-2-a-1** and 2).....

**b. Software**

- 1. For Pilot Trial.....
- 2. For full implementation (not including Pilot Trial).....
- 3. Annual maintenance and upgrades.....
- 4. TOTAL Network Software (Total of Lines **II-A-2-b-1**, 2, and 3).....

**c. Network wiring/optical fiber cabling and connectors**

- 1. For Pilot Trial.....
- 2. For balance of full implementation  
(not including Pilot Trial).....
- 3. TOTAL Network wiring (Total of Lines **II-A-2-c-1** and 2).....

**d. Network-related plant construction**

- 1. Required for Pilot Trial (if any).....
- 2. Required for balance of full implementation  
(not including Pilot Trial).....
- 3. TOTAL Construction (Total of Lines **II-A-2-d-1** and 2).....

**e. Miscellaneous** (furniture, office equipment, and supplies).....

- 3. TOTAL Capital network costs for startup**  
(Total of boxed subtotals for Sections **II-A-2-a, b, c, d, and e**).....
- 4. TOTAL Capital costs for startup** (Total of Lines **II-A-1-c** and **II-A-3**).....

**B. Maintenance and Operation costs for startup**

**1. Payroll costs**

**a. Managerial oversight**

- 1. For Pilot Trial (if any).....
- 2. For balance of first year implementation  
(not including Pilot Trial).....
- 3. **TOTAL Managerial oversight for startup**  
(Total of Lines **II-B-1-a-1** and 2).....

**b. Installation**

- 1. For Pilot Trial (if any).....
- 2. For balance of first-year implementation  
(not including Pilot Trial).....
- 3. **TOTAL Installation** (Total of Lines **II-B-1-b-1** and 2).....

**c. Estimated productivity loss during installation and training**

(salary plus value of work not performed)

- 1. During Pilot Trial (if any).....
- 2. During balance of first year (not including Pilot Trial).....
- 3. **TOTAL Estimated productivity loss**  
(Total of Lines **II-B-1-c-1** and 2).....

**d. TOTAL Payroll costs**

(Total of boxed items **II-B-1-a-3, II-B-1-b-3, and II-B-1-c-3**).....

**2. Consulting costs for installation and first-year maintenance**

- a. For hardware installation** .....
- b. For software installation and customization**.....
- c. For user training**.....
- d. For technical support**.....

**e. TOTAL consulting costs for startup**

(Enter total of Lines **II-B-2-a** through **d**).....

**3. Upgrade costs for first year** (salary plus value of work not performed during installation and shakedown).....

**4. Miscellaneous costs during first year**  
(Additional HVAC, electricity, security costs).....

**5. TOTAL Maintenance and operations costs for startup**  
(Enter total of Lines **II-B-1-d, II-B-2-e, II-B-3, and II-B-4**).....

**C. TOTAL Startup costs of speech recognition**

(Enter total of Capital Costs from Line **II-A-4** and M & O on Line **II-B-5**).....

**III. SPEECH RECOGNITION—ANNUAL COSTS AFTER STARTUP YEAR**

**A. Capital costs**

**1. Desktop**

- a. Hardware upgrades and replacements..... \_\_\_\_\_
- b. Software upgrades..... \_\_\_\_\_
- c. **TOTAL Desktop**.....

**2. Network**

- a. Hardware upgrades and replacements exclusively due to speech recognition requirements..... \_\_\_\_\_
- b. Software upgrades exclusively due to speech recognition requirements..... \_\_\_\_\_
- c. **TOTAL Network** (Total of Lines **III-A-2-a** and **b**).....

- 3. Miscellaneous** (furniture, office equipment, and supplies exclusively due to speech recognition requirements).....

- 4. TOTAL Annual capital costs**  
(Total of Lines **III-A-1-c**, **III-A-2-c**, and **III-A-3**).....

**B. Annual maintenance and operation costs**

**1. Payroll costs**

- a. Managerial oversight..... \_\_\_\_\_
- b. Training..... \_\_\_\_\_
- c. Estimated productivity loss during training..... \_\_\_\_\_
- d. Technical support..... \_\_\_\_\_
- e. Proofreading and correction costs..... \_\_\_\_\_
- f. **TOTAL Payroll costs** (Total of Lines **III-B-1-a** through **e**).....

- 2. TOTAL consulting costs** (typically, outsourced training and technical support)....

- 3. Miscellaneous costs** (HVAC, electricity, security costs).....

- 4. TOTAL Annual maintenance and operation costs**  
(Total of Lines **III-B-1** through **3**).....

<p><b>C. TOTAL Annual operating costs after startup year</b> (Total of Lines <b>III-A-4</b> and <b>III-B-4</b>)..... <input style="background-color: #cccccc;" type="text"/></p>
--

**IV. NET ANNUAL SAVINGS (LOSS) FROM IMPLEMENTING SPEECH RECOGNITION TECHNOLOGY**

**A. Total current annual costs**

- (Enter Total from Section **I-F**).....

**B. Costs of speech recognition startup year**

- (Enter Total from Section **II-C**).....

**C. Net savings (loss) of implementing speech recognition in Year 1**

Compare Line **IV-B** with Line **IV-A**.

- 1. If the amount on Line **IV-B** is smaller than the amount on Line **IV-A**, breakeven will occur during the first year.

*Calculate how many months are required to achieve breakeven:*

Divide Line **IV-B** by Line **IV-A** and multiply the result by **12**.

Enter the result here:.....  months

- 2. If Line **IV-B** is larger than Line **IV-A**, go to Line **IV-D**.

**D. Annual speech recognition operating costs after startup year**

(Enter amount from Line **III-C**).....

**E. Net savings (loss) of speech recognition after 2 years**

1. **Current costs.** Multiply Line **IV-A** x **2**. Enter result here.....

2. **Two-year speech recognition costs.** Add Lines **IV-B** and **IV-D**.  
Enter sum here.....

3. Compare Line **IV-E-2** with Line **IV-E-1**:

- a. If Line **IV-E-2** is smaller than Line **IV-E-1**, breakeven will occur during the second year.

*Calculate how many months are required to achieve breakeven:*

Divide the amount on Line **IV-E-2** by the amount on Line **IV-E-1** and multiply the result by **24**.

Enter result here:.....  months

- b. If Line **IV-E-2** is larger than Line **IV-E-1**, go to Line **IV-F**.

**F. Net savings (loss) of speech recognition after 3 years**

1. **Current costs.** Multiply Line **IV-A** x **3**. Enter product here.....

2. **Three-year speech recognition costs.**  
Add Line **IV-B** and 2 x Line **IV-D**. Enter sum here.....

3. Compare Line **IV-F-2** with Line **IV-F-1**:

- a. If Line **IV-F-2** is smaller than Line **IV-F-1**, breakeven will occur during the third year.

*Calculate how many months are required to achieve breakeven:*

Divide the amount on Line **IV-F-2** by the amount on Line **IV-F-1** and multiply the result by **36**.

Enter result here:.....  months

- b. If Line **IV-F-2** is larger than Line **IV-F-1**, continue adding years to annual calculations until breakeven is achieved.