



WebSphere software

IBM WebSphere Voice Server gives voice to e-business applications

Highlights

- ***Can enhance e-business applications with natural voice input and output, allowing access to customers with a wired or wireless telephone***
- ***Extends Web information to virtually anyone with a telephone and includes powerful IBM speech recognition and Text-To-Speech (TTS) engines***
- ***Uses industry standards so you can easily create new speech-enabled applications or add voice interfaces to existing applications***
- ***Supports a variety of connection environments***
- ***Easily integrates with your Interactive Voice Response (IVR) system***

A leading provider of voice enabling e-business solutions, IBM delivers Web, middleware and telephony solutions that can help businesses quickly deliver information to their customers. As your single point of contact, IBM can help you extend your e-business reach by offering integrated hardware, software and services that support the convergence of voice and data. IBM supports open standards-based VoiceXML 2.0 technology that is scalable and highly compatible. Bring your business to the next level with IBM—providing a solid foundation on which to build integrated, innovative voice solutions.

Voice-enable your e-business applications

IBM WebSphere® Voice Server is a software middleware product that provides breakthrough technology for quickly developing and deploying conversational e-business solutions. Voice-enabled applications give your customers, employees and suppliers greater access to information and

services. WebSphere Voice Server supports industry standards such as Java and VoiceXML 2.0. IBM support for open standards gives you freedom from proprietary technology, enabling you to manage costs and application deployment schedules for your business.

Increase the availability of your services

WebSphere Voice Server provides self-service access for your existing customers as well as new ones. It ensures 24x7 support at all times, dramatically extending your business' level of customer service.

Now you can reach a much larger audience—anyone with a telephone. Your customers can access information and conduct transactions, and they don't have to be connected to the Internet to do it. Using a telephone, they can fully interact with your applications and Web site in a much more natural, user-friendly way.

Extend your investments

You have already made significant investments in developing your legacy applications and your e-business framework. By voice-enabling those existing applications, you can extend these same applications and databases to everyone, including customers and employees who may not have had access previously.

Rapid payback

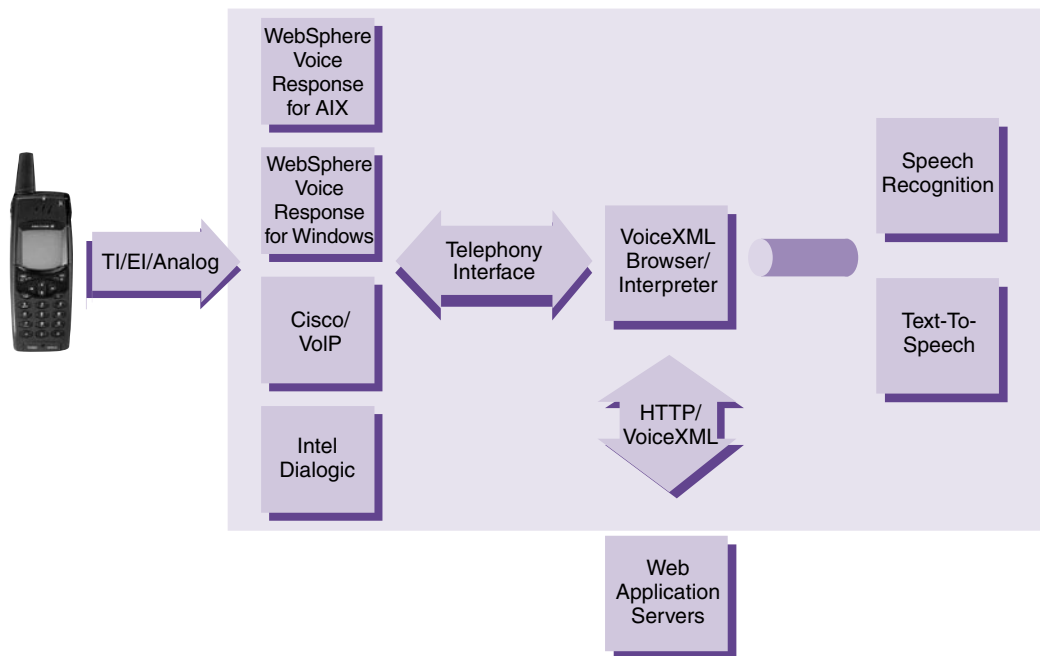
WebSphere Voice Server uses your existing Web and telephony infrastructure, allowing you to reach new customers and provide additional service to existing customers. Speech recognition and

Text-To-Speech (TTS) technology allow you to simplify long menu selections in your existing systems. This means you can reduce the average call duration, cost per call, and decrease call abandonment rates. You can add additional revenue streams by offering new services and use customer service representatives for higher value customer interactions. This allows you to improve customer satisfaction and retention rates. These, and many other factors, can mean a rapid system payback.

Improved usability

Speech recognition and TTS make it easy and natural for your customers to interact with your business. Just as usability is an important factor in traditional applications, it is even more so in mobile applications. Users can access information virtually anywhere, anytime—increasing overall productivity and accelerating communications and responsiveness to customers.

WebSphere Voice Server Connection Environment



Typical voice-enabled Web application scenarios

Today, many applications are used by customers to access information or conduct transactions. WebSphere Voice Server provides the structure for voice-enabled applications, using dialog management which enables the caller to use a natural interface just by speaking. Voice-enabled applications allow customers to call in for a variety of information and transactions, such as the latest weather report, current movie listings, health care provider enrollment, stock information, travel arrangements, employee timecards, appointments and purchases—in short, the entire range of interactive applications.

These four key elements make up the WebSphere Voice Server product offering:

Speech recognition engine—recognizes caller utterances by means of one or more application-specific grammars and converts spoken audio into text.

Text-To-Speech engine—produces an audio stream from text provided by the application for playback over a telephone. There are three TTS offerings available with WebSphere Voice Server:

- *Formant*—the default option, which provides unlimited vocabularies using synthesized voices
- *Concatenative*—an optional feature, which provides unlimited vocabularies using natural-sounding voices
- *Phrase splicing*—an IBM services offering, which provides limited vocabularies using pre-recorded human voices

Voice application development tools

—used to develop and test speech applications using speech recognition and/or TTS. These tools provide support for applications written in VoiceXML and Java.

- *Voice Toolkit for WebSphere Studio* allows application developers to easily add voice technology to middleware applications and includes a VoiceXML editor, grammar editor and pronunciation builder as well as tools for testing VoiceXML applications on a PC without a telephony server.

Telephony platform—the run-time platform for speech applications that connects the voice audio streams from the public telephone network or Voice over Internet Protocol (VoIP) to recognition and TTS engines.

Complimentary products—your WebSphere Voice Server package includes these complimentary products:

- *WebSphere Studio Site Developer V5.0* provides additional tools for a comprehensive Web development tool suite for e-business.

A variety of connection environments and platforms

WebSphere Voice Server offers a wide range of connection environments that can make deployment fast and easy.

Connection environments supported are:

- *WebSphere Voice Response for AIX®*
- *WebSphere Voice Response for Windows® 2000*
- *Cisco telephony platform (supports Voice over IP networks)*
- *Intel® Dialogic Telephony adapters*

Global access

WebSphere Voice Server is available in multiple languages, enabling easy communication around the globe in users' native languages. Languages include Cantonese-Hong Kong, Chinese-Simplified, Chinese-Traditional, Dutch, English-Australian, English-UK, English-US, French, French-Canadian, German, Italian, Japanese, Korean, Portuguese-Brazilian and Spanish. For additional information about language and platform support, go to ibm.com/software/speech/enterprise.

Features

WebSphere Voice Server is used to enable callers to access e-business applications using their voice or the buttons on their telephone. There are a number of features that can enhance your users' experience:

- *Speech-recognition engines, which matches spoken words to a list of possible words in a grammar*
- *TTS engines, which produces synthesized speech data from text provided by the application or stored in databases is shipped with the system (formant is the default — concatenative and phrase splicing TTS are also available)*
- *Support for a variety of telephony-platforms for in-bound and out-bound voice calls*
- *One or more VoiceXML browsers*
- *System Management components, which handles the starting and stopping of browsers*
- *Support for Automatic Number Identification, Dialed Number Identification Service and Call Transfers to ACD or specific agents*
- *Barge-in detection*

Advanced technology converts

Text-To-Speech (TTS)

TTS synthesis provides the ability for machines to convert dynamic text into audible speech. IBM provides concatenative speech synthesis as an optional feature so customers can upgrade the sound quality of the dynamic data played back to a caller on their system. Concatenative TTS uses small pieces of recorded speech to create speech output. This technology sounds more natural because it is comprised from pre-recorded information from both a male and female voice.

The IBM WebSphere Voice Server Concatenative TTS technology offering provides the speech synthesis engine and other components necessary for applications to produce speech output from text input. This offering can be used by solution developers who want to embed the Concatenative TTS into their own speech-enabled solution. Included are the tools, libraries, recordings and application interfaces needed to support concatenative synthesized speech.

Speech technologies

Speech Technologies provides speech recognition technologies and application development tools for developing and implementing speech applications.

Programmers can develop their applications to exploit the following Speech Technologies components:

- *Core speech services, such as signal processing, speech recognition, and speech synthesis (text-to-speech)*
- *A client/server runtime framework, which provides access to the speech services in a client/server environment*
- *A client-side application programming interface (API), enabling the integration of speech-recognition and TTS technologies in new or existing applications*
- *Development and prototyping tools*

Speech Technologies enables prototyping and development on both standalone and client/server systems. Speech Technologies can be used as a standalone package or as an attachment for providing speech services for an existing platform. Speech services provided by Speech Technologies can be implemented as dedicated services or may be allocated dynamically under user control by their applications. Speech Technologies can be used in different modes — as a provider of speech services within a user framework (for example, integration within an IVR).

Natural Language Understanding (NLU)

Natural Language Understanding is the ability of a computer to understand and intelligently react or respond to a query or command from a human in standard language, as opposed to a specially formatted sentence or menu command. People interact with their applications by typing, clicking, speaking or writing; and NLU technology spans all of these input modalities. NLU solutions offer significant benefits to both the enterprise and the end-user by allowing users to do things for themselves (self-service means decreased cost of business) in a manner that's comfortable and pleasant.

NLU combines breakthrough research and development in the fields of voice recognition, linguistics, statistics, human factors and artificial intelligence. There are several types of NLU, including grammar-based and statistical NLU. Grammar-based NLU requires an exact utterance that has a specific match in the system, whereas statistical NLU employs abstract models based on actual conversation, which means a perfect match is not required. Statistical NLU can provide a solid foundation to accommodate increased scalability and application complexity, which can help facilitate future maintenance, expansion and deployments.

IBM NLU systems use statistically-based models which provide more flexibility and robustness than traditional grammar-based methods. IBM NLU eliminates the limitations of speech recognition grammars that are typically used in speech applications today. IBM can provide services to help you realize the benefits of this state-of-the-art technology. Please contact your IBM sales representative for details on NLU.

Custom services

IBM and certified IBM Business Partners offer a complete set of fee-based integration services that can help businesses develop and deploy their voice-enabled applications. There is a fixed-price, Fast-Start program available to help you get up and going quickly.

Custom services include:

- *Requirements workshops*
- *Solution design*
- *Solution implementation*
- *Project planning*
- *Human factors*
- *Prototypes*
- *Proof of concept*
- *Integrated voice solutions*
- *Migration between platforms*
- *Speech recognition and grammar tuning*
- *Enhanced TTS*

WebSphere software platform: building on a firm foundation

WebSphere Voice Server is part of the IBM WebSphere software platform—a comprehensive set of integrated, award-winning e-business solutions. No matter where you are in the e-business cycle, the WebSphere software platform delivers the flexibility you need to grow—at the speed the market demands. Building on this robust platform, you can connect diverse Information Technology (IT) environments to maximize your current investments and leverage existing skills. Deliver your core business applications to the Web using industry standards like Java technology and XML and create next-generation applications that differentiate you from the competition. Advance to a powerful platform for integrated e-business—the WebSphere software platform.

To learn more

For more information about how IBM can help your business take advantage of conversational e-business, call your local IBM Sales Representative, contact an IBM Business Partner specializing in voice at ibm.com/software/voice/partners/list or visit ibm.com/software/pervasive

Hardware requirements

IBM WebSphere Voice Server, minimum requirements

When used with the WebSphere Voice Response for Windows, V3.1 Telephony Platform

- Intel® Pentium® 1GHz processor, or equivalent
- CD-ROM drive
- 512MB RAM minimum required
- 500MB of disk space minimum required per language install
- Local network—10/100Mb fast Ethernet

Note: Requires WebSphere Voice Response for Windows, V3.1

When used with the WebSphere Voice Response for AIX, V3.1 Telephony Platform

- IBM RS/6000 or pSeries™ 223MHz or higher
- CD-ROM drive
- 512MB RAM minimum required
- 500MB of disk space minimum required per language install
- Local network—10/100Mb fast Ethernet

Note: Requires the, separately orderable, WebSphere Voice Response for AIX, V3.1 (with PTF U489834 installed) product. The general WebSphere Voice Response for AIX hardware requirements are identified in the WebSphere Voice Response for AIX: General Information and Planning Guide.

A minimum configuration example is: RS/6000 Systems or pSeries Systems that support the use of the digital trunk adapters (RS/6000 with the IBM ARTIC 960RxD DTXA FC6310 adapter) for connectivity to the telephony networks. Memory and CPU usage are defined in the WebSphere Voice Response publications.

- RS/6000 or pSeries 233MHz or higher
- 512MB RAM required, plus 15MB per Speech Recognition, TTS engine configured and 300MB for any CTTS option except U.S. English, German, French, and Simplified Chinese, which require an additional 600MB
- 2GB Disk drive, or higher
- LAN adapter—10/100Mb fast Ethernet

Voice Toolkit for WebSphere Studio, minimum hardware requirements

- System suitable for running Microsoft Windows 2000
- Minimum of an Intel Pentium 500MHz processor or equivalent (1.0 GHz recommended)
- 768MB RAM
- A display adapter setting of at least 256 colors and 800x600 resolution or higher (1024x768 resolution recommended)
- 175MB free disk space (in addition to the disk space requirement for WebSphere Studio Site Developer V5.1)

Installation requires an additional 200MB temporary space on the drive specified in the user TMP environment variable

Note: The additional hardware requirement is that which is required over and above this to install and run the Concatenative Text-to-Speech for the development language, if so chosen during installation.

If you plan to install IBM WebSphere Concatenative Text-to-Speech, the following additional resources are required:

- 150MB RAM for each CTTS language (excluding Simplified Chinese, which requires an additional 300MB RAM)
- 640MB disk space for each CTTS language (excluding Simplified Chinese, which requires an additional 600MB of Disk space), of which 300MB are for temporary storage

When used with the Cisco or Siemens Hipath Telephony Platform

This connection product remains at the V2.0 level

When used with the Intel Dialogic Telephony Platform

This connection product remains at the V2.0 level

WebSphere Voice Server Speech Technologies for Windows minimum requirements

For all language versions except U.S. English with CTTS this connection product remains at the V2.0 level

Speech Technologies Windows 2.0 with Third Generation U.S. English CTTS Voice minimum requirements

Speech Technologies—Client:

- Intel-compatible 500 MHz, or faster, processor
- CD-ROM drive
- 512 MB RAM
- 200 MB free disk space (minimum)
- Connection to a LAN
- Additional processor speed, RAM, and disk space required by your telephony environment

Speech Technologies—Server:

- Intel-compatible 500 MHz, or faster, processor
- CD-ROM drive
- 1 GB RAM minimum
- 1 GB free disk space (minimum)

IBM WebSphere Voice Server Speech Technologies for Linux minimum requirements (client and/or server)

- Intel Pentium 500 MHz processor, or equivalent
- CD-ROM drive
- 1GB RAM minimum required
- 1GB free disk space
- Local network—10/100 Mb fast Ethernet

Note: 1GB RAM required, plus 15MB RAM per Speech Recognition and TTS engine configured

- 1GB Disk drive, or higher
- LAN adapter—10/100Mb fast Ethernet

Software requirements

WebSphere Voice Server when used with WebSphere Voice Response for Windows, V3.1 Telephony Platform

- Microsoft Windows 2000 Server with Service Pack 3 applied
- IBM WebSphere Voice Response for Windows, V3.1

IBM WebSphere Voice Server when used with WebSphere Voice Response for AIX, V3.1 Telephony Platform

- AIX V5.1 Maintenance Level 4
- Java 2 Technology Edition JRE for AIX, Version 1.4.1
- IBM WebSphere Voice Response for AIX, V3.1 (with PTF U489834 installed)
- Voice XML 2.0 and Java Enhancements for WebSphere Voice Response for AIX installed

Voice Toolkit for WebSphere Studio requirements

- Microsoft Windows 2000 Professional with ServicePack 2, or higher, applied
- WebSphere Studio Site Developer V5.0.1 (or WebSphere Studio Application Developer) installed

Deployment environment for WebSphere Voice Response for AIX and Windows

- PSTN (T1/E1) or SS7 connection for WebSphere Voice Response for AIX product to switch network
- WebSphere Voice Server Client software installed on the WebSphere Voice Response system
- WebSphere Voice Server software and Language software installed on separate pSeries machines for Speech Recognition and TTS capabilities
- Dedicated high-speed LAN for Distributed WebSphere Voice Response systems and WebSphere Voice Server systems (100 Mb Ethernet LAN depending on size of system)
- Web application server hosting VoiceXML application recommended

When used with the Cisco or Siemens Hipath Telephony Platform

This connection product remains at the V2.0 level

When used with the Intel Dialogic Telephony Platform

This connection product remains at the V2.0 level

IBM WebSphere Voice Server Technologies for Windows minimum requirements

This connection product remains at the V2.0 level

IBM WebSphere Voice Server Technologies for Linux minimum requirements

- RedHat 7.3
- SuSe 8.0

Languages

Operating System	Telephony platforms	Languages
AIX	WebSphere Voice Response	Cantonese Chinese-Simplified Dutch English-UK English-US French French-Canadian German Italian Japanese Korean Portuguese-Brazilian Spanish
Windows 2000	WebSphere Voice Response	English-Australian English-UK English-US French German Portuguese-Brazilian Spanish
	Speech Technologies	Chinese-Simplified Chinese-Traditional English-UK English-US (updated TTS voice) French German Japanese
Linux	Speech Technologies	English-US German



© Copyright IBM Corporation 2003

IBM Corporation
8051 Congress Avenue
Boca Raton, Florida 33487

Printed in the United States of America
09-03
All Rights Reserved

IBM, the IBM Logo, AIX, DirectTalk, RS/6000 and WebSphere are trademarks of International Business Machines Corporation in the United States, other countries or both.

Intel and Pentium are Trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Java is a trademark of Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.