



WebSphere software

Incorporating IBM speech solutions into a service oriented architecture.

*By Steve Cawn, Baiju Mandalia and Wendi Nusbickel,
IBM Software Group*

Contents

- 2 Executive overview**
- 3 SOA: An overview**
- 4 SOA patterns for self-service and speech-phone access**
- 6 The IBM speech-solution framework**
- 12 WebSphere Application Server**
- 13 Technologies**
- 17 Development tools**
- 21 An architecture for today and for the future**
- 22 IBM value add**
- 23 Conclusion**
- 23 For more information**

Executive overview

Service oriented architecture (SOA) has gained strong momentum in recent years. As companies begin to understand the benefits of On Demand Business, standards-based development and SOA as a part of product offerings, SOA has shifted from concept to a viable approach to managing disparate IT resources.

IBM has been a leader in embracing the SOA approach, with a comprehensive portfolio of hardware, software and services designed to enable companies to implement SOA. IBM WebSphere® speech solutions further extend IBM's leadership in delivering middleware solutions that integrate with the IT infrastructure. As integral components of the IBM On Demand Business strategy, IBM speech solutions can help companies automate processes so that information can be accessed from practically any connection – through a Web browser, a telephone or a mobile device – and so that transactions can be completed in a highly secure, efficient manner.

The telephone – whether wireline or wireless – is often the handiest, user-friendliest access device. As a result, you can increase access to your business services and applications by speech-enabling existing applications. The voice user interface (VUI) has become a more intuitive, informal and customer-friendly method than ever before – helping to humanize automated interaction. Also, remote data access by voice can enhance employee productivity and accelerate customer service. Improved conversational-access technology can help speed transactions, shorten calls, handle increased call volume and reduce customer service costs.

With more than 40 years of research and product development experience and more than 300 speech patents, IBM delivers breadth and depth of expertise, helping our clients and Business Partners to deploy conversational solutions that take advantage of the most natural way to communicate.

This white paper demonstrates how IBM has used its premier speech technologies, tightly integrated with IBM WebSphere Application Server, to create a standards-based platform for speech-based services through SOA. The discussion begins with an overview of SOA and speech patterns within SOA, and then progresses into the speech-solutions framework. This white paper also addresses how WebSphere Application Server and IBM WebSphere Voice Server can provide your business with a doorway into a new and innovative way of interacting with the SOA framework.

SOA: An overview

SOA is an IT approach that enables your organization to better align IT with your business strategy. Let's start with some of the basics. First of all, what is a *service*? It is a repeatable business task, a part of a business process. Don't think about software or IT. Think about what your company does on a day-to-day basis, and break these business processes into components. Implementing an SOA is much like connecting puzzle pieces (see Figure 1). Services can snap together like individual puzzle pieces. Business processes are a series of services, much like sections of a puzzle, that can also be snapped together. SOA is an architectural style that makes this possible by providing capabilities such as governance, standardization, flexibility and, most of all, making services independent of the underlying technology. Thus, SOA supports integrating a business as linked services by making it easy to connect services into a business processes just like snapping together puzzle pieces to create a complete image.

To learn more about SOA, visit ibm.com/soa. The next section of this white paper discusses IBM speech solutions, what they are and how they fit into SOA.

What is...

<p>... a service? A repeatable business task, such as check customer credit or open new account</p>		<p>... service oriented architecture (SOA)? An IT architectural style that supports integrating your business as linked services</p>
--	--	---

Figure 1. SOA is much like a puzzle, where individual services can be connected to create business processes.

SOA patterns for self-service and speech-phone access

IBM has created a standards-based infrastructure for speech interaction with business applications. This infrastructure includes tools that enable dynamic application development for information and self-service capabilities. Key elements of an application that provides self-service for a customer include clear navigational directions, extended search capabilities and useful links. One popular component is a direct link to online representatives who can answer questions and offer a human interface if needed. The following examples describe how self-service applications can be used across a variety of industries:

- *An insurance company makes policy information available to users and allows them to apply for a policy by automated phone service.*
- *A mortgage company publishes information about its loan policies and loan rates online. Customers can view their current mortgage information, change their payment options or apply for a mortgage by phone.*
- *A bank allows customers to access their accounts and pay bills by phone.*
- *A company allows its employees to view current human-resource policies online. Employees can change their medical plan, tax withholding information and stock-purchase plan by automated phone service.*
- *A company automates a password-reset function.*
- *A company automates call-attendant functions to answer all its telephone communication needs, including telephone calls, faxes and messages.*
- *A company provides facility services (such as where is the nearest office, ATM and so on).*

Figure 2 illustrates a company that wants to implement a new business process to support customers who place orders by automated phone services.

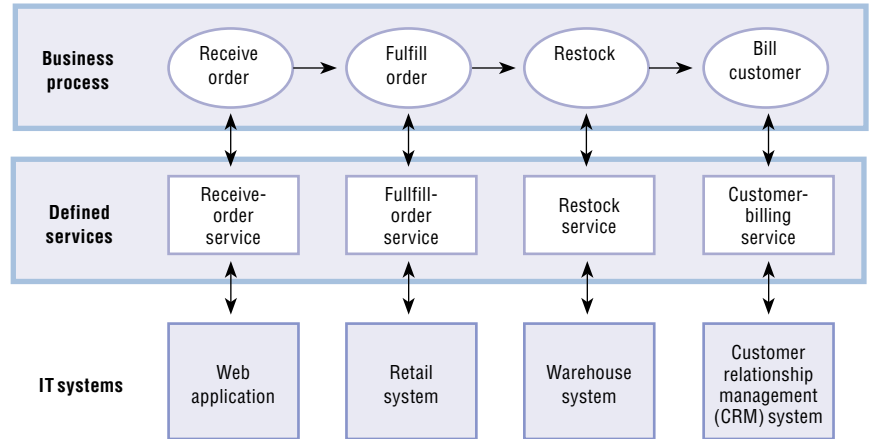


Figure 2. Example of a company's order-placement process

The company already has existing retail, warehouse and billing systems. It would like to build the new process by reusing the functionality that is provided by these systems, rather than having to write new applications or new interfaces to the existing systems. If the company has already taken an SOA approach, it will have defined the interfaces to its existing systems in terms of the functions or services that these interfaces can offer to support the building of business processes. The defined interfaces make building the new Web front end to the system very simple. All the company needs to do is to develop an application that makes calls to the services to complete the new business process.

The SOA approach means that the company can build horizontal business processes that integrate systems, people and processes from across the enterprise quickly and easily in response to changing business needs. As a result, the company can maximize its previous IT investments by reusing existing systems to implement new business processes that extend the use of the systems beyond the processes that they were originally written to support.

IBM offers flexibility to implement an SOA approach by providing SOA-compliant products and services for traditional Web applications and speech solutions.

The IBM speech-solution framework

The convergence of voice and data networks is leading to an integrated infrastructure that can satisfy many business challenges. This new level of integration in contact centers and backbone networks can enable new services and applications to be provisioned easily and with the potential to significantly reduce cost. Companies are upgrading these new applications to provide multichannel access because of the increased use of mobile devices by customers, employees and businesses in the marketplace. The IBM middleware strategy supports this trend through IBM's continued commitment to open standards and common tools in its voice and data products.

IBM clients are challenged every day to cut costs, add new functional capabilities and make self-service systems easier to use while maintaining or improving service, availability and reliability levels. Open standards become an important component to deployment of new services and products. The broad adoption of VoiceXML has provided economies of scale for programming resources and platforms for voice applications and for customer care. Time to market and reuse of skills are critical to a rapid development and deployment of these areas. But IBM standardization goes beyond just the voice and self-service functions of the enterprise.

You must have a robust infrastructure to meet the overall business needs of your enterprise. The IBM software strategy is designed to deliver componentized, standards-based technologies that integrate with the entire portfolio of IBM software offerings. These capabilities enable our clients to reuse their investments through common development tools and consistent application environments across the enterprise. To help ensure that these objectives are achieved, IBM has taken an SOA approach to enterprise solutions (see Figure 3).

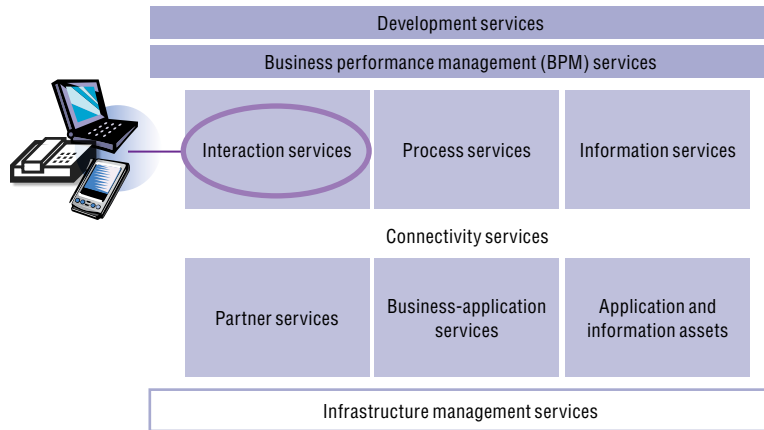


Figure 3. The IBM SOA reference architecture

The IBM SOA reference architecture defines every element needed to build, deploy and manage enterprise applications. Each component is built on industry standards, and communication among components uses standardized communications protocols and message sets. SOA is the blueprint for solutions delivery for all software components across IBM – and it enables interoperability and seamless access to other vendors’ data, provisioned in real time. The IBM speech solutions described in this white paper are integral elements of the IBM SOA reference architecture, serving as the core of the interaction-services component. Interaction services are about connecting users on any device, using the device’s particular interface (such as speech, keyboard, stylus and so on) to a common set of business services, applications, transactions and data.

Integrating these software services using a common set of development tools is critical to lowering the total cost of ownership (TCO) of these business services, especially when extending them to a speech environment. IBM speech middleware products, WebSphere Voice Server and IBM WebSphere Voice Toolkit, provide a holistic approach for building integrated voice and data solutions. These products provide interaction services for telephone-based applications. They are built around the key industry standards, including VoiceXML, Speech Synthesis Markup Language (SSML), Speech Recognition Grammar Specification (SRGS), Java™ and Java 2 Platform, Enterprise Edition (J2EE), TCP/IP, Media Resource Control Protocol (MRCP), Real-time Transport Protocol (RTP) and Call Control XML (CCXML), which are described later in this white paper.

The three-tier application model

As Figure 4 illustrates, consistent with the architecture structures defined by the World Wide Web Consortium (W3C), and further refined in SOA, the basic architecture of the enterprise has evolved to a three-tier model where the business data (Tier 3) is separated from the business logic (Tier 2), which is in turn separated from the rendition of the user interface (Tier 1).

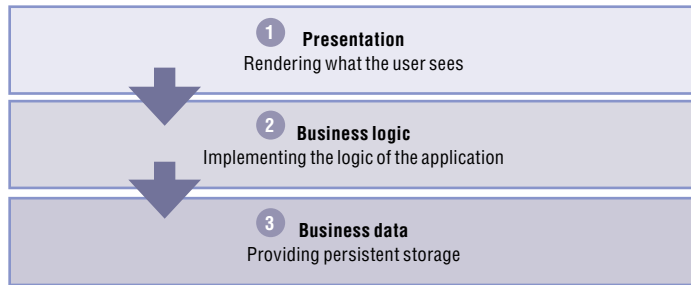


Figure 4. The three-tier architectural model

For example, on the Web, the user interface is defined in HTML and rendered by the Web browser on a PC or similar device. The Web browser thus constitutes Tier 1 of the model and is remotely connected to the Web or application server through the TCP/IP network using HTTP. The Web server, or more commonly, the application server, receives HTTP requests from the Tier 1 devices and implements the application logic associated with the request. The application logic (Tier 2) is commonly built as a J2EE or Microsoft® .NET application and delivers HTML documents that describe the next set of possible interactions with the user back to the Tier 1. As part of the implementation of the application, data is read and written to data stores and databases, which constitute Tier 3 of the model.

The use of VoiceXML encourages the adoption of the three-tier application model where the presentation of the dialog to the caller is separated from the business logic of the application and from the data storage. Defining voice capabilities using VoiceXML enables you to create voice applications and solutions using the same tools and techniques as other access channels such as the Web (using HTML), and Wireless Application Protocol (WAP) and Wireless Markup Language (WML).

Tier 1 provides the VoiceXML interpretation, telephony interfaces and speech resource interfaces. IBM advocates the use of MRCP to connect speech resources to the telephony connection and the VoiceXML environment. To this end, WebSphere Voice Server software supports MRCP and can be used with a number of interactive voice response (IVR) or media gateways. MRCP increases flexibility by enabling you to choose speech resources based on quality, cost, language availability or previous investments with limited impact on your business applications.

The IVR is no longer the deciding factor in speech technology

When developing applications in VoiceXML, the role of the IVR or media gateway is significantly altered. Formerly, choosing an IVR was a major factor in the deployment of any voice-based service, because it determined the application programming model, the tools and skills needs, and the back-end integration requirements. But with the use of VoiceXML, your choices are much more open.

Using VoiceXML, the application runtime environment is a Web application server with all its attendant security capabilities, transaction safety mechanisms and back-end integration infrastructure. By building the application using open standards such as J2EE, you can use a variety of off-the-shelf tools, including tools based on Eclipse (www.eclipse.org) technology like those in the IBM Rational® family of software-development products, to develop and deploy voice-based applications. Furthermore, the use of these standards enables you to build voice solutions alongside their Web and other channel equivalents by people with mainstream skills (such as Java, XML, HTML and so on).

Figure 5 illustrates the major components of a three-tier voice solution.

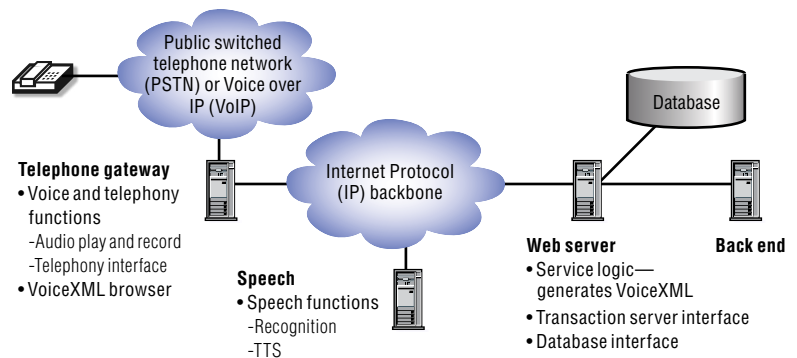


Figure 5. The three-tier voice solution

The components of the three-tier voice solution include:

- Telephony gateway – *IVR platforms from IBM strategic alliance partners or IBM WebSphere Voice Response. Learn more about these partners at ibm.com/software/pervasive/voice_server/ivrgateway.html.*
- Speech server – *WebSphere Voice Server (WVS) on Intel servers running Linux® or Microsoft Windows® 2003, or IBM System p™ machines running the IBM AIX® operating system.*
- Web server for VoiceXML – *WebSphere Application Server or any other Web application server (such as BEA, Tomcat, JBoss or Apache).*

Industry standards play a critical role in any solution, and voice solutions are no exception. The voice solution described in this white paper is built on several key industry standards:

- *VoiceXML, a speech application programming model*
- *CCXML, a call-control programming model*
- *MRCP, an interface between the VoiceXML interpreters and speech servers*
- *SRGS, a speech grammar format*
- *SSML, a semantic representation used by text-to-speech (TTS) systems*
- *Java and J2EE, application programming models for Web applications and applications that render dynamic VoiceXML*
- *Java Management Extensions (JMX)*
- *RTP, an audio interface between telephony gateway and speech servers*
- *TCP/IP and other network protocols*

Using these components with these standards can deliver a cost-effective solution. The ability to extend application functions within these components is provided through the standards compliance of the components and adherence to the standards from an application programming perspective. IBM is committed to all of these standards, from a product-implementation perspective and through its deep involvement in W3C and Internet Engineering Task Force (IETF) committees to develop these specifications and drive them to acceptance status. Since the beginning of the VoiceXML Working Group in 1998, IBM (who was one of the four founding members of this committee) has maintained a director's position in this key standards body. IBM is also a major contributor and voting member to other standards specifications, including CCXML, MRCP, SRGS and XML.

Figure 6 takes a closer look at the specific components that make up this next generation of speech solutions.

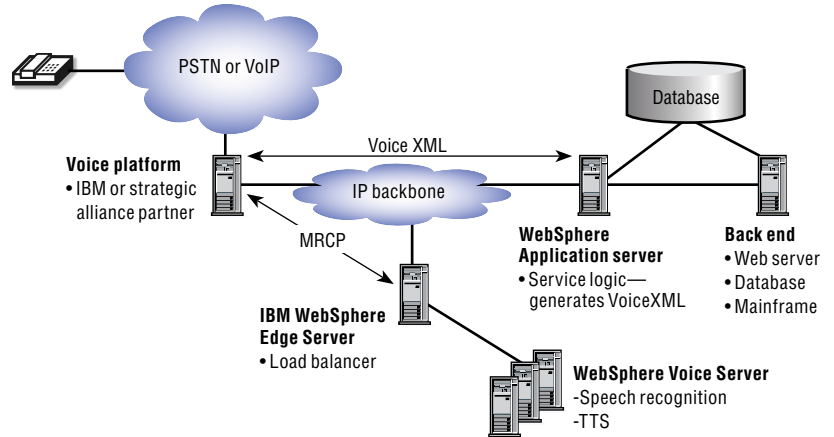


Figure 6. Next-generation speech-solution components

This solution uses a Web-based configuration, load balancing and management structure. When speech requests are made by the VoiceXML application, an MRCP request is initiated by the voice platform. This request is passed to the load-balancing component, which determines which server should process this request. The load balancer then routes the request to the next available server, based on load-balancing options. The server dispatches the request to a WebSphere Voice Server instance that is capable of handling that request. Using this structure means that there isn't any single point of failure within these resources, because you can configure redundant edge load balancers. Also, the availability of an individual speech server is transparent to the VoiceXML application, which enables you to add or remove capacity as needed. The WebSphere Application Server administrative console handles speech-resource management by extending standard JMX functions to enable a single, Web-based management console for the cell.

WebSphere Voice Server management is handled in the same, consistent fashion that WebSphere Application Server is managed and deployed. IBM Tivoli® performance-monitoring tools, provided with WebSphere Application Server, manage the performance of each WebSphere Voice Server cluster. Along the same lines, runtime statistics and data are propagated to a common management repository, which simplifies the management reporting aspects of the installation. By using WebSphere Voice Server in Tier 1 and WebSphere Application Server in Tier 2 and Tier 3, you can take advantage of one management framework to monitor all servers within the same cell.

WebSphere Application Server

WebSphere Voice Server runs as an enterprise application in WebSphere Application Server. This capability extends WebSphere Application Server reliability, scalability and availability to WebSphere Voice Server. With WebSphere Application Server as its foundation, WebSphere Voice Server can take advantage of the following features:

- *Administrative ease of use through the WebSphere Application Server administrative console, to configure, monitor and troubleshoot speech resources.*
- *Continuous access using WebSphere Edge Server, to cluster WebSphere Voice Server machines for load balancing, failover and support serviceability.*
- *Rapid problem determination using the WebSphere Application Server administrative console, which uses graphical alarms and messages to display the current status of a system. WebSphere Voice Server also uses the WebSphere Application Server logging facility so that all trace and log files are kept in a central repository. WebSphere Voice Server extends the WebSphere Application Server collector tool to gather logs and other problem-determination information for IBM support.*
- *Scalability through WebSphere Application Server and IBM WebSphere Application Server Network Deployment manager so that you can add WebSphere Voice Server machines to or remove them from the WebSphere Application Server management cell without having to stop other systems in the cell.*
- *Security that helps enable you protect your WebSphere Voice Server systems by using the WebSphere Application Server global security feature.*

WebSphere Voice Server is not a niche solution. Rather, with WebSphere Voice Server, IBM brings the speech access point into the IT infrastructure through WebSphere Application Server. Along with being a J2EE platform, using WebSphere Application Server as the foundation of your IT infrastructure enables you to get the most value from products like WebSphere Voice Server. The benefits are mutual. WebSphere Voice Server adds MRCP and speech components to your WebSphere Application Server technology-based infrastructure. And you can bring all the functions of WebSphere Application Server to your speech solution. All in a cost-effective manner.

Reducing system-management overhead

Two major factors help reduce system-management overhead in speech solutions: WebSphere Application Server and VoiceXML. Along with being a J2EE platform, WebSphere Application Server provides key functions, including load balancing, failover and recovery, serviceability, management and administration – all on a highly scalable architecture. Using WebSphere Application Server across all three tiers can help simplify system-management tasks. All WebSphere Application Server systems in the network are part of the WebSphere Application Server management cell, and they are all managed by the same WebSphere Application Server administrative interface (console or scripting). As a result, you can now manage your voice server, voice applications, visual applications, business applications and data all using the same WebSphere Application Server management interface – giving you a single way to administer several systems across the entire solution.

VoiceXML uses the Web programming-model paradigm. In the VoiceXML application model, no application resources are deployed to the IVR or media gateways. Instead, all resources, including prerecorded audio clips, speech-recognition grammars and the VoiceXML dialogs themselves, are retrieved as needed from a Web or application server using HTTP. This capability helps reduce your application deployment and systems management challenges. To deploy new or updated services, the systems operator need only publish the updated files to the Web or application server, and each of the IVRs or media gateways picks up the updated files automatically on the next usage. As a result, you can help reduce the need for day-to-day management of IVR or media gateway machines.

Technologies

A number of developments in recent years have resulted in technologies that offer many benefits and can help you expand your self-service, while at the same time improving usability. This section outlines some of these key technologies and shows how you can apply them to the problems of delivering first-class speech self-service solutions.

VoiceXML

VoiceXML, originally created under the auspices of the VoiceXML Forum, but now developed as part of W3C, defines a markup language for the description of machine-to-human dialogs. VoiceXML documents are retrieved by a VoiceXML interpreter, or browser, from a standard HTTP-capable Web or application server. The VoiceXML browser runs directly on, or on a server associated with, some form of telephony communication hardware or software, such as an IVR system or media gateway. A dialog described in VoiceXML can play prerecorded audio or text, using a TTS engine, and can take input as touch tone (also known as Dual Tone Multi-Frequency [DTMF] keys), speech recognition or audio recordings. The VoiceXML specification defines robust, flexible capabilities that cover the majority of the traditional IVR capabilities in terms of dialogs, thus enabling a broad range of applications to be automated using a telephone and your voice as the single input device. Visit www.w3c.org/voice/ to learn more about this technology.

Automatic speech recognition

Speech recognition software can listen to words spoken over a telephone, recognize these words and pass the recognized words as text to an application. This is called *automatic speech recognition (ASR)*. ASR technology matches spoken input to the expected input (grammar) in a speech application. WebSphere Voice Server includes the ASR capability supporting eight languages, and supports a wide range of telephone connection environments (such as speaker phones, cell phones and noisy background environments).

ASR features in WebSphere Voice Server include:

- *Robust speech, noise and silence detection*
- *Very high recognition accuracy with enhanced acoustic models*
- *Support of SRGS grammar external references, weights and repeats*
- *Standards-based lexicon support for customized pronunciations*
- *Fast grammar load time*

TTS

TTS software can convert written text into spoken words and often provides the “voice” of a speech application. WebSphere Voice Server provides fifth-generation synthesis and signal processing algorithms. The WebSphere Voice Server TTS engine supports:

- *SSML*
- *Customized, standards-based pronunciation dictionaries (lexicons)*
- *Multiple, natural-sounding TTS voices*

The WebSphere Voice Server TTS engine also offers several advanced TTS male and female voices. It also uses what is called *concatenative TTS technology*, which divides a large database of actual speech into very small units. These units include words, syllables, phonemes and so on. Multiple representatives (segments) are available for each unit. The WebSphere Voice Server TTS engine selects the best sequence of segments, and splices them together to produce a voice that sounds more natural.

TTS is typically used as part of the system voice in speech applications. It is often used for variable items in a dialog, such as a user address or name, or specific dialog items that are dynamic and thus cannot be predicted by the application developer. The often-preferred approach, in a working speech interaction, is to use TTS segments that can be spliced into prompts spoken by a professional voice talent to create unique messages for each user. For example, the sentence “Your balance is 7513 dollars and 45 cents” can use a recorded voice talent for the static “your balance is” and TTS for “7513 dollars and 45 cents,” because the amount of money in each user’s account is dynamic information and too variable to record professionally in a cost-effective way.

Additionally, both the ASR and TTS engines support a multilanguage capability, which means that more than one ASR and TTS language can be installed on a single WebSphere Voice Server machine. As a result, the same WebSphere Voice Server machine can handle multiple requests for different languages simultaneously.

Natural Language Call Routing

Natural Language Call Routing, also known as Natural Language Understanding (NLU), technology from IBM enables a user to interact with computer applications in much the same way as they would when dealing with a person. IBM NLU technology adds a natural dimension to speech technology by supporting unstructured, conversational dialog rather than requiring the user to speak specific commands. It puts the user in control of the interaction, helping to eliminate long menu options, transactions driven by keypads and commands that are hard to remember – enabling easy, transparent access to information and services.

Natural Language Call Routing is a broad subject in the area of statistical linguistics. IBM has developed and deployed NLU technology for conversational telephony. Recognition of natural-language or free-form speech is handled by the WebSphere Voice Server ASR engine using an application-specific Statistical Language Model (SLM). The language model defines the speech-recognition domain for the system, as well as the potential utterances that might be spoken by the user and recognized by the system. The SLM is built using a body of textual utterances that are domain and application specific and that are gathered through a data-collection process. Systems employing NLU let callers specify their requests in their own words rather than select from a menu of keywords that is typical in a directed dialog system. NLU is available through IBM services to enhance speech solutions.

Conversational biometrics and speaker verification

Conversational biometrics (CB) technology provides a nonintrusive and highly accurate mechanism for authenticating users, based on the analysis of their voice. This technology is called *speaker verification*. Speaker verification uses *voiceprints* to verify users (much like a fingerprint is in the tactile sense).

A voiceprint can be analyzed purely from an acoustic perspective, without considering the content being spoken. In other words, users can say anything because grammar doesn't restrict what they can say. Also, when the voiceprint recognition is implemented in a text-independent mode (for example, no constraints on the words to be spoken or the language spoken), you can gain many other advantages. Users do not have to remember passwords or pass phrases. They also do not have to go through a separate process for verification, because anything they say as part of the transaction dialog can be used to verify their identity. The result is an integrated and nonintrusive verification process.

Verification can take place continuously or periodically in the background as needed (when fraud is suspected in the middle of the dialog, for instance), or at any time after the transaction is completed by analyzing the spoken command for a voiceprint match. Verification can also take place in an incremental manner, and the user can be granted higher privileges if higher verification scores are obtained with more speech data collected as the dialog progresses. IBM offers speaker verification with all the features described in this section as a services offering.

Development tools

Integration of software services and development tools is a very important element in helping to lower the total cost of solution development. IBM has invested considerable effort in developing tools for all aspects of solution development. As part of this effort, IBM contributed the source code that is the basis of the Eclipse project, which provides an open-source infrastructure on which partners, or indeed anybody, can build tools that work together. The IBM Rational framework is an integrated development environment (IDE) that provides a family of tools along with a number of IBM supplied plug-ins, all built on the Eclipse framework, providing development, debugging and deployment tools for the multichannel solution.

IBM's voice tools strategy is built around the tasks associated with voice and speech application development. Major tasks are shown in Figure 7.

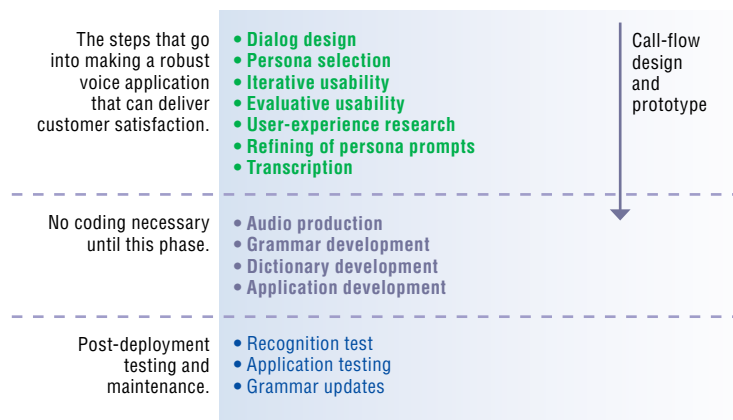


Figure 7. Major tasks involved in voice and speech application development

Understanding how people communicate and want to interact with business services is important when implementing speech applications. Likewise, the tasks associated with call-flow design are critical when developing speech applications. The first set of tasks involve human-factors engineering, which is the science of understanding how people interact and communicate. WebSphere Voice Toolkit is designed with these specific tasks in mind, and provides components specific to these disciplines and tasks. The specific tool for handling the user-interface design is the IBM Communication Flow Builder, shown in Figure 8.

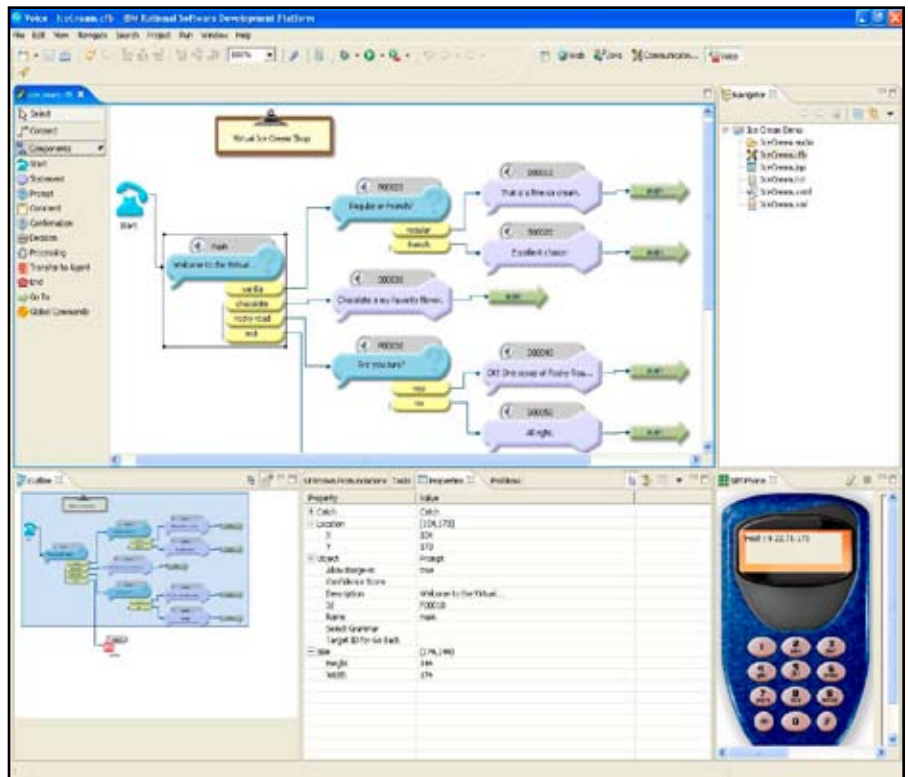


Figure 8. IBM Communication Flow Builder

This tool enables a business analyst or human-factors engineer to quickly prototype a call flow, saving time in the development process.

Figure 9 shows the components of WebSphere Voice Toolkit.



- **Graphical VoiceXML application-development environment**
 - Plugs in to Rational Application Developer
- **VoiceXML editor**
 - Provides a wizard that allows users to select and customize the VoiceXML application
 - Includes reusable dialog components
- **CCXML editor**
 - Enables editing of call-control functions
- **Grammar editor**
 - Supports VoiceXML, Version 2.1 formats
- **Pronunciation builder**
 - Enables you to create and modify pronunciations
- **Precompiled grammar and native pronunciation support**
- **Voice portlet tools**
 - Develops VoiceXML fragments and other voice elements for voice portals or portlets
 - Deploys functions, such as JavaServer Pages (JSP) debugging from portal tools
- **GUI-based application debugging**
- **Audio recorder**
- **Analysis capabilities**
 - Provides call logs
 - Allows you to analyze audio quality
- **National Language Support (NLS)**
 - Lets you develop applications in the languages the voice middleware supports

Figure 9. Components of WebSphere Voice Toolkit

The components of WebSphere Voice Toolkit include:

- *A VoiceXML, Version 2.1 technology-compliant editor*
- *A CCXML editor*
- *A grammar editor*
- *A pronunciation editor that enables you to interactively create custom pronunciations for both speech recognition and TTS conversion*
- *A VoiceXML and MRCP simulator to assist and simplify the functional application testing*
- *Trace tools*
- *Debug tools*
- *The ability to use an open-source family of reusable dialog components that make up the VoiceXML and grammars for many commonly used functions.*

Dialog flows can be built and tested along with the associated speech-recognition grammars and TTS pronunciations. By using the parts of the voice toolkit in concert with other Rational components, you can build the voice dialogs as a tightly integrated part of a multichannel solution. You can use components such as XML editors, HTML editors, and Java and J2EE editors to generate an integrated set of applications for the enterprise. Additionally, tools from IVR partners, other tools developers, systems integrators and open-source organizations can be integrated into WebSphere Voice Toolkit through customizable document type definition (DTD) and the development of Eclipse plug-ins. This feature enables you to take advantage of the unique value provided by these partners.

WebSphere Voice Toolkit also provides a powerful GUI called the *voice trace analyzer*. WebSphere Voice Server allows you to enable a lightweight trace for this tool that can be collected and imported into the voice trace analyzer at any time. The voice trace analyzer enables application developers to quickly troubleshoot or tune applications whether during application development, pilots or production (see Figure 10).

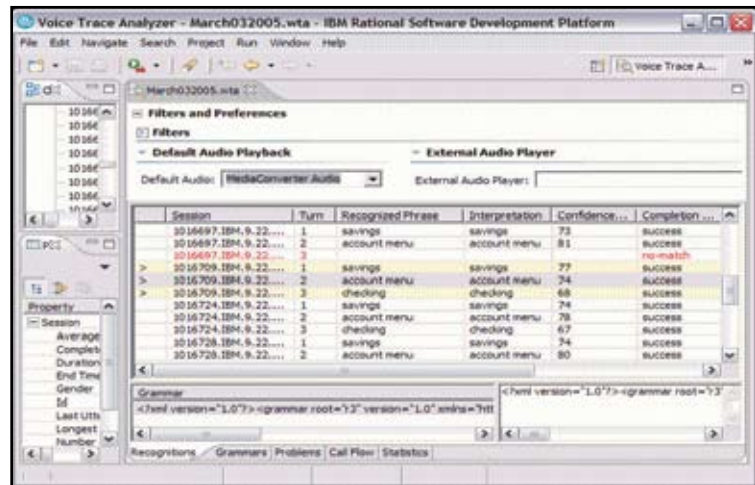


Figure 10. Voice trace analyzer

As the speech market matures, many other voice-development tools are being developed and sold by other vendors. IBM has expanded its tools offerings to make its speech tools accessible to other vendors. This new component strategy gives you more flexibility to select application-development tools independent of speech tools. IBM speech tools support WebSphere Voice Server, and can be used to develop speech applications for speech servers that are compliant with VoiceXML.

An architecture for today and for the future

With all of these functionally rich capabilities, IBM has just begun to embark on a strategy for robust speech self-service solutions that is designed to help ensure interoperability and application compatibility today and into the future. The core of this framework is the extensive use of IBM’s strategic middleware platform, WebSphere Application Server (see Figure 11). As stated previously, this framework provides a robust foundation for highly available, scalable solutions that meet the needs of On Demand Business. And because WebSphere Application Server is based on industry standards, you can increase flexibility and interoperability for the multiple components involved in these solutions.

Using a common framework improves the interoperability and management of these solutions. The network connection is provided by IBM vendor partners, such as Genesys, Avaya, Cisco or VoiceGenie, which enables multiple network connections, helping to improve the flexibility of the self-service environment.

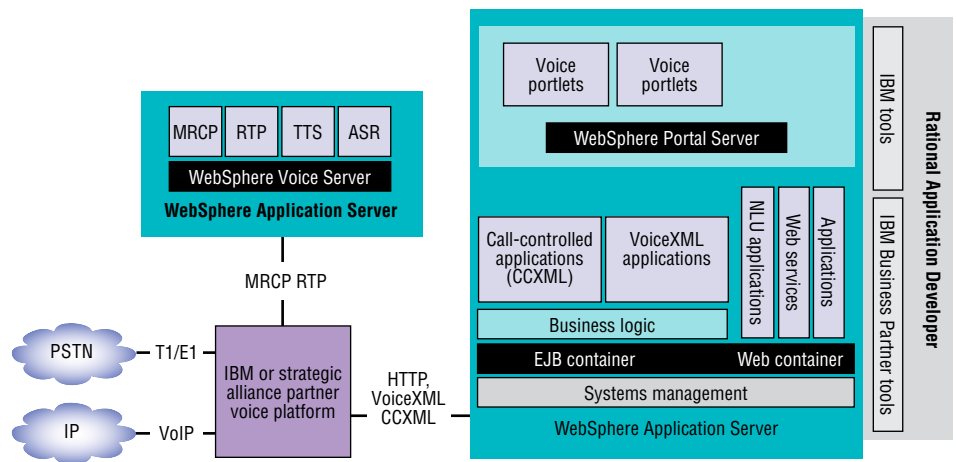


Figure 11. The IBM voice-solution architecture

This feature-rich platform also enables you to deploy new classes of self-service applications. Using the framework provided by WebSphere Application Server and WebSphere Portal Server fully enables applications that are both Web and portal based. And a broad set of standards-based applications capabilities are all managed under a common systems management infrastructure and defined by an SOA framework.

Finally, note the common development framework for all of these capabilities. A key part of IBM's short- and long-term strategy is to provide the most-robust experience for the application developer, helping to ensure that all necessary tools are integrated within a single development environment. This capability can improve programmer productivity, enabling the programmer to focus on installation-specific tasks, and use common components that are available within the installation or from the industry.

WebSphere Voice Server, Version 5 was the first step IBM took to make these capabilities a reality. IBM continues to evolve this framework in a carefully planned, staged set of deliveries, helping to ensure that upgrading to new features and capabilities is a smooth process as the products and standards evolve, helping to enable maximum value and investment protection.

IBM is also addressing the short- and long-term needs for self-service functions within your enterprise. Based on solid technology, the standards-based IBM solution provides a robust, dynamic infrastructure that can deliver high-quality speech solutions for a broad range of applications and users.

IBM value add

IBM is positioned to deliver this type of solution for the following reasons:

- *Innovation. IBM uses its extensive research, experts and leading practices, along with its extensive vendor alliances to deliver short- and long-term innovations. With more than 100 speech researchers alone, IBM's ability to innovate is unmatched by any other speech vendor today.*
- *Accountability. IBM can perform the prime contractor role for this service and would therefore be accountable to manage the implementation of the project and the subcontractors to completion, along with all related requirements and delivery commitments.*
- *Service and delivery flexibility. An IBM solution can scale rapidly in accordance with your evolving business needs. IBM can provide a detailed upgrade path that enables you to move from your current IVR infrastructure to a new open-standards-based solution.*
- *Open platform and standards, best-of-breed vendors. IBM's open solution includes sophisticated speech technology and integration with many other components provided by industry leaders.*
- *Proven, large-scale transformation experience. IBM has a proven track record with large-scale contact-center transformation.*

Conclusion

IBM speech solutions with WebSphere Application Server implemented as part of an SOA approach can provide the right step for the development and deployment of speech self-service solutions. SOA provides the integrated, open-standards-based environment to help organizations meet the complex challenges with quick response to new business requirements and aids in reducing the cost of IT to the business.

Speech technology is going through an inflection point. It has matured, enabling users to interact with enterprise self-service applications in a natural way. You can use speech-based self-service to differentiate your business, and as a catalyst for potential cost savings in your contact center. An WebSphere Application Server technology-based speech offering using WebSphere Voice Server is an industry-leading solution that brings the robust management and reliability capabilities to speech that are already offered by IBM today for visual business access (through a Web browser).

IBM is fully committed to speech as a core element of its SOA strategy. As infrastructures for self-service applications converge for data and voice, IBM solutions can provide rapid enablement of new services.

For more information

To learn more about IBM speech solutions, contact your IBM representative or IBM Business Partner, or visit:

ibm.com/pvc

To learn more about SOA, contact your IBM representative or IBM Business Partner, or visit:

ibm.com/soa



© Copyright IBM Corporation 2006

IBM Corporation
8051 Congress Avenue
Boca Raton, Florida 33487
U.S.A.

Produced in the United States of America
06-06
All Rights Reserved

AIX, IBM, the IBM logo, the On Demand Business logo, Rational, System p, Tivoli and WebSphere are trademarks of International Business Machines Corporation in the United States, other countries or both.

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

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

Linux is a trademark of Linus Torvalds in the United States, other countries or both.

This publication provides addresses to non-IBM Web sites. IBM is not responsible for the content on such sites.



G224-7605-00