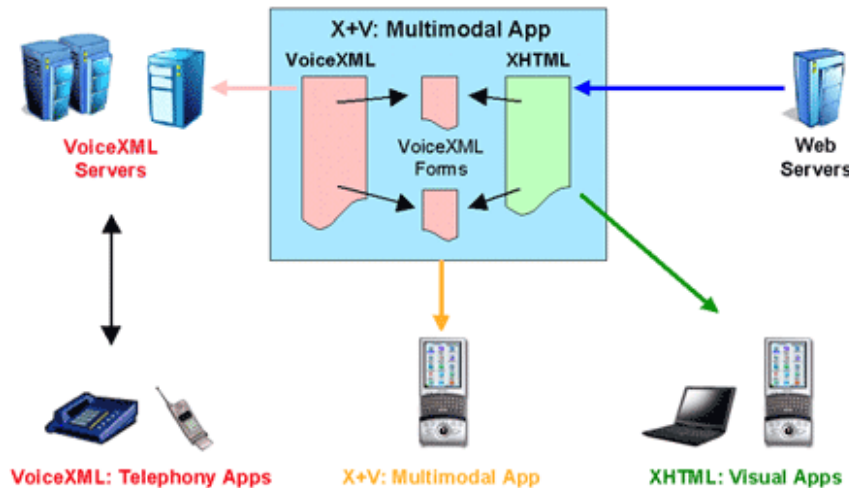


IBM Embedded ViaVoice, Multimodal Browsers and the Multimodal Toolkit for WebSphere Studio.



interface enables users to obtain and manage information as the situation dictates—whether through speech or visual displays—anytime, anyplace. Now a developer can deliver a graphical user interface (GUI) and speech recognition in a single application.

Single, fully integrated architecture

The modular Embedded ViaVoice architecture provides fully-integrated speech recognition and speech synthesis engines supporting the full feature requirements of an application with minimal central processing unit (CPU) utilization.

A single architecture with consistent application interfaces allows Embedded ViaVoice to support solutions from low resource PDAs through high performance in-car solutions to Java-based or markup language-based connected solutions. This single architecture implementation is a particular advantage to applications that need to span a broad range of platform capacities as well as solutions where significant growth in capacity is a requirement.

Highlights

- **Allows fully integrated speech recognition and text-to-speech on one device**
- **Supports speech recognition vocabulary lists of up to 100,000 words in real time, and in a broad range of languages**
- **Supports different scalable programming models, Java™ clients and multimodal browsers**
- **Helps to minimize the skills and time needed to develop state-of-the-art applications for personal digital assistants (PDAs), automotive navigation systems, hand-free phones and other handheld, wireless devices**

Simplifying communications

As computing continues to extend from PCs onto devices, the business and consumer worlds are expecting more from devices. Users want more natural ways of communicating with machines or with other people via machines.

Together, IBM Embedded ViaVoice® Version 4.2 and the Multimodal Toolkit for WebSphere® Studio allow developers to use fully-integrated automatic speech recognition (ASR) and text-to-speech (TTS) on one device—from set-top boxes, to automotive navigation systems, hands-free phones, PDAs and other small devices. The resulting

Broad language base

Embedded ViaVoice is available in a broad set of languages both for speech recognition and speech synthesis through the support of a worldwide network of IBM speech development laboratories.

With the latest release of Embedded ViaVoice, high quality embedded concatenative text-to-speech (eCTTS) was introduced to provide more human sounding speech synthesis for more advanced applications. Development of additional language models and voices as well as the perpetual improvement of existing languages continues. For the most current language plans, please contact your local sales representative.

High recognition accuracy

The Embedded ViaVoice recognition engine is phoneme-based using finite state grammars to support continuous speech recognition. Through a comprehensive research and development effort, IBM has reduced the word error rate of Embedded ViaVoice by an average of 35 percent annually over the last three years.

Large vocabulary recognition

The maximum supportable vocabulary has grown by a factor of 25 over the last four years. The most recent version of Embedded ViaVoice supports recognition of lists of up to 100,000 words in real time — allowing, for example, unconstrained recognition of any street in California or any place name within Germany.

Multiple programming models

Embedded ViaVoice base ASR and TTS engines support several different scalable programming models. Many small footprint embedded applications can use Embedded ViaVoice through its C/C++ language application interfaces. PDA voice personal information management (PIM) enablement and voice phone dialers require extremely low footprints.

In car navigation, hands-free phone and personal assistance services require more robust systems. Alternatively, Java applications can be written to the standard Java Speech Application Programming Interface (JSAPI) to perform recognition and synthesis as a component of a standards-based connected system. Automobile information retrieval systems are an example of these connected solutions.

Finally, the Multimodal Toolkit for WebSphere Studio, combined with the Multimodal Browser, uses the W3C proposed XHTML+Voice (X+V) standard and Embedded ViaVoice speech recognition and speech synthesis to enable both visual and voice access to Web

applications. Just like visual Web applications, Multimodal applications are developed once and can be rendered on multiple devices such as PDAs, cell phones, laptops, kiosks, or custom (in-car) information systems. Examples of applications include financial account management, in-car navigation systems, personal name dialers and warehouse inventory management.

Supports Java clients

Embedded ViaVoice supports Java clients through the standard Java Speech Programming Interface. This latest release of Embedded ViaVoice includes several key extensions of the JSAPI standard.

Multimodal Web Browsers supported

Multimodal applications using the X+V proposed standard offer a natural evolution from today's Extensible Hypertext Markup Language (XHTML) applications and Voice Extensible Markup Language (VoiceXML) applications. Multimodal Browsers have been developed in strategic relationships with Opera Software (Opera Browser V7.3) and ACCESS Systems (NetFront Browser V3.1).

Multimodal Toolkit for WebSphere Studio uses the X+V proposed standard and Embedded ViaVoice speech recognition and speech synthesis to enable both visual and voice access to Web applications.



The Multimodal Web Browser rendering a multimodal application

An example is an airline flight reservation Web page that can be filled in with a combination of voice, keyboard, and stylus. The latest version of the toolkit and browser supports mixed initiative input to a Web page, which allows a user in one utterance to provide some or all the data required in any order. The application can then use directed input for any items still required. This latest version also speech-enables the browser's own navigation functions (e.g. "browser go home"). The current Multimodal Browsers support Pocket PC and Embedix clients.

Developer tools

The Embedded Voice Toolkit and the Multimodal Toolkit each add extensions to Eclipse-powered WebSphere Studio that provide multimodal and embedded voice functionality through a user interface that can minimize both the skills and time needed to develop state-of-the-art applications for in-hand and in-vehicle devices.

The tools provide an integrated development environment that lets you integrate visual and voice applications efficiently without requiring expertise in all the development languages. The toolkits provide multiple tools, editors, and views that are operated using standard menus, icons, toolbars, and basic grammar, XHTML and VoiceXML programming.

Key components

- *Java and C/C++ IDEs for developing applications*
- *Multimodal X+V Editor for creating XHTML and VoiceXML in the same application*
- *Grammar Editors including syntax checking and automatic grammar generation*
- *Pronunciation Builder to create/modify pronunciations as well as to easily generate pronunciations from audio*
- *Reusable Dialog Components (RDC) and Wizards to assist the developer in creating applications*
- *Application deployment, audio recording, grammar testing and analysis tools*
- *Multimodal Browser Launcher to test your applications on the desktop*

Services and workshops

Porting and Integration Services includes porting to a new operating system, recompilation for a different processor architecture, or modification of the embedded audio layer to use a new driver or codec. Alternatively with the Audio Adaptation Kit, we supply the tools for the customer to perform and test the audio adaptation themselves.

Five day classes for application developers are available for the Embedded ViaVoice Software Development Kit (SDK). In addition, customized development workshops provide skills transfer and instruction on application development, evaluation methodology and tools so customers can design and tune their system.

IBM usability experts can provide reviews of user interface design, vocabulary and grammar optimization (including creation and testing of alternative grammars) and execution of voice recognition accuracy tests. Using a state-of-the-art audio studio, consultation is also available on selection and placement of microphone and other components for optimizing the audio signal. Assistance is available in application design, implementation and testing, specialized tool creation, and creation of validation plans and programs. Support is available for your technical interactions with other solution partners and with your customers.

IBM expertise in voice

IBM's sustained 30-year research and development investment in speech recognition and synthesis has resulted in multiple product offerings including Embedded ViaVoice and the Multimodal Toolkit for WebSphere Studio. For today's embedded applications, highly functional speech-based or multimodal systems can be developed using these industry-leading products.

IBM provides products for a wide range of voice-enabled applications through multiple programming models including C/C+, Java and X+V. IBM Embedded ViaVoice products enable customers to have a true competitive advantage in today's fast moving marketplace as well as a clear path for future growth through a single, fully integrated architecture.

For more information

Please contact your local IBM sales representative for more information or a demonstration of our embedded and multimodal products. Or visit our Web site at:

ibm.com/software/pervasive/
multimodal



© Copyright IBM Corporation 2004

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

Printed in the United States of America
05-04
All Rights Reserved

IBM, the IBM Logo, e-business on demand, "e (logo) business on demand" lock-up, ViaVoice and WebSphere are trademarks of International Business Machines 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.

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.



Printed in the United States on recycled paper containing 10% recovered post-consumer fiber.

