

Telephony Call Control Now Available for HTML, XHTML, and XML

Thom Sawicki
Sr. Manager Business Development
Intel Labs
Intel Corporation

Kamal Shah
Engineering Program Manager
Intel Labs
Intel Corporation

Table of Contents

(Click on page number to jump to sections)

TELEPHONY CALL CONTROL NOW AVAILABLE FOR HTML, XHTML, AND XML	3
OVERVIEW	3
WHY SALT?	3
TELEPHONY	3
CALL CONTROL	4
A MULTIMODAL PC	4
SUMMARY	4
MORE INFO	5
AUTHOR BIOS	5

DISCLAIMER: THE MATERIALS ARE PROVIDED "AS IS" WITHOUT ANY EXPRESS OR IMPLIED WARRANTY OF ANY KIND INCLUDING WARRANTIES OF MERCHANTABILITY, NONINFRINGEMENT OF INTELLECTUAL PROPERTY, OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT SHALL INTEL OR ITS SUPPLIERS BE LIABLE FOR ANY DAMAGES WHATSOEVER (INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF PROFITS, BUSINESS INTERRUPTION, LOSS OF INFORMATION) ARISING OUT OF THE USE OF OR INABILITY TO USE THE MATERIALS, EVEN IF INTEL HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. BECAUSE SOME JURISDICTIONS PROHIBIT THE EXCLUSION OR LIMITATION OF LIABILITY FOR CONSEQUENTIAL OR INCIDENTAL DAMAGES, THE ABOVE LIMITATION MAY NOT APPLY TO YOU. INTEL FURTHER DOES NOT WARRANT THE ACCURACY OR COMPLETENESS OF THE INFORMATION, TEXT, GRAPHICS, LINKS OR OTHER ITEMS CONTAINED WITHIN THESE MATERIALS. INTEL MAY MAKE CHANGES TO THESE MATERIALS, OR TO THE PRODUCTS DESCRIBED THEREIN, AT ANY TIME WITHOUT NOTICE. INTEL MAKES NO COMMITMENT TO UPDATE THE MATERIALS.

Note: Intel does not control the content on other company's Web sites or endorse other companies supplying products or services. Any links that take you off of Intel's Web site are provided for your convenience.

Telephony Call Control Now Available for HTML, XHTML, and XML

Thom Sawicki
Sr. Manager Business Development
Intel Labs
Intel Corporation

Kamal Shah
Engineering Program Manager
Intel Labs
Intel Corporation

Overview

New SALT spices up the Web—literally. SALT (speech application language tags) lets developers add speech and telephony call-control features to existing and new Web-based applications. These exciting, new features are the next step in making the computing interface more human-friendly or natural. With SALT, users can use speech (instead of the keyboard and mouse or stylus) to access information online, make calls through their computer, order products, and so on. Speech makes using the Web more convenient and accessible through a wide-variety of devices including desktops, laptops, PDAs, and phones.

SALT does not require developers to rewrite or rebuild an existing application in a new markup language. Instead, SALT is an extension of existing standards, offering developers a set of new, powerful HTML tags. These tags let developers seamlessly embed speech enhancements in existing HTML, XHTML, and XML pages. The 10 new tags are grouped for use as:

- Speech Output or TTS: <prompt>, which may contain <value>, <content>, and <param>
- Speech Input or ASR: <listen>, which may contain <grammar>, <bind>, <record>, and <param>
- Touchtone Input or Keypress: <dtmf>, which may contain <grammar>, <bind>, and <param>
- Messaging: <smex>, which may contain <bind> and <param>

With SALT, developers can continue using languages, technologies and toolkits they have been familiar with for years. And, with Intel's SALT Call Control Technology Preview that includes open-source libraries and documentation, developers can start experimenting with SALT now. The SALT specification will make both multimodal and telephony-enabled applications and services faster and easier to create, deploy, and use.

Why SALT?

Telephony call-control features have been in use for several years. However, today's telephony call-control applications are unique and custom-designed. There is no standard set of tools for developers to use to extend an existing Web page to include or create a new one that has telephony call-control features.

SALT solves this problem by extending HTML, the most common Web markup language in use around the world. SALT defines a set of speech related tags for HTML, an ECMAScript call-control object, as well as a scripting methodology. Instead of re-inventing the wheel for each company's application, developers can now implement speech-I/O and call-control features from their existing work model. In short, SALT gives Web developers a vehicle to develop the next generation of Web communication.

Telephony

Case studies have shown that four times as many phone callers are opting for voice entry when such features are available from a business (Telephony Voice User Interface Conference, February 7, 2001.) However, even though there are 1.6 billion phones in the world, only a small fraction of Web applications are reachable by phone. SALT lets developers include voice input to Web sites via wireless devices, speech-enabled applications, and phones.

For example, with a SALT-enabled Web site, employees could talk to their online catalog to get information for customers, instead of clicking and scrolling through the pages. In a car, one could use the speech-I/O feature of his wireless PDA (personal digital assistant) to locate the nearest store and get directions. A shopper could receive an automatic cell-phone call that the custom configuration he ordered an hour ago is now ready for pick-up. With SALT, end users can use speech to search online catalogs, make airplane reservations, buy tickets for an upcoming concert, and so on.

SALT telephony features also help resolve one of the major issues with a truly worldwide Web: the use of different alphabets and ideographs. SALT-based applications are very compelling for environments where keyboard-based I/O is inherently complex. For example, a Russian who uses the Cyrillic alphabet might be offered only an Arabic character-set keyboard. In this instance, he or she could use speech-I/O features to complete a Web site application.

With today's computers, the types of input and output that are available are limited. Telephony features remove limitations so that users can easily access and use Web applications from anywhere.

Call Control

SALT lets e-Businesses offer telephony call-control features to employees and customers. These features let a Web site establish phone calls, take calls, put a customer on hold, establish a conference call, and more. With SALT, developers can extend existing e-applications to link call centers, service functions, and more. Call-control takes the PC from being a passive interface into the realm of proactive computing.

For example, before leaving home for an airline flight, many people check the flight's departure times by phone or computer to make sure the airplane is on time. With call control, an airline's Web application can automatically let passengers know whether the flight is on time or will be delayed, and for how long. Notifications are delivered via each passenger's choice of communication device—wired or wireless telephone, voice-enabled pocket/laptop/tablet/desktop PC or PDA—and passenger's desired modality—text, speech, graphics or a combination of modes. Another example: A catalog customer might be trying to place an order for a product that doesn't seem to be available. While still online at the catalog site, the user could place a call to customer service, via the catalog application to try to resolve the problem.

A Multimodal PC

In today's world, the development of Web applications through standard markup languages (such as HTML) is limited to text and graphical elements. Unfortunately, the visual orientation of today's Web applications makes it hard for users to interact with mobile devices that have small displays and keypads. SALT resolves this problem by enabling natural, multimodal (speech and visual) access to the Web.

With SALT, tasks can be easily switched from keyboard input/output (I/O) to speech I/O. For a mobile work force, this is an immensely valuable capability. For example, when the keyboard or stylus isn't effective or appropriate (such as in a car or walking), the user could switch to speech I/O. The user then simply talks to his computer to interact with the e-application. This is the realm of ubiquitous computing, in which users can access information-based services anytime, anywhere. It means increasing both efficiency and the cost-effectiveness of e-applications. For example, in one case study, a company that added speech features to its Web site realized an 84 percent cost reduction in customer service, and a drop of more than 50 percent in customer complaints (Telephony Voice User Interface Conference, February 7, 2001.)

With SALT, multimodal access can include input through speech recognition, keyboard, keypad (DTMF), stylus, and/or mouse. Output can be provided in the form of recorded audio, synthesized speech (text-to-speech), on-screen text, graphics, and/or motion video. All input and output modes may be scripted to occur independently or concurrently.

Summary

Voice access to the Web is currently limited to a telephony-centric interface. And, where voice access has been achieved, it has been via separate development methods and with tools that have unique scripting languages or proprietary development environments.

In contrast, SALT eliminates the need for custom languages and unique application development tools. SALT is a royalty-free, platform-independent specification that extends existing HTML, XHTML, and XML Web markup

languages. By adding new tags for speech to the HTML language along with a call-control object, SALT allows programmers to use common tools to add speech I/O and call-control features to new and existing Web sites. In essence, SALT marries speech and the Web together.

A preliminary SALT platform and set of tools for speech features are already available from Microsoft. With the preview collateral that is now available from Intel, developers can start experimenting with speech I/O and call-control. Make sure to look at the SALT specification to see how SALT can enhance and extend your Web applications and services. By checking out the standard now, developers can get a head start on business and technology plans to implement speech I/O and call-control in their Web projects.

More Info

The new SALT tags and the ECMAScript call-control object methodology are described in the SALT specification. The most recent draft of the specification is available at the SALT Forum Web site. The draft specification offers a preview of what SALT can do for developers, businesses, service providers, and end users.

The "Speech Application Language Tags (SALT) Technical White Paper" is also available online at the forum's site. The white paper describes design principles, the three main (top-level) SALT elements, event handling, telephony capabilities for call-control, and the flow of dialog in voice-only scenarios.

Intel's "SALT Call Control Technology Preview" is now available on the Intel® Developer Services site. The preview includes open-source libraries, example code, tools, and documentation for developers to learn more about SALT Call Control and develop proof-of-concept pages. Microsoft's beta-release SALT tools are also currently available at Microsoft's speech.NET Web site.

Author Bios

Thom Sawicki has been with Intel for 20 years. He is currently a senior business development manager for Intel Labs. Thom has worked in a variety of roles at Intel including sales, marketing, and strategic investments. Today, he defines and implements marketing programs and industry ecosystems for next-generation telephony, speech, and high-speed network technologies. Thom has also been active in the SALT Forum since its formation in 2001, and is currently Intel's marketing representative and board liaison. Thom received his B.S. in computer science from Oregon State University, and an M.S. in science and technology management from the Oregon Graduate Institute.

Kamal Shah has been with Intel for 18 years and is currently an engineering program manager at Intel Labs. He led the Intel team responsible for the technical contributions to SALT and W3C CCXML (Call-Control XML, World Wide Web Consortium). His work is focused on Internet, Voice and High-Speed Networking initiatives. Kamal has managed a variety of operating-system and IP-based, multimedia communication projects. Specifically, he was responsible for the definition, successful launch, and deployment of an innovative IP telephony-service business initiative. Kamal received his B.S.E.E. from Gujarat University, India. He received his M.S. in computer engineering from the University of Texas at Austin. In March 2002, he delivered an industry perspective session on behalf of Intel at the "Spring Voice on the Net" conference. Kamal has received numerous divisional recognition awards and is co-author on two patents pending.

—End of Intel Developer Update Magazine Article—