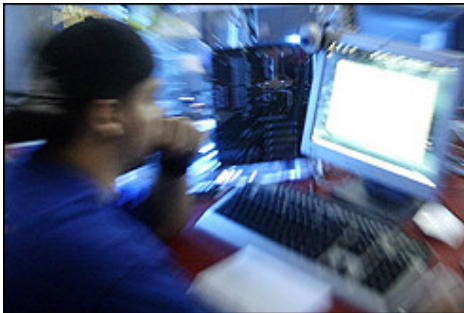


W3C looks to improve speech recognition technology for web transactions



W3C, the standards-setting body for the Internet (World Wide Web Consortium), has completed a draft for the important VoiceXML 3.0 - technology enabling voice identification verification. While normally associated with voice commands, it has the potential to greatly speed and improve the accuracy and positive proof of online transactions.

Some larger net businesses are even using it to confirm orders and verify identity. Many, however, have become increasingly worried about the reliability and security of these transactions with fraud and identity theft on the rise. Error rates have been around 1 to 2% - unacceptable for ironclad business transactions.

W3C does not actually make software but produces standards. They now have a working draft, said James Larson, co-chair of the W3C VBWG - Voice Browser Working Group.

The standard also addressed the issue of extending its Speech Synthesis Markup Language (SSML) functionality to certain languages including Mandarin Chinese, Japanese and Korean.

SSML is important because it allows software makers to control speech from pitch to volume to pronunciation. This insures the software will hear the right tones and pitches so critical in languages where a tiny change in pronunciation can affect the whole meaning of a word.

SSML is also used to tag areas of speech with different regional pronunciations. It is based on JSpeech Grammar Format (JSGF).

A technical description and how to use SSML version 1 on a web page can be found here:

<http://www.xml.com/pub/a/2004/10/20/ssml.html>

Microsoft Agent website is another source for would be speech interface developers.

<http://www.microsoft.com/MSAGENT/downloads/user.asp>

Opera browsers can be programmed for speech recognition with some XHTML (Extended Hypertext Markup Language) extensions. <http://my.opera.com/community/dev/voice/>

Working with web-based speech applications can be frustrating. While the speech recognition software works well, poor quality microphones and PC speakers combined with slower Internet connections can put a damper on effectiveness. These issues will be difficult to address due to being largely beyond the control of the developer. New speech compression algorithms and simple responses like yes or no make the job much easier.

Trained systems – ones that are accustomed to the user's voice – have been much more successful, but users typically do not have the patience to complete the training and the time factor makes it impractical.

Expect to find the first complex VoiceXML 3.0 technology mostly in telephone-connected and cell-phone activated systems – ones that have more controllable voice quality.

Hopefully, with the new W3C standards, companies can dedicate more to useful speech recognition and less to reinventing the wheel. Standards usually lead to software tool kits for programmers and these often end up in popular packages like Microsoft's Frontpage and Adobe's Macromedia Dreamweaver.

Amateur and professional web designers alike may soon find a compelling reason to upgrade to voice enabled web design suites.

Maybe one day you can toss that pesky keyboard and mouse and talk to your machine instead – a promise made since the late 1980s and not yet satisfactorily realized.

by Philip Dunn, Copyright 2005 PhysOrg.com

This document is subject to copyright. Apart from any fair dealing for the purpose of private study, research, no part may be reproduced without the written permission. The content is provided for information purposes only.