

THINK SERVICES A DIVISION OF UNITED BUSINESS MEDIA

CUSTOMER MANAGEMENT **insight**

EVENTS TRAINING CONSULTING NEWSLETTERS WEBCASTS BLOGS

- SHOPPING CART
- LOG IN
- MEDIA CENTER
- MAILING LIST
- CONTACT US
- HOME

Speaking Tour: ASR And TTS Follow The Money

Free CallCenter Insider Newsletter

- [Magazine](#)
- [New Customer Management Insight Digital Publication](#)
- [Call Center Magazine](#)
- [Archives](#)
- [FREE Webcasts](#)
- [Podcasts](#)
- [Vendor Guide](#)
- [Multimedia Content](#)
- [Online Advice](#)
- [eNewsletter Alerts](#)
- [ICMI Forum](#)
- [ICMI Jobs](#)
- [Training](#)
- [Consulting](#)
- [Newsletters](#)
- [Publications/Tools](#)
- [Industry Research](#)
- [QUEUETIPS](#)
- [Conferences](#)
- [ICMI Global Report](#)
- [Call Center Knowledge Online](#)
- [Join ICMI](#)
- [Members Only](#)
- [Vendor Guide](#)
- [Industry Information](#)
- [About ICMI](#)
- [Download the ICMI Catalog](#)

A survey of the current speech market shows voice-activated dialing and call center offload are still the easiest sell. These may open the door for voice portals.

By [Elen Muraskin](#)

06/05/2001, 12:08 PM ET

Platform Vendor/ASPs

VocalPoint

Platform vendor/ASP [VocalPoint Technologies](#) (San Francisco, CA - 415-615-0600) sells a proprietary, patent-pending voice browser as software.

Vocalpoint's particular claim is in generating its voice application dynamically from HTML, leveraging the increasingly dynamic nature of the visual web itself. Its proprietary software, generating what it calls a Navigation Object Model, doesn't scrape screens; it uses algorithms that make best guesses at menu structure and compiles branched menus of prompts and grammars on the fly. Deployers can refine these trees of menus through cascading style sheets.

"With a myYahoo! page, with say, 2,000 lines of source HTML, you can generate a high-quality experience with 30 lines of style sheets. Even if you're super-skeptical that you can render a great experience on the fly, the reality is that you can get a good experience with a minimum of effort on any web page that's out there today," says CTO Garry Chinn. A typical Amazon.com page, much simpler, takes three lines of style sheets, he says. The style sheet tells the user where the core information is. The algorithm does a very good job on first pass, he says.

"We try to identify some basic elements of content on that page, and it's not that hard to ID contiguous units of information. After we identify these blocks of information, we synthesize a hierarchy above it. When the browser goes

UTILITIES

- [print this article](#)
- [e-mail this article](#)

RELATED LINKS

[Nuance Releases Recognizer V9](#)

[Nuance Unveils Voice Search](#)

[Microsoft to Acquire Tellme Networks](#)

[Aspect and Microsoft Create Speech Application](#)

[Nuance Acquires BeVocal](#)

[Aspect's Voice Portal Receives Nuance Certification](#)

[Sarbanes-Oxley vs. Hosting?](#)

[Speech Recognition Comes of Age](#)

[Aspect Intros VoiceXML Portal](#)

[Paul English, IVR Consultant?](#)

Your Email Address

Optional Areas of Interest

- International New s
- Advice/Tips
- Technology
- Agent Development
- IVR

TechEncyclopedia

Define It

through this exercise, it builds out this hierarchy. It may not be perfect, but then the user can go in there and prune out a whole subtree, because it's something that no one would ever want to hear in an audio experience." The user can also tweak the automatically generated prompts to "liven up" the experience and augment the grammars.



They do not currently use VoiceXML on their host platform. VocalPoint's own ASP platform uses NMS boards, Nuance speech rec, and L&H Realspeak TTS. They're migrating to SpeechWorks' Speechify TTS, however, and from NT to Unix. Christine Koncol, vice president of marketing, says that the Browser software is sold by license and maintenance contract, with yearly upgrades. The ASP model sells on usage, and number of ports. They're expecting a 50/50 mix of hosted vs. browser-sale business, she says.

"Today we're working with HTML - down the road we'll be working with XML and SHTML, as well," says Chinn. "We can take that Navigation Object Model (NOM) and create an XML representation of it as well. In the old days, you saw monolithic HTML pages. The web evolved towards more dynamic applications, and that's where we're starting out.

A lot of people using VoiceXML pages today are revisiting the beginning of the web, starting with these big, static apps. With the evolution of VoiceXML, they will move to dynamic apps, (and they will use the spec when it matures, he adds) but our platform already works with those. If you have a dynamic HTML application, you have a dynamic Vocalpoint application as well."

Telecom Italia: Mobile carrier Telecom Italia is VocalPoint's star customer to date, illustrated to good effect on the company's website. Telecom Italia used the browser to voice-enable selected websites and launch a suite of new voice services, including familiar consumer-type portal information services, and messages, under the name VoxNauta. In addition to their own preexisting (NMS-based) IVR platform, Telecom Italia hitched the browser to their own "home-grown" Italian ASR and TTS. Also: Abovehealth, a health-plan ASP, is now in soft launch. EasyBenefits.com, an employee benefits web ASP, is using Vocalpoint for its voice equivalent.

New Respect For Speech Rec

We were en route to Amherst, MA, making wonderful time to the first stop on Daughter Number One's college tour. Traffic suddenly slowed. Then it crawled. We were in Waterbury, CT - a place that had surely never seen the shadow of Shadow Traffic. Was it rush hour? Was there construction up ahead? Were the gods of college admissions telling us we were meant to pay New Jersey in-state tuition?

I could only whip out my cell phone and call 800-555-TELL. Barging past its friendly menus, I said "travel," then "traffic," then "Waterbury, Connecticut." TellMe recognized the town name with perfect accuracy. Unfortunately, it didn't know anything about traffic in Waterbury. Danbury, it knew.

TellMe asked me for the Interstate we were following. I said "I-84." TellMe told me there were three incidents reported on I-84: The third one, a fuel spill, was happening three exits ahead, at Exit 25. We drove off the next Exit (22) and backroaded our way to Exit 29, where all was clear. We were half an hour earlier than the cousins we were supposed to meet near Exit 30 (leaving us plenty of time to take in all the sights - pharmacy and coffee shop - of Plantsville, CT). "Wow," said my husband. "Cool," said Daughter Number Two.

Future Vision: Rosy . . .

Steve Ehrlich, vice president of marketing for Nuance, says, "Have we got perfect business models yet? No. Are they starting to take shape? Absolutely." His vision of the voice web falls somewhere in between the ad-sponsored consumer voice portal that tells you everything for free, and the simple offload app that saves the call center \$35,000 per agent seat per year.

"Every enterprise will evolve to have a set of voice applications which will become their voice site," he explains. And you, as unsuspecting customer, may meet these applications on your way into a simple 800-number call. A clever, speech-enabled airline 800-number may confirm your flight to Chicago, then go on to ask if you'd like to make restaurant reservations there, and if so, which cuisine, please? Or offer Chicago's weather forecast. Or driving directions to your hotel. "These will all be content that a TellMe, for example, already has," says Ehrlich. "So the ability to link from the enterprise agent-replacement app to some of this content is going to start occurring within the next year. The call center becomes part of this bigger picture, like a website is." A voice host, Ehrlich posits, may even offer to build this app (mostly prewritten, anyway) for an airline client for free, and build the customization cost into the per-call or per-minute hosting charges.

"That's really appealing for the airline, because it makes the customers more inclined to call them. TellMe gets greater per-minute revenue. Nuance sells more speech ports. Everybody's happy and the voice web happens."

. . . and thorny

Ehrlich thinks that the speech-hosting pie is large enough to accommodate a bunch of players. The more immediate question is whether the economy will float long and high enough to keep them in business until the hosting model takes off as it should in the Rosy scenario, above. "Guys with less funding are really going to struggle to stay in business," he concedes. They may be bought out. New players - web portals like Yahoo and MSN - will start competing for the applications business. Companies like Net2Phone will enter the ASP market." Indeed, Net2Phone has already made that intention clear, in a joint announcement with HeyAnita. Under the agreement, Net2Phone and HeyAnita will jointly market end-to-end voice hosting solutions to enterprises, call centers, broadband service providers, and its established user base. Net2Phone supplies telephony

VoiceXML Developments

Hand-in-glove with the voice-host proposition is the development and standardization of the common voice markup language, VoiceXML. The idea is for a VoiceXML browser to run on Voice web servers the way Apache and Microsoft's IIS serve up HTML to PC-based browsers.

Last December we reported that it had begun to "rain" VoiceXML interpreters, as momentum gained behind a voice app gen with HTML-wide appeal. This affected IVR platform providers on the vendor side, and buyers as well. Steve Ehrlich of Nuance confirms: "A lot of enterprises are now thinking, 'Do I stay with my IVR or do I go to an open, VoiceXML platform?' We're in that transition state now, and starting to see movement towards more ASPs. The unwillingness of many IVRs to implement VoiceXML is starting to hurt them."

Version 2.0 of the VoiceXML specification is now in progress. According to Stuart Patterson, SpeechWorks CEO, the core recognition vendors have to tread carefully in deciding how far to develop a VoiceXML browser (software that interprets VoiceXML and also includes the navigation tools to move from site to site, make bookmarks, and note other user preferences). The core vendors also have to choose their degree of "open source" spec compliance. He says that over 20 platform partners have expressed interest in incorporating [VoiceXML components](#) from SpeechWorks, but want to use them in their own architectures and platforms. "We've had a strong request over the past six months and optimized our recognition engine and the way we deliver our VoiceXML stuff to deliver this as components, so that people can choose the level at which they want to integrate with our technology. We've gotten excellent reception to that approach."

Where do SpeechWorks Dialog Modules, (prepackaged grammars, code, and exception routines for common recognition tasks) fit in with the markup language? "There are several ways to use Dialog Modules in a VoiceXML environment," says Patterson. "As currently implemented, in half a dozen to a dozen different environments, they're obviously not standard or fully compliant VoiceXML. Part of our partnership with VoiceGenie is to come up with a more VoiceXML-compliant version of Dialog Modules. We're very bullish about that, because VoiceGenie is one of less than a handful of companies that can offer you something on-premise that is nearly or completely 100% compliant.

"On the other hand, there are a ton of companies that can offer you VoiceXML within their ASP offering. In many cases, they're not set up to support VoiceXML Dialog Modules yet, but they may be attracted by them. All of our hosting partners are interested in incorporating VoiceXML into their offerings." Nuance is currently testing VoiceXML compatibility within its own camp of developers and ASPs, to make sure anything built with its V-Builder VoiceXML-creation tool runs on BeVocal and TellMe's platforms. "It's a painful but short-term issue," says Ehrlich. "It has forced some application companies to choose their ASP

Adapting Core Recognition Technology

Adapting core recognition technology to VoiceXML is not simply a matter of converting the modules themselves into VoiceXML code. It also means adjusting to a client-server architecture in which the core recognition engine may not reside in the same platform as the browser, or the VoiceXML application itself. This is a departure from SpeechWorks historically tight integration. "How do you do dynamic grammars, for example," notes Patterson. "This is not fully specified, in part because it's handled differently by different speech companies. We've tried, in optimizing for VoiceXML, to make that more explicit and non-proprietary, so that companies that are going VoiceXML A-Z can take advantage of how we do it without feeling that they're totally locked in."

"Some aspects of speech that aren't yet fully defined (in VoiceXML) will be an issue when someone says they're agnostic about whether they support IBM, SpeechWorks or Nuance," says David Nahamoo, of IBM Speech. "Because VoiceXML is not explicit in all the things that are really needed to do the highest-end applications." Almost all platform companies will be forced to add extensions to do things like dynamic grammars, and that will commit them to a particular ASR or TTS, he says. "Complex applications that can't be done with VoiceXML 1.0 we'll do with other implementation models, such as the state-table model, the scripting language that we have for Direct Talk," says Nahamoo. "As we go forward, VoiceXML will support all those complex mixed-initiative interactions and natural-language understanding."

"I look at VoiceXML as something that is stabilizing and will allow a much broader community of developers to start implementing applications," he adds "We just finished a contest with a set of college students who had no familiarity with this programming language. They managed to take the VoiceXML desktop toolkit, write an application, and get to a point that they could be reasonably implemented from the desktop with ViaVoice interaction." (Awarded by the VoiceXML Forum and IDC in addition to IBM, the \$25,000 grand prize went to Brent Metz, 21, originally from Fairfax, VA, and a student at Virginia Tech University, who wrote a diet-planning application that recognized foods and calculated calories.)

Planes, Trains, Bar Codes, and Voice Verification

[Aeritas](#) (Irving, TX - 214-260-0400) is a startup with a "[line-busting](#)" application aimed at speeding business travelers through airport check-in. It's due to start U.S. pilots soon, and is scheduled to take off on Lufthansa this September. They have a technology partner in the Sabre travel system, used by travel agents and covering 91% of flight, hotel, and rental car online information, says Malik Mamdani, Aeritas CEO.

A close integration with OpenWave's WAP servers lets Aeritas Express for Travel accomplish some edgy multi-modal wireless interaction, well before 3G's ETA. The app works as follows: Business traveler registers for the service over the web, arriving at the airline site via Travelocity, Expedia, or other portal. Registrants enter mobile phone number, among

other things. Aeritas' server (NMS and Nuance 7.0-equipped) calls them and asks them to speak in a phone number to establish a voiceprint. The call ends. Customer proceeds to buy herself an E-ticket over the web or even through an agent.

On the day of departure, the passenger receives a WAP text message on her phone with terminal, the flight, and the gate assignment on it. It also has a link; click on it, the WAP session ends, and the phone automatically makes a call into the server, which asks the caller to speak her phone number. It performs the pattern matching that verifies identity - more reliable than fingerprints, says Mamdani. The server also asks the passenger all the FAA-required questions about packing one's own bags and accepting articles from others. The call ends, and Aeritas' platform sends a bar code to the phone's screen - one that can be read with the existing boarding-pass scanners now found at most gates.

With your voice registered and all your baggage carry-on, just think how close you can cut it now before you hail that airport cab.

<< [Previous Page](#) | 1 | 2

										
										