

Solutions Profile

Company: Vicorp

HQ: UK

Segment: Service Creation Environment (SCE), Speech Technology

Date of Profile: October, 2004

Author: Mark A. Plakias, Principal Researcher

Companies mentioned in this Profile:

Audium (US), Fluency Voice (UK), TuVox (US), Vicorp (UK), Voice Objects (Ger.)

Key Concepts:

Automated Speech Recognition (ASR), Interactive Voice Response (IVR), Integrated Development Environment (IDE), VoiceXML (Voice Extensible Markup Language), MRCP (Media Resource Control Protocol), Java, Application Server, WebServer, Client-Server, Object-oriented Computing, Reusable Java Software Components.

Abstract:

This Solution profile discusses the evolution of phone from a separate, 'specialized' cost center in customer management operations to part of an enterprise-wide multichannel customer interaction utility that uses web-based software technologies. Moving speech recognition-based self-service and information access into this multichannel utility favors standards-based service creation tools that can provide well-behaved Java components to existing native application servers and web servers. Legacy IVR platform providers have been slow to wake up to this requirement, leaving the door open to 'pure-play' third-party IDE vendors who are not tied to any one runtime platform. Another issue is that few players in this service creation segment have been able to avoid imposing additional runtime requirements, or limiting the customer to predefined 'cookie-cutter' application templates.

Reality: Multichannel Customer Interaction is now the Norm

We all live the multichannel reality as consumers everyday: you go online to make price comparisons, you call a toll-free number possibly on a mobile) to get the nearest store location and confirm product availability and you engage in 'offline' interactions with the physical store. Afterwards, to get help in installation, you may be on the phone again to a technical support help desk.

While you may notice some glitches traversing this multi-channel customer service fabric – such as having to give your telephone or account number multiple times -- there is no question that this is the way it will be, there is no going back for us as consumers. Each channel makes sense for a specific purpose or context, and we want them all at various times.

The question is whether the company you are doing business with has caught up to where you are.

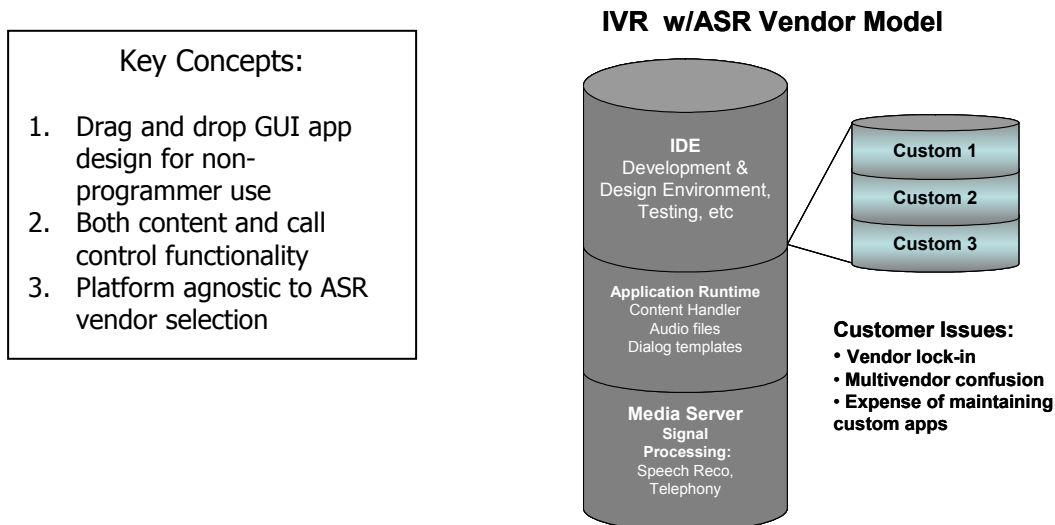
Increasingly, the answer is yes, but as enterprises bring customer service reps to customers via text chat on the web, or email, or telephone, the hardest channel to update has been the interactive voice response (IVR) telephone self-service platform. The same virtue of these touchtone platforms, durability in the face of high call center volumes, is keeping them in service long past the point where they are cost-effective to maintain, from a vendor or customer perspective. Indeed, the inflexibility of IVR technology is such that many very large installations are approaching 'end-of-life' status as vendors discontinue making or supporting them.

Enter speech recognition as an obvious upgrade to the end-of-life touchtone platform problem.

Problem: Telephone self-service lags behind

Early Speech/IVR vendor-driven engagements set the tone for the first stage of the market: a proprietary IVR platform which provided a highly-integrated (albeit specialized) stack of vendor-specific application development tools (typically written in C++ or other lower-level language), an efficient but proprietary runtime for executing applications and running recognition engines, and a voice response cage that included the necessary signal processing boards for decoding and encoding tones and utterances for the recognizer.

Typically, custom apps were developed by the IVR and speech vendors' professional services teams, resulting in multivendor team coordination issues, as well as others shown below. However, several useful key concepts are present even in this model.



The key point to understand here is that the model above contradicts the client/server model which drives most web development today – hence phone self-service has lagged behind, trapped in an obsoleting vertical stack.

The New (Speech) Application Development Market

Since the late '90s, speech recognition has started to move to a position as an enterprise utility, just as HTML-based and web-based information delivery is today. This involves making speech tools available to the community of existing developers, administrators, brand and product managers that all touch web applications today. An important part of making speech control over everyday business processes possible is integrating the

unique aspects of voice interfaces with the existing, typically Java-based, tools and infrastructure used to support web applications.

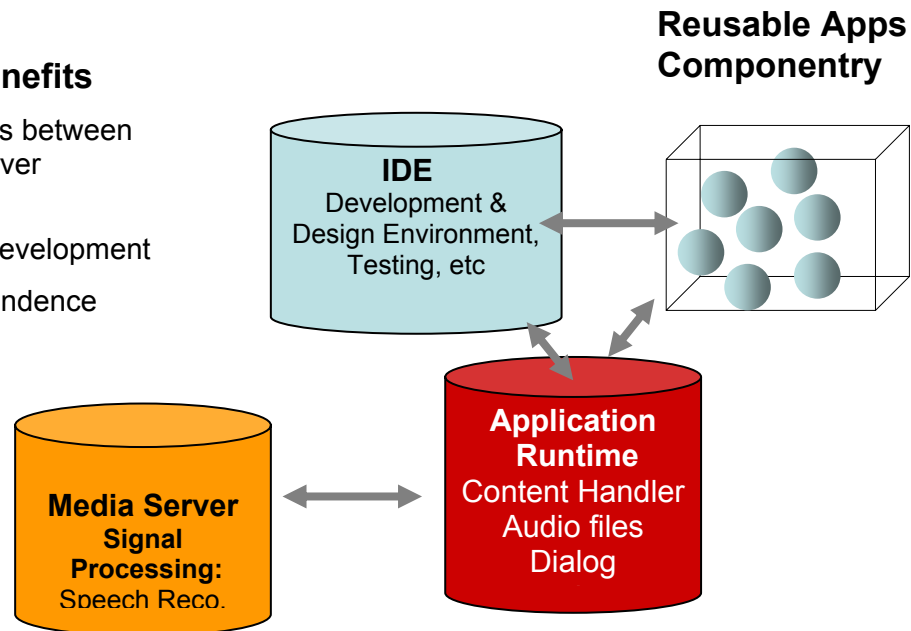
An important parallel trend in making speech a utility is the growing amount of enterprise-like software development and maintenance within telephone companies of all types. With increasingly complex customer service, supply chain, and HR challenges, today's telco looks a lot like an enterprise – but bigger. In addition to high availability and scalability requirements, telco's have the additional requirement of wanting the option to host applications for enterprise customers. Speech is a particularly attractive candidate for hosting, given its complexity, and helps keep enterprises tied to the carrier. This has implications for service creation tools we will discuss below.

The logical outcome of this larger trend and how its affects the problems with speech development described above is to separate out the development environment (referred to in this report as either an Integrated Development Environment – IDE, or as a Service Creation Environment—SCE) from the rest of the platform stack. This is illustrated below.

Distributed Speech App Server Model

Customer Benefits

- W3C standards between commercial server platforms
- Less custom development
- Vendor independence



Recommended Approaches: Progress towards Open Development

Today, a small but growing class of solutions providers have emerged to make this happen. This solution set benefits multiple players in the ecosystem involved with speech recognition, which as we have seen above, sits in a larger complex of multichannel customer management solutions.

The fundamental characteristics shared by the solutions discussed below are:

1. Standards-based multi-vendor operation: None of the solutions discussed here are from a speech recognition vendor, and none of them are designed to work specifically with one or another. In 2003, the emergence of the MRCP protocol allows these solutions providers to work with any speech engine.
2. Putting design best practices into software: Behind all of these solutions is the value proposition that VUI best practices can be learned and embodied in software, rather than built custom every time. This proposition connects to faster development cycles and lower costs.
3. Use of mainstream application and web server infrastructure: All of these solutions allow customers to use the same app servers and web servers used for their web operations to hold the speech templates, related content and object libraries, and serve them up as the equivalent of logical 'pages' over the phone.
4. Third-party development tools provide choices: Finally, the 'unbundling' from the speech engine allow customers to mix and match best of breed components from speech engine, text-to-speech, and media servers.

In addition to the above general principles, a general tendency exhibited by new companies in this space is to move towards prepackaged capabilities. These capabilities range from 'components' that can be commonly-recurring aspects of any application, such as name and address collection, to full-blown applications such as dealer locator or ATM finder. While it is understandable that any software company making tools for applications development can be pulled into making a few themselves, it is viewed by Opus as a slippery slope that can lead to dilution of focus, and worse, competition with channel partners who see their value-add as providing finished applications. By this criterion, reusable components are good, finished packaged apps are bad.

These broad themes need to be viewed through the lens of specific architectural decisions that individual vendors have made in creating their products. The key architectural issues identified by Opus Research are as follows:

1. Runtime Involvement

Essentially, this decision comes down to how much of an additional load the solution puts on the application server. In one scenario, the speech solution is involved in the runtime, and customers will require additional training, in monitoring and maintaining the solution. The other approach is to eschew any additional metadata and allow the components to be handled as any other content by the native application server. *This is much less disruptive to ongoing web and app server administrative processes.*

2. Separation of Presentation from Logic

Opus Research looks for several architectural decisions to deliver on this promise. One is the separation of the call flow design component from actual elements within the flow – *users should be able to create a reference to an object in the flow without having to define the specifics of the object.* Further on, the contents of the objects should be open to modifications without affecting the flow, and without requiring programming knowledge.

3. Roles-based Service Creation

What flows from the above architecture decision is the obvious result that expensive programmers should not be required for every change, because they are hard to get a hold of. *Opus Research looks for very clear examples of how simple forms fills can be used by non-technical personnel to change content, and ideally, macros or rules that automatically bind different content based on events or schedules.*

4. Partner Ecosystem

The two directions here are really quite simple: either the IDE is viewed as the 'kingpin' in the overall speech solution, or it is a well-behaved component – albeit an extremely important one – in the process. What Opus Research looks for here is a business model that is tightly integrated with partner objectives, and allows the solution vendor to remain in the background. If the partner is a platform vendor, e.g.: Nortel or IBM or HP, then the solution should be capable of being offered as an IDE that will generate well-behaved applications at a minimum level of vendor involvement. This is good on two fronts: (1) it incents the platform vendor to resell the solution, and (2) promotes customer satisfaction.

5. Multi-tenant Administration and Application deployment

The concept of multi-tenant operation means that large outsourcing providers or telcos with multiple customers (tenants) to manage need additional tooling for managing these customers and their individual applications in an efficient way. Unless this segment is addressed architecturally, chances are the solutions provider will miss out on a lucrative market.

6. ECLIPSE Conformance

One key technical detail that relates to the partnering discussion is the software basis for the IDE/SCE platform itself. Since most of the solutions studied here involve manipulation of Java libraries, they should be considered in the context of an open source community-based underlying IDE framework known as ECLIPSE. The ECLIPSE foundation has become the standard for anybody building development environments

for the Java world. The use of the ECLIPSE open source framework as the basis for the IDE is thus a legitimate basis for evaluating these platforms.

7. Redistributing the Value Chain

A key commercial objective of most players is to insinuate themselves into an already-crowded value chain, with the usual metric being a port, essentially a single phone channel. Other cost drivers include the actual speech recognition software (\$700-\$1500 per port), text-to-speech software (\$250 – \$450 per port), and the telephony interface card (\$200 - \$450). Some of the solutions providers covered here have exacerbated this problem with pricing in the \$2,000 per port range (VoiceObjects), based on the premise that the application development-and-runtime model is paramount. Others, such as Vicorp (which avoids any runtime involvement) have taken a more measured view and priced for volume scenarios, rather than a strategic pricing model. Opus Research looks for volume pricing to be in the \$4-500 per port range for a realistic market adoption model.

How Does Vicorp Fit into the market?

In this Solutions Profile, Opus Research presents the results of an examination involving five current solutions providers, shown with their respective base of operations:

- Audium (US) (www.audiumcorp.com)
- Fluency (UK) (www.fluencyvoice.com)
- TuVox (US) (www.tuvox.com)
- Vicorp (UK) (www.vicorp.com)
- Voice Objects (Ger.) (www.voiceobjects.com)

The vendors shown above were evaluated to ascertain which is the most applicable for the following market situations:

- Departmental-scale single application (or limited number) within a defined vertical
- Large Enterprise
- Service Provider

Below we show the aggregated results of the evaluation, after a discussion of the Vicorp XMP platform from an architectural perspective, as well as a business model view.

Vicorp

This UK-based solutions provider Vicorp is something of an anomaly in a segment populated by startups, having been in business since the early '80s, with a strong background in the telecom service creation environment (SCE) business. As such, it has the benefit of the broadest customer footprint of any player in the space, with installations numbering in the hundreds for its legacy touchtone-based systems. These days, Vicorp has reinvented itself as an open-standards based pure-play SCE, with the signal distinction of being the only player studied that truly conforms to the no-runtime

paradigm discussed above. This is a remarkable feat, given its legacy business required it to have software code addressing all layers of the solutions stack, including deep into the runtime.

As discussed above, Vicorp achieves the no-runtime footprint by contributing Java libraries to the application server runtime, without imposing extraneous metadata or other runtime management tools. This makes it highly attractive to app server companies such as BEA (WebLogic), IBM (WebSphere), and Sun. With respect to the other architectural criteria discussed in this report, we find that there is a clear separation of presentation from business logic, and the roles-based maintenance of applications by non-technical personnel is clearly delineated in the user interface, not just in the drag-and-drop GUI (which is virtually universal across all platforms studied) but in the forms-based structure for customizing content. This easy customization can be defined by event triggers or business rules that are also easily manipulated.

Perhaps due to its telecoms lineage, the Vicorp platform has multi-tenant capabilities deeply architected into its structure. Applications can be 'white label' instances of a common root application, each maintained within its own administrative domain with multiple layers of reporting and administrative access all subject to permissions. This includes call detail generation that is sufficient for telecom applications, pursuant to the old adage 'if you can't bill for it, it's not a product.'

Other positives include a relentless partner orientation, matched only Audium among this class of players. With no danger of alienating its channels, Vicorp has clearly seized on its service provider-scale capabilities and sewn up HP and its OpenCall Media Platform as a strategic partner. It has also responded to the IBM RDC initiative. Technical issues that remain to be sorted out are upgrading the platform to be an ECLIPSE-based environment, which management has committed to doing.

Key Findings

The following tables show the basis for the evaluation of the five solutions providers referenced in this Solution Profile (the full text of this report can be requested from Opus research at www.opusresearch.net). In these tables we summarize the results of the analysis based on the architectural issues described above, and on the commercial issues surrounding the solution providers' market decisions. Those fields flagged with + indicate a marketplace advantage.

Technical Analysis

	Runtime Component	Separate Presentation	Roles-based Development	Multi-tenancy	ECLIPSE-based
Audium	Yes	Yes	Medium	Medium	Yes +
Fluency	Yes	Yes	Limited	Medium	No
TuVox	Yes	No	Limited	Low	No
Vicorp	No +	Yes	High +	High +	Roadmap
VO*	Yes	Yes	Medium	Medium	No

Business Analysis

	Customer Footprint	Partner Metric: # of OEM Relationships	Pre-Pack Apps Focus	Sales Strategy
Audium	Strong +	5: Cisco, Holly, Syntellect, VG*, Other	Moderate	Indirect +
Fluency	Low	1: Avaya	High	Direct, Indirect
TuVox	Low	3: Genesys, Verascape, VG	Medium	Direct, Indirect
Vicorp	Strong +	4: HP, VG, IBM, ScanSoft	Moderate	Indirect +
VO*	Low	3: Comverse, Genesys, VG	Low	Direct

*VO = VoiceObjects VG = VoiceGenie

The above analysis shows in the high-end of the market there are two leaders, Audium and Vicorp. This segment includes ASPs and telcos, where Audium and Vicorp both have reference customers. Both have substantial wins in the US telco market, and abroad (Audium at Telstra in Australia, Vicorp at multiple accounts including Bell Canada). We have already noted that four out of the five solutions profiled herein claim service provider customers, but the architectural decisions made at Vicorp have made it clear to us that they are gearing for the long-haul in the service provider market, which we view as in the early stages of adoption. Specifically, Vicorp's carrier-centric merits stem from conscious architectural provisions for multi-tenant operations, and a runtime-agnostic design that promotes scalability from the application-serving infrastructure, rather than inhibits it.

Summary: Vicorp for Carrier-Grade Multi-tenant Deployments

It is a truism in high-tech generally that the scope of opportunity for new players narrows as the marketplace becomes more educated, and looks for 'carrier-grade' solutions. We like Vicorp's strong, clearly thought-out multi-tenant features, as well as its no-runtime administrative clarity, and Audium's success at carrier-grade ASP accounts such as Convergys and West cannot be denied. From a technical perspective for this segment we would follow HP in giving Vicorp the nod over Audium, especially with its pedigree in telecommunications. Both firms have the right go-to-market approach for service provider, showing the respective management teams are both very clear on their place in the ecosystem. Continued success in the service provider market will constitute validation of the Vicorp business model, while it's adoption of the ECLIPSE framework will close out one of Audium's most salient differentiators.

About OPUS Research

Founded in 1985, Opus Research has been at the forefront of advances in interactive communications for two decades. As an OEM research supplier, as a consultant to telecom providers and their suppliers on three continents, and as a commentator on contact center technologies whose principals have been quoted in the *New York Times*, *Wall Street Journal*, *Businessweek*, as well as in appearances on CNN and Tech TV. Mark Plakias writes a monthly column for Speech Technology Magazine. Visit Opus Research today at www.opusresearch.net .