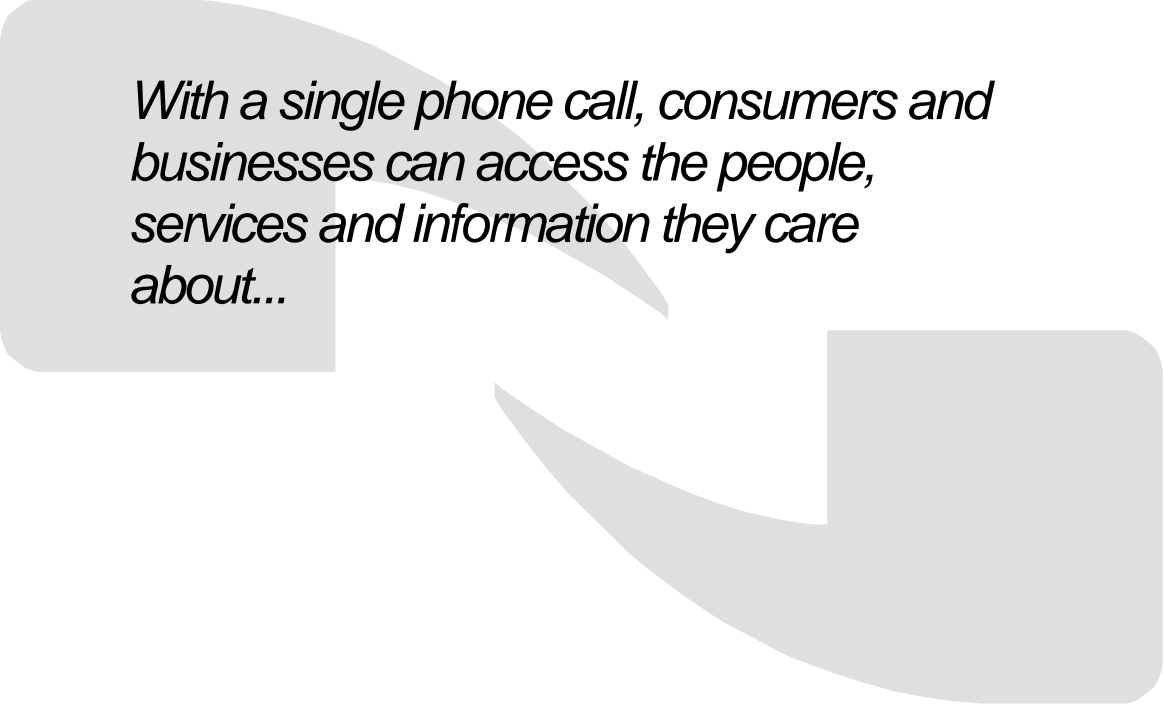


## White Paper

# The Voice Web

---

*With a single phone call, consumers and businesses can access the people, services and information they care about...*





# The Voice Web

*With a single phone call, consumers and businesses can access the people, services and information they care about...*

## Introduction

The richness of the Web. The accessibility of the telephone. The convergence of these networks is forming a vast new network called the Voice Web. What is the Voice Web? It's an interconnected network where Internet-based information, e-commerce, and personal communication services can be accessed from any phone using human voice. The Voice Web delivers the ability to make phone calls and access Internet content in one continuous call session.

Spoken conversation is still one of the most natural and effective ways for people to access information and conduct business. With the rapid advances in computer processing and speech algorithms, telephony-based voice recognition systems can now understand natural spoken conversation. Since the mid-1990s, many market-leading companies have operated V-Commerce™ (voice-driven e-commerce) systems to deliver services like flight information and real-time stock trading using voice authentication. And just as individual IT applications gave way to the Internet, these distinct voice-driven applications are evolving to a new network called the Voice Web.

What will the Voice Web sound like? Imagine picking up the phone on the commute home to voice browse the latest high-tech job openings, listen to and buy the latest CD, or call a client by name from an electronic address book.

The Voice Web is happening now. A growing number of companies are moving at Internet speed to aggregate services (unified messaging, address book, shopping) and Internet-based content (stock quotes, movies, Yellow Pages) and make them available over any telephone. Enterprise and e-commerce companies see tremendous opportunity to expand their customer reach to the vast number of global phone users, and they're engaging the help of infrastructure providers to voice-enable their content and services. The availability of standards-based technology and products will fuel the rise of Voice Web services in the same way that industry standards enabled the rapid growth of the World Wide Web.

This paper is intended to provide a basic understanding of the Voice Web and give an overview of the kind of services businesses and consumers can expect. After reading this paper – enterprises, telecom service providers, Internet, and infrastructure companies will understand how they fit into the growing Voice Web ecosystem.

## Market Forces Driving Demand for the Voice Web

A number of market forces are coming together to drive demand for intuitive voice interfaces to content and services that can be used anytime, anywhere.

### *The pervasiveness of telephones, especially mobile telephones*

There are over 850 million wireline phones and 450 million wireless handsets in use today as compared to only about 200 million Internet-capable PCs.<sup>1</sup> And with the increased affordability of mobile service, the growth of wireless handsets is estimated to top 1 billion by 2003.<sup>2</sup> This makes telephones by far and away the most ubiquitous network access device. Since voice recognition servers are accessible from any existing telephone, telephones are the perfect device to offer services to the mass market.

### *The explosive growth of the World Wide Web*

In less than 5 years, the Internet has revolutionized the world as a new medium for information delivery. Today, more than 260 million people use the Internet as an indispensable part of their daily lives - for convenience, entertainment, and enhanced productivity in the workplace.<sup>3</sup>

The Internet is attractive because it provides one stop access to obtain practically any type of information. But obtaining information on the Web requires sitting at a desk, which is becoming less desirable in a highly mobile society. Consumers' Internet addiction is driving device manufacturers to accommodate mobility through personal digital assistants (PDAs) and smart phones with Web connectivity. Palm™ handheld devices and Wireless Application Protocol (WAP) cellular phones are popular examples of these.

These appliances are raising the awareness that mobile access to Internet content is possible. While text-based Internet access is helpful, slow transmission speeds and the difficulty of interacting with small screens limits the value of these devices.

The proliferation of the telephone and the appeal of anywhere, anytime access to Internet content is sparking the demand for simple, but personalized services that let individuals and businesses obtain fast and easy access to the people and information they care about. The Voice Web will transform telephones overnight into the world's most convenient communication vehicle for voice and data retrieval.

## Speech Recognition Technology

While telephones have existed for over 120 years, commercially viable speech recognition is a relatively recent development. In fact, telephony-based speech recognition software that provides for large vocabularies, requires no caller training, and supports any type of telephone was not even obtainable until the mid-1990s. Current speech systems support naturally spoken phrases such as "I wanna buy 33 shares of 3Com at the market".

People like using their voice instead of pushing buttons or even talking to humans. In 1999, Evans Group Research collected and analyzed the feedback of 250 users of various deployed speech systems.<sup>4</sup> Quantitative results revealed that 83% of callers preferred speech to touch-tone and 74% preferred voice-driven systems to talking to an agent.

Speech recognition is processed on a network server, not the actual telephone itself. This configuration makes systems robust and scalable to support millions of callers and to furnish the horsepower to recognize all of the possible ways callers might phrase their requests for the same information. Callers can also barge-in (interrupt system prompts) if they already know what they want.

Flexibility and ease-of-use are mandatory requirements for mainstream adoption of any new service. Businesses are building voice-driven services because speech recognition offers

an intuitive way for people of all ages and lifestyles to access the information and people they want from any telephone.

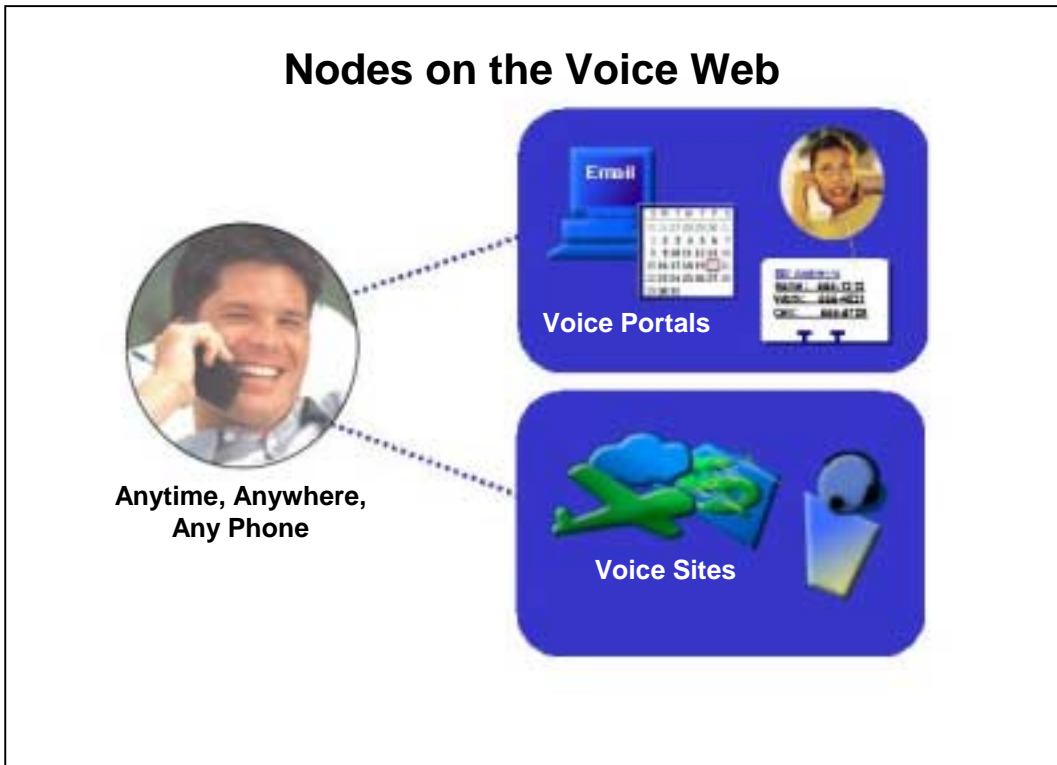
## The Voice Web Vision

The Voice Web vision is to unite two networks that individuals and businesses rely on every day—the Internet and the telephone network. This new, interwoven, voice-driven network will offer a magnificent array of Internet-based content, commerce, and personal communication services that are easily accessed with a single call and a helpful voice browser. In the near future, service providers will offer off-hook access to the Voice Web, replacing the age-old dial tone with a friendly “how may I help you”?

Teenagers and corporate executives alike will use the Voice Web to access the latest concert information or their company’s daily performance data. Of course it is not necessarily appropriate or even desirable to access every Web page over the phone. Instead, the Voice Web will deliver localized and personalized Internet-based information and services that are relevant to consumers and businesspeople when they are mobile. Voice Web applications will be defined and controlled by callers through the use of a standard voice browser that enables them to traverse sites or navigate content easily.

## Nodes on the Voice Web

Content and services will emerge in two ways on the Voice Web — as individual voice sites and as voice portals. Both voice portals and voice sites are nodes on the network and the connectivity between them is enabled through a standard voice browser, which is described later.



Voice portals are aggregated services (unified messaging, long distance, shopping) and Internet-based content (stock quotes, movies, Yellow Pages). Callers reach voice portal services through a toll-free number or by dialing a dedicated network-based access code.

Voice sites are nodes on the Voice Web that contain voice-enabled enterprise or e-commerce applications and V-Commerce services.

## Voice Portal Providers

Voice portals are the first category of application services to appear on the Voice Web. Fast moving portal providers are defining the initial consumer experience and opening the door for other companies to develop and market services for specific industries or for specific functions (shopping, for example).

There are dozens of companies focused on taking voice portal services to market. Among these early players are BeVocal and Tellme.

For the startup company focused solely on offering voice portal services, the Voice Web presents the opportunity to leverage successful Internet business models over the phone.

For the service providers, voice portal services present the opportunity to increase network usage and customer loyalty, attract new market segments, and differentiate existing service offers.

Initial voice portal services are largely horizontal - offering smart, synergistic information services and commerce applications centered on entertainment and convenience. Portal providers are packaging these content and services with an eye toward promoting successful business models that link useful information with secure transactions.

Many of these early voice portal services are targeted at busy mobile professionals. For some, that content will take the form of services that help them balance daily demands of work and personal life; for others it will be content that entertains them on the commute home after a hectic day.

Below are some examples of the times and places that voice portal services are helpful to the average businessperson.

In the car on the commute to work...

- To check headline news or local traffic
- To arrange for convenience services such as an oil-change
- To look up Outlook® contacts and voice dial a colleague by name

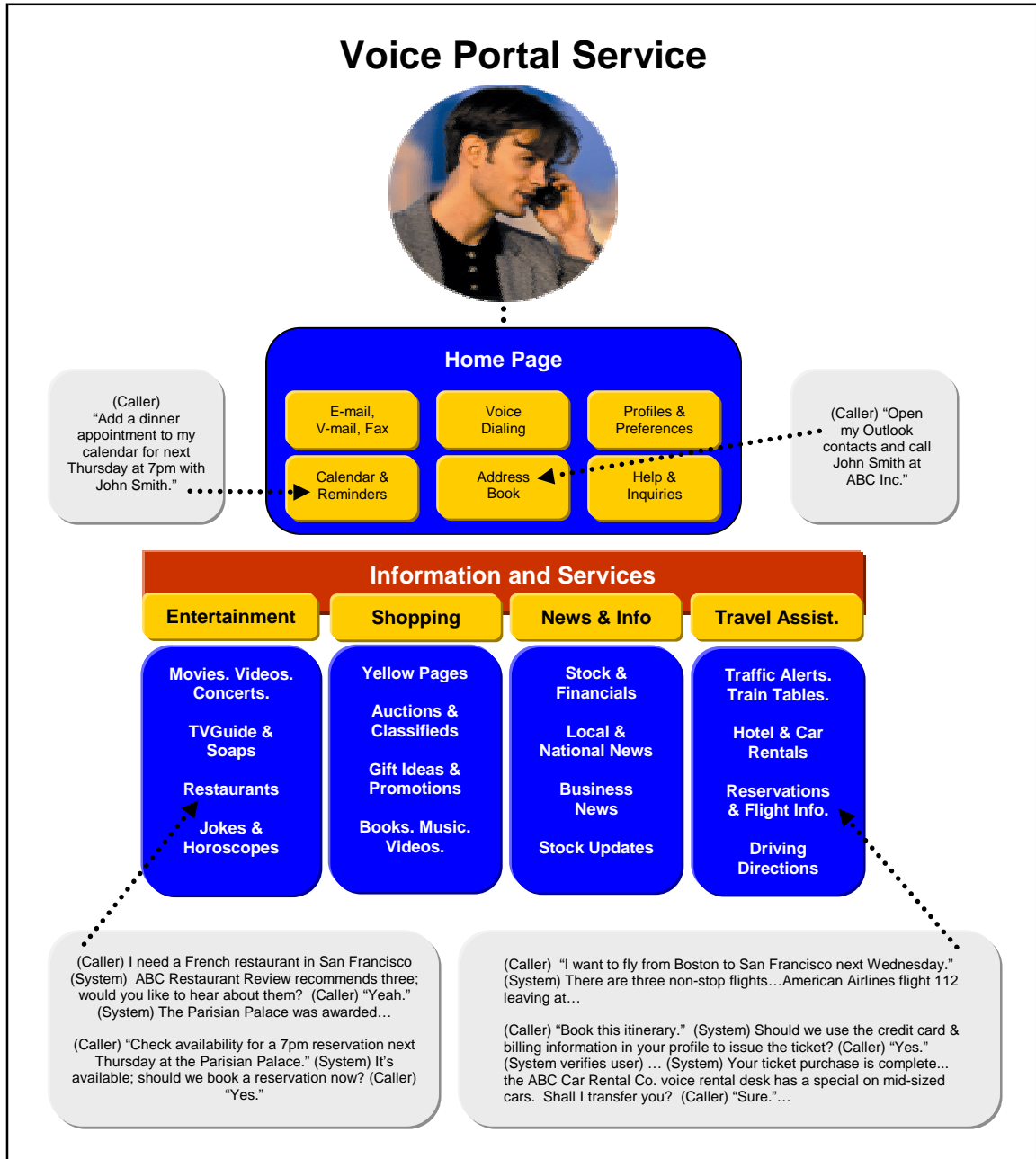
While traveling to another city for business or pleasure...

- To retrieve and respond to daily e-mails, faxes, and voice mail
- To find and book a local restaurant for an important client meeting
- To get driving directions to a customer site
- To handle travel arrangements, get upgrades, and obtain gate information

Anytime...

- To get local movie listings, critics' reviews, and buy tickets for tonight's show
- To conduct queries of favorite Web sites (job banks, classifieds, auctions)
- To get up-to-the-moment sport scores, stock quotes and ski or surf reports

One of the most appealing aspects of the Voice Web is the ability to do multiple tasks in a single call. The picture below shows how a businessperson completes a sequence of tasks related to a future business trip. He accesses helpful data that leads to a successful V-Commerce transaction, taps into his personal information and office aids, and makes a telephone call — all in one synchronous session.



## Voice Sites on the Voice Web

Much of the initial content on the Voice Web will be provided by voice portal service offerings, not unlike the CompuServe and AOL services of the early 1990s. But as the Voice Web matures, more content will come from e-commerce and enterprise voice sites.

Many enterprises are moving to reduce costs inherent in agent-based call centers and e-business operations.<sup>5</sup> Major airlines and financial institutions are already ahead of most enterprises, using speech recognition to increase customer satisfaction and obtain dramatic cost savings.

E-commerce companies will see their customer base and sales revenues burgeon with the increasing number of V-Commerce transactions, as the Voice Web turns telephones into instant marketing devices that personalize and localize the shopping experience. V-Commerce revenue is expected to exceed \$30 billion by 2005.<sup>6</sup>

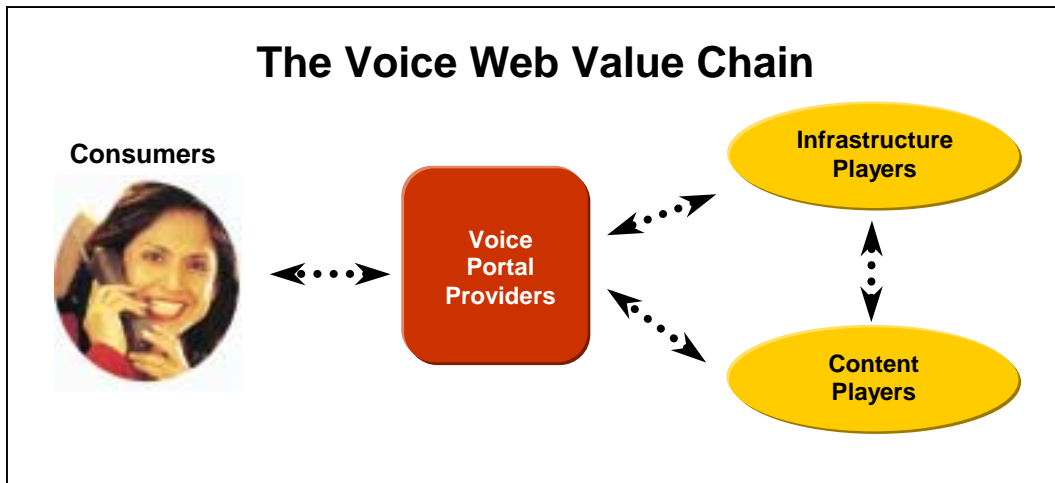
These V-Commerce solutions will evolve to become nodes on the Voice Web, accessible either directly or through connectivity from a voice portal service. This latter access method will impact how e-commerce companies and enterprise companies manage customer interactions today. By establishing supplier relationships with portal providers, enterprise and e-commerce companies will allow portal providers to link to an enterprise's voice site in much the same way a Web site is accessed from a Web portal today. Instead of a Web browser providing the connectivity, a voice browser enables callers to navigate from site to site. Revenue opportunities aside, enterprise and e-commerce companies are motivated to participate in the Voice Web, partly because it transfers the burden and cost of telephony infrastructure to the portal provider. In return, the enterprise or e-commerce company establishes a business relationship with the portal provider to process V-Commerce transactions and inquiries on their behalf or to link to their sites.

In the same way competition spurred the requirement that every company have a presence on the Web, the growth in number of voice sites will drive the demand to have an alternative way of reaching content and services that is convenient to people on the Voice Web.

## The Voice Web Value Chain

Industry players are coming together in their efforts to bring Voice Web services to consumers and businesspeople on the go. Effective partnering is a key success factor on the Voice Web because no one company can deliver everything that people and businesses will want in Internet time.

Various voice portal providers are distributing Voice Web services to people and businesses, while content and infrastructure players are supplying supporting technology, product, and hosting services to these portal providers.



The Voice Web value chain is made up of the following industry player categories:

Voice Portal Providers – Several categories of companies are offering or planning to market voice portal services to individuals and businesses:

*Web Portals* – Leading Internet portals that wish to extend their on-line reach to every phone and drive new revenues with a stronger affinity brand.

*Network Service Providers* – Telecom or Internet service providers who want to drive increased customer loyalty and network usage with branded portal services that they host in the network and/or obtain from third-party suppliers.

*Startup Voice Portals* – Companies whose fundamental business is building, hosting and marketing voice portal services targeted to particular audiences. These services may also be distributed through a network service provider with private-label agreement.

*Automobile* – Auto manufacturers are preparing to offer communication portals on wheels to Internet-generation drivers, to drive new streams of revenue from subscription and V-commerce services.

Content Players – Enterprises, e-commerce, and information services companies interested in expanding their service footprint at a lower cost. These companies are voice-enabling their own information as voice sites, or aggregating/syndicating various content to sell to voice portal providers.

Infrastructure Players – Technology software providers (speech recognition, location identification); hosting companies and service bureaus; Web developers (Web application servers, Web integration tools); hardware and platform providers; and wholesale network transport providers. These companies will pursue business-to-business relationships with voice portal providers.

## Opportunities on the Voice Web

There is enormous economic potential for all players on the Voice Web. This section highlights the business relationships and revenue opportunities that are being forged.

**Business-to-Business Opportunity:** Enterprise and e-commerce companies will use the Voice Web to expand their client base and grow revenue at lower cost. The Voice Web presents them the opportunity to sell their products and services through portal providers, and forgo costly advertising campaigns or expensive call centers to handle the transactions over the phone. Payment to portal providers will range from simple lead referrals to a percentage of the total commerce transaction. Portal providers will have supplier relationships with infrastructure and content players to purchase “Voice Web Ready™”<sup>7</sup> content and services; enabling technologies and products; hosting services; professional services; and wholesale network transport. Application Service Providers will host voice sites and voice portal services and sell product suites to merchants that make their retail applications Voice Web Ready. Startup portal providers will earn most of their revenues from V-Commerce transactions as well as from advertising, sponsorship, and third-party distribution contracts with network service providers.

**Business-to-Consumer Opportunity:** Voice portal providers will market directly to the public and manage the customer relationship with individuals and businesses that sign up for their service. To drive viral adoption, some of these portal providers will offer these initial services free-of-charge, thereby deriving the majority of their revenue from V-Commerce transactions, network usage, and advertising.

Ovum, a leading U.K. consulting firm, estimates the global voice portal services market will reach \$26 billion by 2005.<sup>8</sup> In the United States, analysts in the voice and mobile commerce arena estimate the market for voice portal services and technologies will exceed \$12 billion by 2005.<sup>9</sup>

The chart below summarizes revenue categories on the Voice Web for both business-to-business (B2B) and business-to-consumer (B2C) relationships.

Voice Web Player	Business Model	Software, Hardware, Platforms, Professional Services, Hosting, etc.	Advertising, Sponsorship	V-Commerce (lead referrals, transaction processing, purchased item)	Monthly Subscriptions (per service, per user, per bundled package, etc.)	Network Minutes
Portal Providers	B2B, B2C	✓	✓	✓	✓	✓
Content	B2B	✓	✓	✓		
Infrastructure	B2B	✓	✓	✓		✓

## Technologies Enabling the Voice Web

HTML and Netscape® Navigator ignited the growth of the World Wide Web by making it easy for developers to publish content that was then accessible by a mass audience through an easy -to-use graphical Web browser. The result is that more than 9.5 million Web sites have been published to date.<sup>10</sup>

Many companies seek to provide similar leadership in the Voice Web, providing standards-based technologies and products that will speed the construction of an interconnected Voice Web. A highlight of just some of these technologies and enabling products follows.

Speech Recognition – A robust and scalable software server engine that understands what someone is saying and then performs some action based on that understanding.

Voice Authentication – Provides secure Voice Web access, whether for enrollment identification or transaction authorization, using a person's biometric voiceprint. Voice authentication and the capability to securely store/utilize personal shopping information will make it convenient for busy consumers to do comparison shopping, make frequent purchases, or try new promotional offers.

Voice Browser\* – A standard user interface used by all industry participants, and to which all applications can conform. A standard voice browser such as Nuance Voyager™ enables navigation, context management and transaction completion for consumers on the Voice Web through standard facilities such as bookmarks, profiles, user authentication, etc. It is installed at the portal provider and not on the telephone itself.

Text-to-Speech (TTS) – Technology that converts machine-readable text into audible synthesized speech. TTS is necessary for large directories, frequently changing information, or when recording audio playback is deemed to be cost-prohibitive.

Audio – Audio content (live satellite broadcast feeds, streaming audio, syndicated radio, recorded sound, etc) that provides information output and helps create the branded persona of a Voice Web portal service.

VoiceXML – An emerging standard in markup languages for creating voice applications.

SpeechObjects™ – SpeechObjects are open standards-based, reusable speech application components that can be used independently and with VoiceXML to make rich Voice Web applications.

Developer Tools – A full range of tools will accelerate the growth of voice sites and Voice Web applications. Nuance V-Builder™ is one such tool, which uses VoiceXML and SpeechObjects to aid in the rapid development and deployment of voice sites.

Location Identification – Technology that determines the geographic position, direction of travel, and velocity of transmitters in mobile phones. Location identification will enhance Voice Web applications such as driving directions, V-Commerce, and local news and traffic.

\* To listen to real audio demonstrations of voice browsing capability, [click here](#).

## **Evolution of the Voice Web**

The Voice Web is in a nascent stage, but elements to create explosive growth are at hand. Voice portal providers are working with established infrastructure and content players to provide a wide array of information and services. Voice sites are being deployed today to support customer access to enterprise services and voice portals are launching almost daily. Sustainable business models are attracting generous venture capital, and standards-based technologies and products are laying the foundation for an interconnected Voice Web.

Ultimately the Voice Web will evolve to network service providers offering “intelligent dial tone” or immediate off-hook access to the Voice Web where services are available upon lifting the handset at home or upon turning on a mobile phone. With this milestone, the telephone network and the World Wide Web will be fully integrated such that consumers can seamlessly access the people and information they care about, 24 hours a day, from wherever they are.

There is an immense amount of opportunity on the Voice Web. Those that get involved now and help bring together the two most powerful networks—the telephone and the Internet— will surely reap the benefits.

## About Nuance

Nuance develops, markets and supports a voice interface software platform that makes the information and services of enterprises, telecommunications networks and the Internet accessible from any telephone. Every day, millions of people interact with Nuance systems at blue chip companies like Sprint PCS, American Airlines, Charles Schwab, The Home Shopping Network, Lloyds TSB, Sears and United Parcel Service. Nuance is also driving the creation of the Voice Web and delivering software for V-Commerce (voice-enabled e-commerce) services and applications. Nuance is headquartered in Menlo Park, Calif. with global sales offices and partners supporting solutions in multiple languages around the world. For more information about Nuance, visit [www.nuance.com](http://www.nuance.com) or call 1-888-NUANCE-8.

© 2000 Nuance Communications Inc. Nuance, Nuance Voyager, SpeechObjects, V-Commerce, V-Builder, and Voice Web Ready are trademarks or registered trademarks of Nuance Communications, Inc. All other trademarks, registered trademarks or service marks are the property of their respective owners.

## Endnotes

- 
- 1 Goldman Sachs Investment Research
  - 2 Yankee Group
  - 3 CyberAtlas.com
  - 4 February 1999 Evan Group Research study on deployed systems using Nuance software
  - 5 Goldman Sachs Investment Research estimates average savings per call at \$2.80
  - 6 "Foundations of Speech Portals" report, The Kelsey Group, February 2000
  - 7 A Voice Web Ready solution is built using standards-based VoiceXML and/or SpeechObjects, or other such other technologies such that the solution is rendered accessible via a standard voice browser, such as Nuance Voyager.
  - 8 "Personal Assistant Services" report by Dan Ridsdale Ovum October 18, 1999
  - 9 The Kelsey Group
  - 10 Iconocast, Winter 2000 Report