A wink, a wave, a touch, a glance, a word, and all your wishes will be granted.

W3C WORLD WIDE WEB *consortium*®

Multimodal Web

Web pages you can speak to and gesture at.

◾ The Multimodal Web will transform how we interact with applications:

• **In your hand**: portable access to multimedia communication, news and entertainment services.

• **In your car**: integrated dashboard system offering hands free navigation and infotainment services.

• **In your home**: remote control of your everyday appliances, including television, videorecorder, fridge, etc.

• **In your office**: choose how you interact with your computer, using a pen, keyboard or spoken commands.

For this, W3C is developing **standards** that support **multiple modes of interaction:** aural, visual and tactile. You can access the Web using your voice or your hands via a key pad, keyboard, mouse, or stylus. You will also be able to listen to spoken prompts and audio, and to view information on graphical displays.

◾ The Vision: W3C is developing the Multimodal Interaction Framework in order to:

• **extend the Web to enable multiple modes of interaction,**

• **augment Human to Computer and Human to Human interaction,**

• **improve Web usability on mobile devices.**

W3C Members involved:

Access, Agfa-Gevaert, Alcatel, Aspect Communications, AT&T, Avaya Communications, BeVocal, Boeing, Canon, Cisco Systems, Comverse Technology, Electronic Data Systems, Ericsson, France Telecom, Fraunhofer Gesellschaft Institute, Hitachi, HP, HeyAnita, IBM, INRIA, Intel, IWA/HWG, Kirusa, Loquendo, Matsushita Electric, Microsoft, MITRE, Mitsubishi Electric, MobileAware, NEC, Nokia, Nortel Networks, Nuance Communications, OnMobile Systems, Openstream, Opera Software, Oracle, ScanSoft, Sharp, Siemens, Sky Co., SnowShore Networks, Sun Microsystems, Telera, Tellme Networks, T-Online International, Toyohashi University of Technology, V-Enable, Verascape, Vocalocity, VoiceGenie Technologies, Voxeo.

MULTIMODAL WEB APPLICATIONS

Industries have a new exciting range of applications in the Multimodal Web to explore:

Mobile

Multimodal applications are of particular interest for mobile devices. Speech offers a welcome means to interact with smaller devices, allowing one-handed and hands-free operation. Users benefit from being able to choose which modalities they find convenient in any situation.
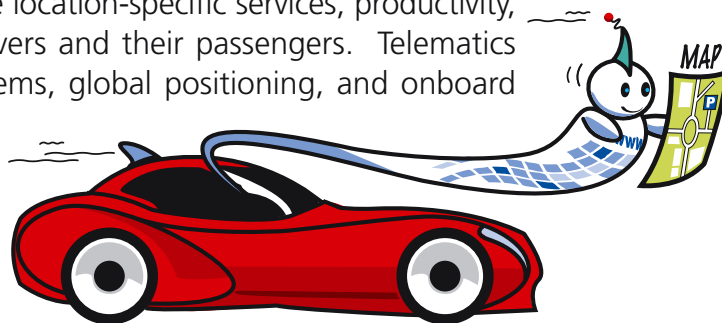
Interested companies are:
- **companies developing smart phones and personal digital assistants,**
- **companies developing browsers and authoring technologies,**
- **content providers for the Mobile Web.**

Automotive Telematics

Telematics is an emerging market of automotive communications technology that combines wireless voice and data to provide location-specific services, productivity, and in-vehicle infotainment services to drivers and their passengers. Telematics refers to the convergence of wireless systems, global positioning, and onboard automotive embedded electronics.

Interested companies are:
- **companies working on developing the next generation of in-car systems and e-vehicles.**

Ambient Intelligence

In addition to desktop access to the Web, Multimodal interfaces are expected to add value to remote control of home entertainment systems. Ambient Intelligence enables sensors, interactive screens, input devices for speech, handwriting and tactile information to directly interact with each house or outdoor device.
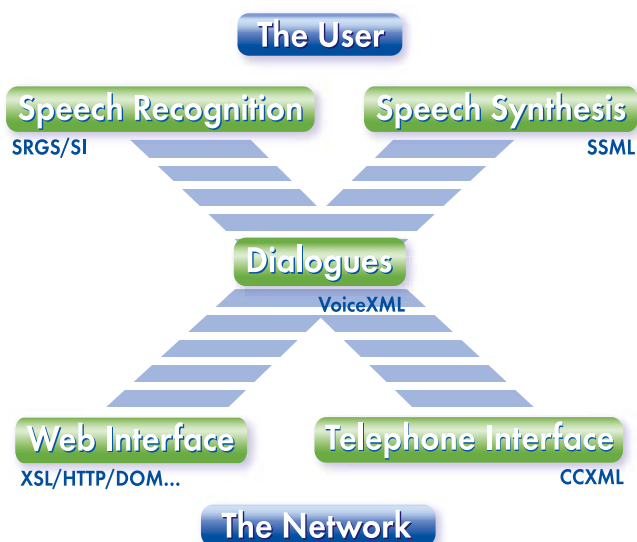
Interested companies are:
- **companies involved in developing embedded systems and consumer electronics.**

MULTIMODAL ACCESS

W3C is developing the Multimodal Interaction Framework. It is intended as a basis for developing multimodal applications with markup, scripting, styling and other resources.

Voice Interaction

Voice interaction can escape the physical limitations of keypads and displays as mobile devices become ever smaller. Voice provides an accessible alternative to using the keyboard or screen. This can be especially important in automobiles or in other situations where hands- and eyes-free operation is essential.

The User

Speech Recognition
SRGS/SI

Speech Synthesis
SSML

Dialogues
VoiceXML

Web Interface
XSL/HTTP/DOM...

Telephone Interface
CCXML

The Network

In the W3C Speech Interface Framework, W3C's VoiceXML controls how the application interacts with the user, while the Speech Synthesis Markup Language (SSML) is used for spoken prompts and the Speech Recognition Grammar Specification (SRGS) for guiding the speech recognizers via grammars that describe the expected user responses.

Other specifications in the Framework include Voice Browser Call Control (CCXML), which provides telephony call control support for VoiceXML and other dialog systems, and Semantic Interpretation for Speech Recognition (SI), which defines how speech grammars bind to application semantics.

Stylus Interaction

Complementing speech, a stylus can be used for handwriting, gestures, and drawings. Also, it will enable users to input specific notations for mathematics, music, chemistry and other fields, otherwise difficult to enter via a keyboard. Handwriting is expected to be popular for form filling and instant messaging on a mobile device.

Delivery Context

The framework enables applications to dynamically adapt to the current device capabilities, device configuration, user preferences and environmental conditions, such as low battery alerts or loss of network connectivity, muting the microphone and disabling audio output. Dynamic configurations include snapping a camera attachment onto a cell phone or bringing devices using a wireless network, e.g. a camera phone and a color printer.

W3C WORLD WIDE WEB *consortium* ®

Information Society Technologies