



# Voice eXtensible Markup Language (VoiceXML)



***Fabrizio Gramuglio***  
***VoiceXML Italian User Group***  
***[gramuglio@dotvocal.it](mailto:gramuglio@dotvocal.it)***





# VoiceXML Special Thanks to:



Programmi@mo

[www.programmiamo.net](http://www.programmiamo.net)

vaisulweb

[www.vaisulweb.it](http://www.vaisulweb.it)

**VOVOCAL**  
VOICE TECHNOLOGIES

[www.dotvocal.com](http://www.dotvocal.com)



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML

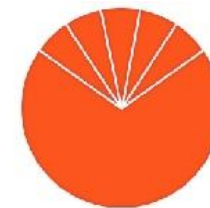


- **Corsi VXML presso eCLASSES.org**



- **Workshop**

- **Sabato 10/05 ore 16,30**
- **Domenica 11/05 alle 11,30**
- **Domenica 11/05 alle 11,30**



**Vacanze nel Mondo**  
*By i Viaggi del Ventaglio*

- **Sito VoiceXML Italian User Group**

**VOICE XML**  
ITALIAN USER GROUP



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML Index



## *Voice Applications / Index*

- **Why develop speech applications for the telephone?**
- **How speech applications work**
- **Introduction to VoiceXML**
- **Making the computer talk**
- **Making the computer listen**
- **A step forward: MultiModal website**
- **Voice-Commerce Application**



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML

## Why develop



*Voice Applications / Index*

- **Speaking and listening is fundamental for our lives**
- **Speaking and listening to a computer is natural and convenient**
- **When it may be inappropriate**
- **Speech applications challenges**
- **What can you do with a single phone call?**
- **Voice Portal applications**



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML Advantages



## Voice Applications / Features

- **Despite physical handicaps such as blindness or poor physical dexterity**
- **Bypass the limitation of small keyboards and screens**
- **Devices with no keyboard**
- **While callers work with their hands and eyes**
- **Anytime during the day**
- **Instant connection without being placed on "hold"**



# VoiceXML Advantages



## *Voice Applications / Features*

- **Using languages that do not lend themselves to keyboarding**
- **Convey emotion**
- **Access time-sensitive data**
- **Access location-sensitive data**
- **Access public and personal information**
- **Control computerized processes and activities**



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML Disvantages



## *Voice Applications / Features*

- **Hear graphics**
- **Point to edit**
- **Listen actively for a long period of time**
- **Overcome "Human memory"**
- **Be heard and hear in noisy environment**
- **Privacy**
- **Callers' false expectations**





# VoiceXML

## Type of Speech Aps



### Voice Applications / Features

	<b>Touchtone -only</b>	<b>Speech - Only</b>	<b>Touchtone and speech</b>	<b>Multimodal</b>
<b>Speech in</b>	No	Yes	Yes	Yes
<b>Keypad</b>	Yes	No	Yes	Yes
<b>Mouse and keyboard</b>	No	No	No	Yes
<b>Speech out</b>	Yes	Yes	Yes	Yes
<b>Display monitor</b>	No	No	No	Yes



# VoiceXML Applications



## Voice Applications / Examples

- **Mass-Market voice portals:**

E-mail, driving instructions, traffic conditions, weather, telephone number search, business reminders, local information services, stock quotes, personalized news, sport news, horoscopes

- **Corporate external voice portals:**

Call centers, telephone attendant, product and service information, order entry, troubleshoot product failure, banking

- **Corporate internal voice portals:**

Customer relationship management, product availability and pricing, Order status, SCM, Human resource informations



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML

## Introducing VXML



### *Voice XML / Features*

- **A markup language for building speech interface**
- **Simplifies application development**
- **Minimizes internet traffic interaction**
- **Separates user interaction code from application logic code**
- **Provides portability and simplicity**
- **Supports rapid prototyping and interactive refinement**



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



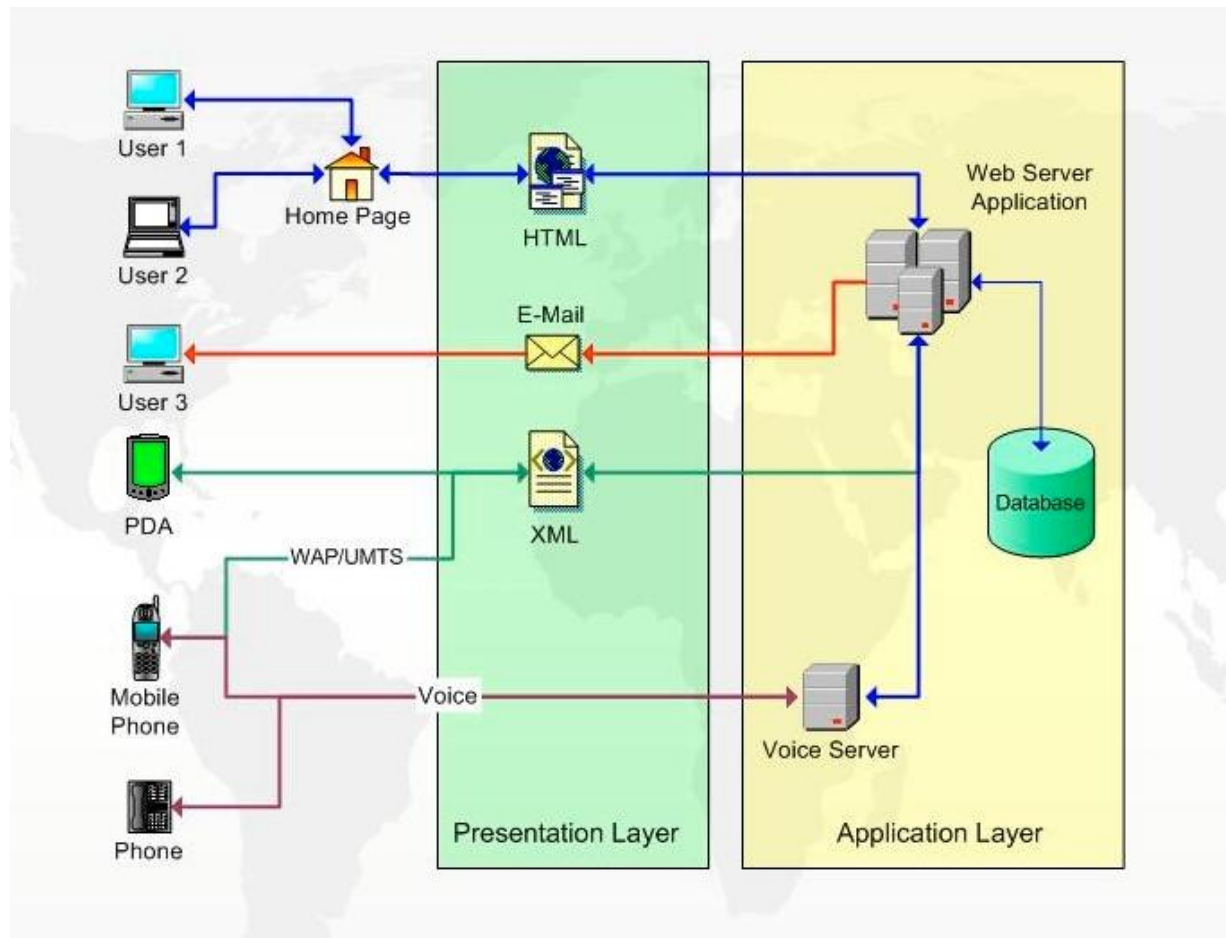
WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



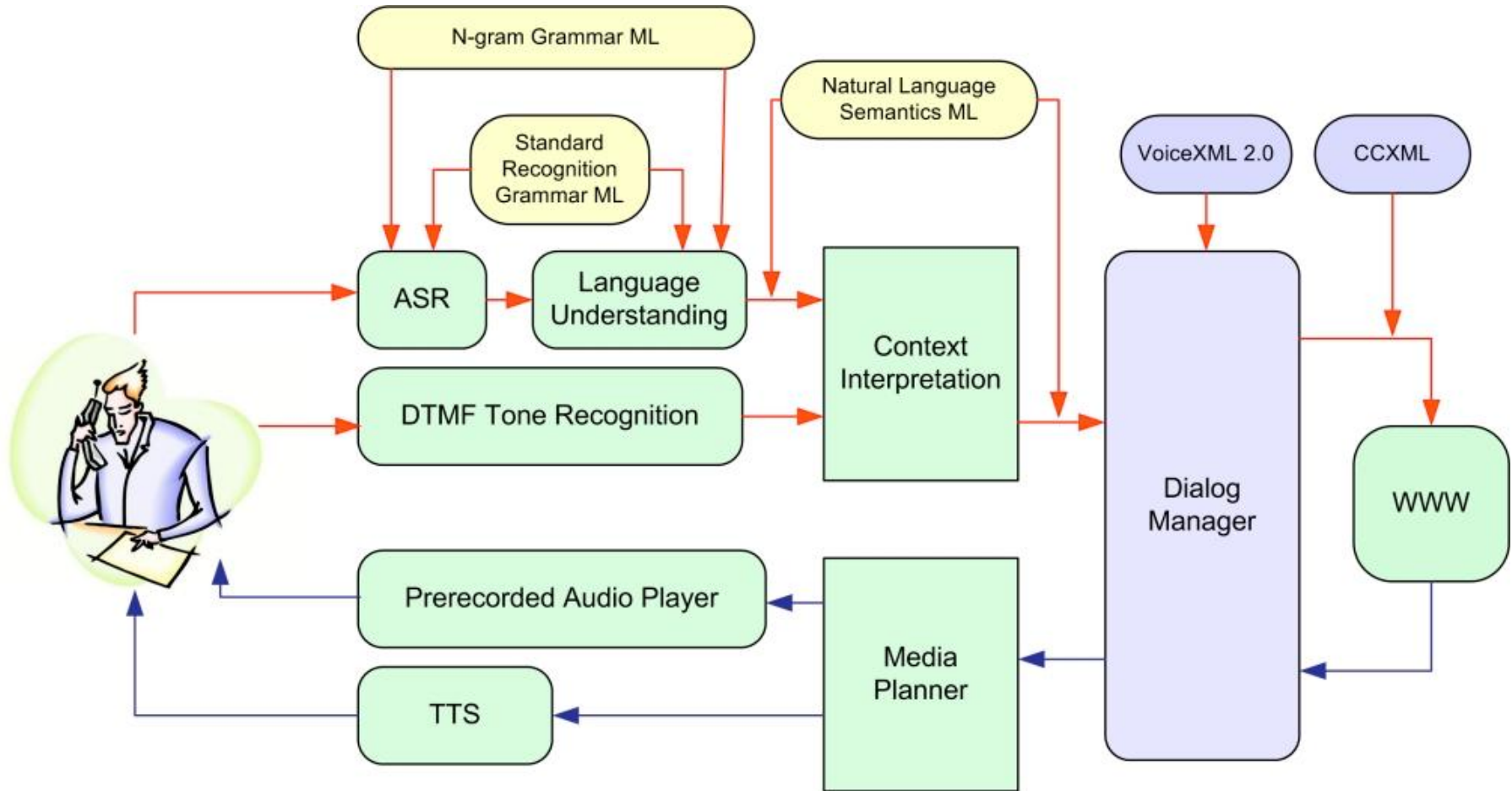
# VoiceXML Architecture



## Voice Applications / Architecture



## Voice Applications / Architecture





# VoiceXML Genealogy



## Voice XML / Introduction

- Standardization body
- Develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential



[www.w3.org](http://www.w3.org)



- Industry organization founded by AT&T, IBM, Lucent and Motorola.
- Inventors and promoters the Voice Extensible Markup Language (VoiceXML)

[www.voicexmlforum.com](http://www.voicexmlforum.com)



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML Overview



## Voice XML / Introduction

- Many companies involved
- Working Groups:
  - About 39 groups, mainly related to:
    - HTML evolution
    - XML technologies
    - Voice technology integration
    - Accessibility
    - I18N
  - Working for stable specification
- Voice Browser Working Group (VBWG) - Voice enabling the Web
  - [www.w3.org/voice](http://www.w3.org/voice)
- Multimodal Interaction (MMI) – Anywhere, Any device, Any time
  - [www.w3.org/2002/mmi](http://www.w3.org/2002/mmi)



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML Brief History



## *Voice XML / Introduction*

- 1999: Forum submit the VoiceXML language to the W3C
- 2000: Forum VoiceXML 1.0 version
- 2000-2002: W3C VBWG VoiceXML 2.0 specs working on
- 2001: Forum review journal, meetings, and tools
- 2003: W3C VBWG VoiceXML 2.0 expected Recommendation
- 2003: Forum quality mark



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>





# VoiceXML How to develop



*Voice XML / Development*

- **Separate GUI and VUI**
- **HTML verbal browse**
- **Transcoder**
- **Combined GUI and VUI**



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



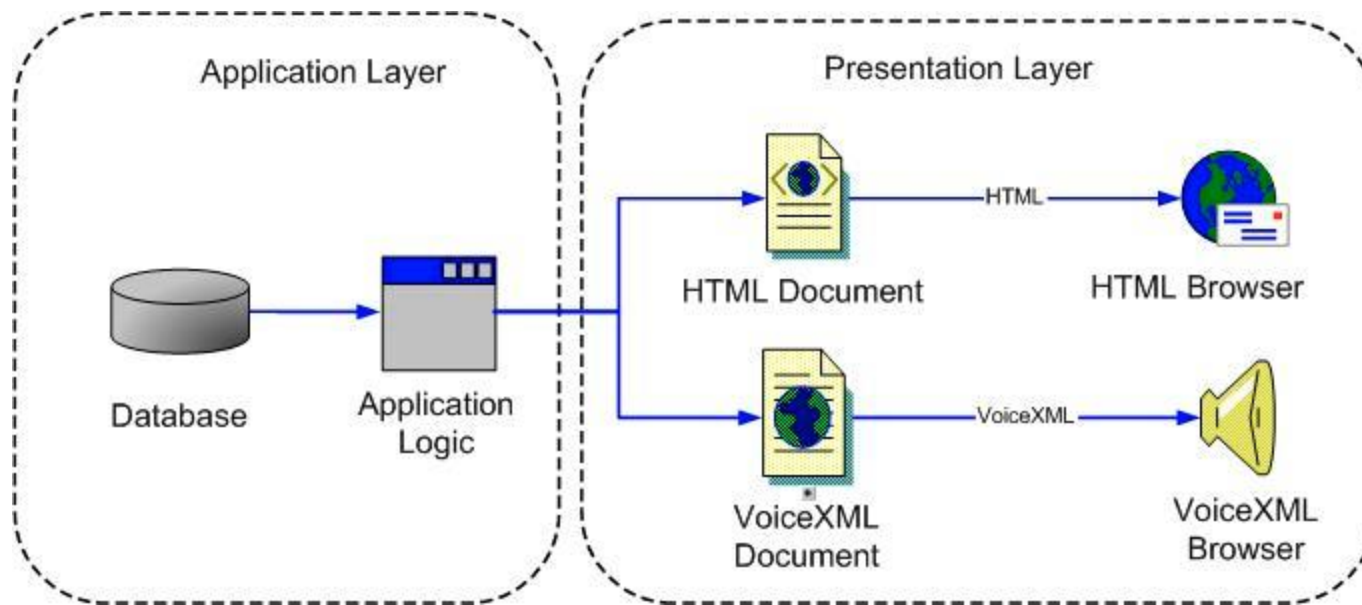
WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



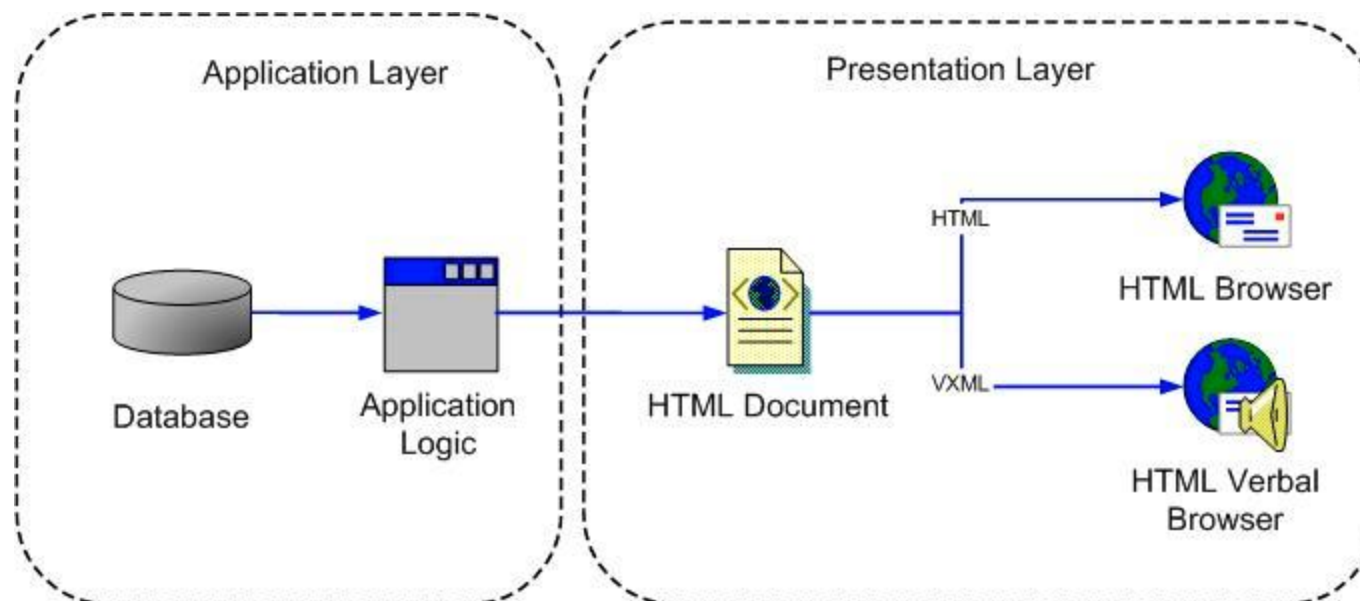
# VoiceXML Separate



## Voice XML / Development



## Voice XML / Development



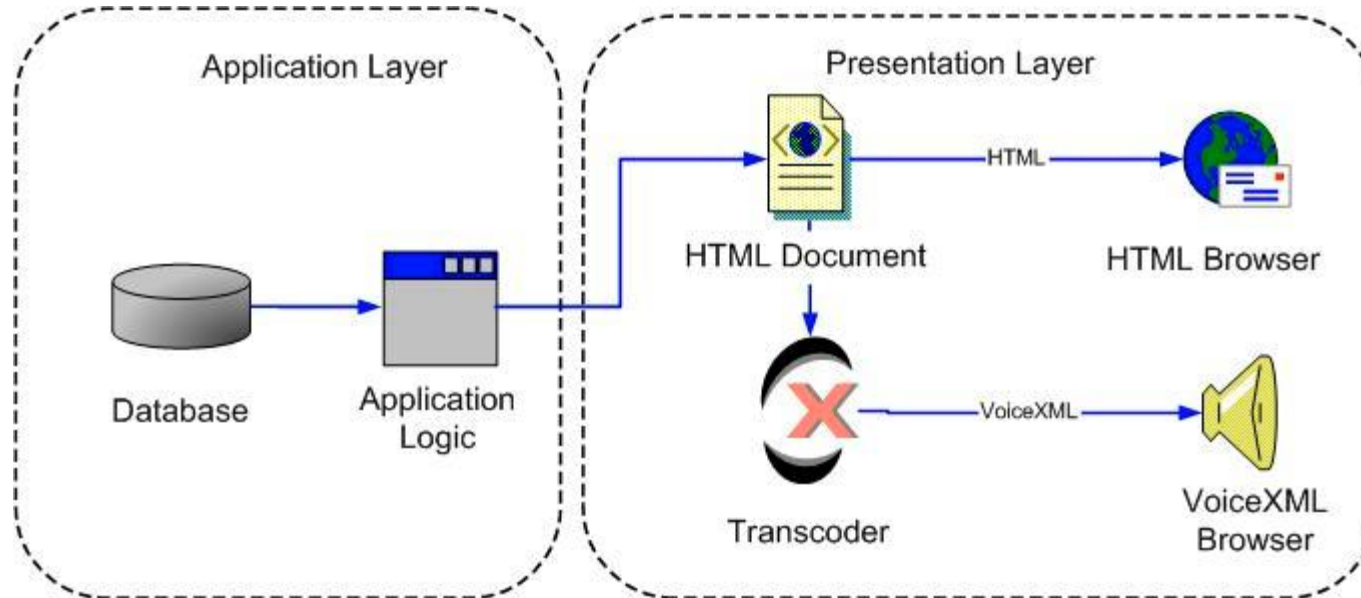
- **Extract required informations**
- **Transform visual control to verbal control**
- **Sequence the presentation of controls**



# VoiceXML Transcoding



## Voice XML / Development



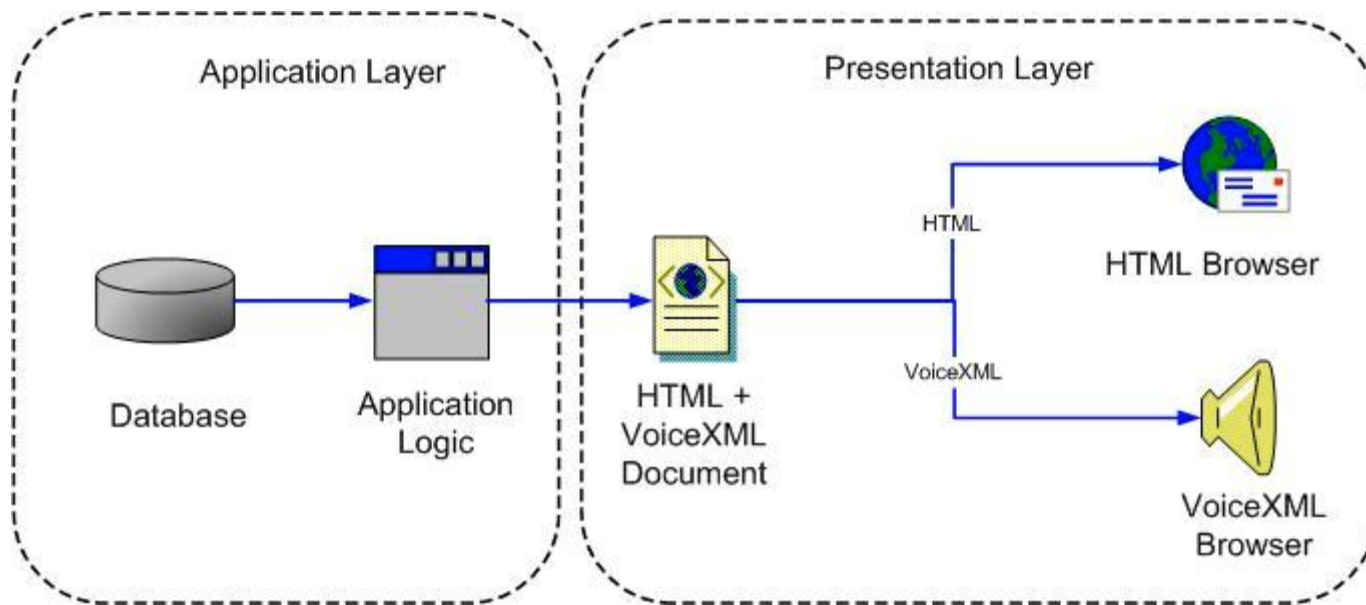
- **Customize the verbal interface**
- **Extend the grammars to include more words**
- **Event handlers and more instructions for TTS**



# VoiceXML Combined



## Voice XML / Development





# VoiceXML How to develop



## Voice XML / Development

	Separate	HTML-V	Transcoder	Combined
<b>VUI Quality</b>	High	Low	High, if customized	High
<b>Sensitive html changes</b>	No	Yes	Yes	Only GUI
<b>Effort to create VUI</b>	High	Low	Medium, if customized	Very High



# VoiceXML Introduction



## Voice XML / Genealogy

```
<?xml version="1.0"?>
<vxml version="2.0" mode="TTS">
<!-- This example -->
<form>
  <block>
    <prompt>Hello World. This is my
    first VoiceXML application
  </prompt>
</block>
</form>
</vxml>
```



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>



# VoiceXML Speech



## Voice XML / Features

### Text To Speech

- Design the persona
- Design the sound
- Earicons
- Improve tts (pause, emphasize)

### Prerecorded Audio

- Improve prerecorded speech



Fabrizio Gramuglio  
[www.vxmlitalia.com](http://www.vxmlitalia.com)  
[gramuglio@dotvocal.com](mailto:gramuglio@dotvocal.com)



WEBBIT 2003  
Seminari IWA/HWG  
<http://webbit.iwa-italy.org>





# VoiceXML Prosody



## Voice XML / Features

<b>&lt;audio&gt;</b>	<b>Replay a prerecorded file</b>
<b>&lt;break&gt;</b>	<b>Insert a quite pause</b>
<b>&lt;emphasis&gt;</b>	<b>Emphasize word</b>
<b>&lt;paragraph&gt;</b>	<b>Identify a paragraph</b>
<b>&lt;prosody&gt;</b>	<b>Specify values for volume, pitch and rate</b>
<b>&lt;sayas&gt;</b>	<b>Specify pronunciation</b>
<b>&lt;sentence&gt;</b>	<b>Identify a sentence</b>
<b>&lt;voice&gt;</b>	<b>+ age, gender</b>



# VoiceXML Listen



## Voice XML / Features

	Usage Model	Requirement
<b>Touchtone menu</b>	Numeric data entry	DTMF Tone recognition
<b>ASR Menu</b>	AlphaNumeric data entry	Conversational speech recognition
<b>Forms</b>	Data Entry	Conversational speech recognition
<b>Query</b>	Caller formulate the query	ASR-large vocabulary, TTS
<b>Dictation</b>	Caller dicatate sentences	Dictation RS and large vocabulary
<b>Mixed Initiative</b>	Dialog	Speech recognition – small vocabulary



# VoiceXML Grammars

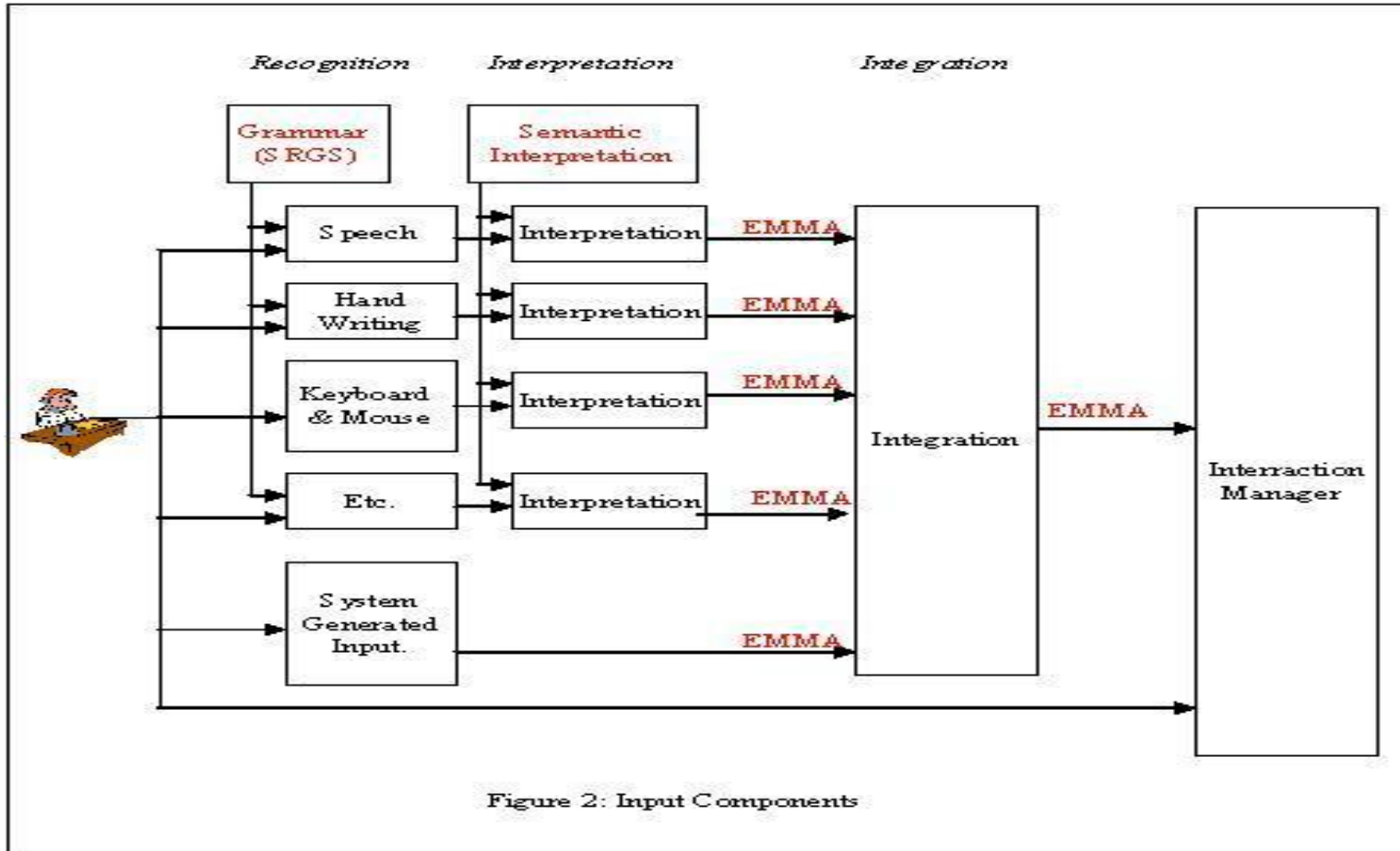


## Voice XML / Development

```
45. <!--
46. In-line Grammars would need to be built based on your company
47. directory. You must enter one of the names listed in the
48. in-line grammar shown or replace the grammar with your own.
49. -->
50. <form id="without">
51.   <field name="you">
52.     <grammar>Casey|Eric|Jan|John</grammar>
53.     <prompt>Please say or enter a name</prompt>
54.     <prompt>Your choices are: Casey, Eric, Jan or John.</prompt>
55.     <nomatch count="1">I don't understand. The program will
56.       now exit.
57.       <goto nextitem="end"/>
58.     </nomatch>
59.     <noinput count="1">I have received no input. The program will
60.       now exit.
61.       <goto nextitem="end"/>
62.     </noinput>
63.     <filled>
64.       <prompt>I heard<value expr="you"/></prompt>
65.     </filled>
66.   </field>
67. </form>
```

# VoiceXML MM Input

## MultiModal



## MultiModal

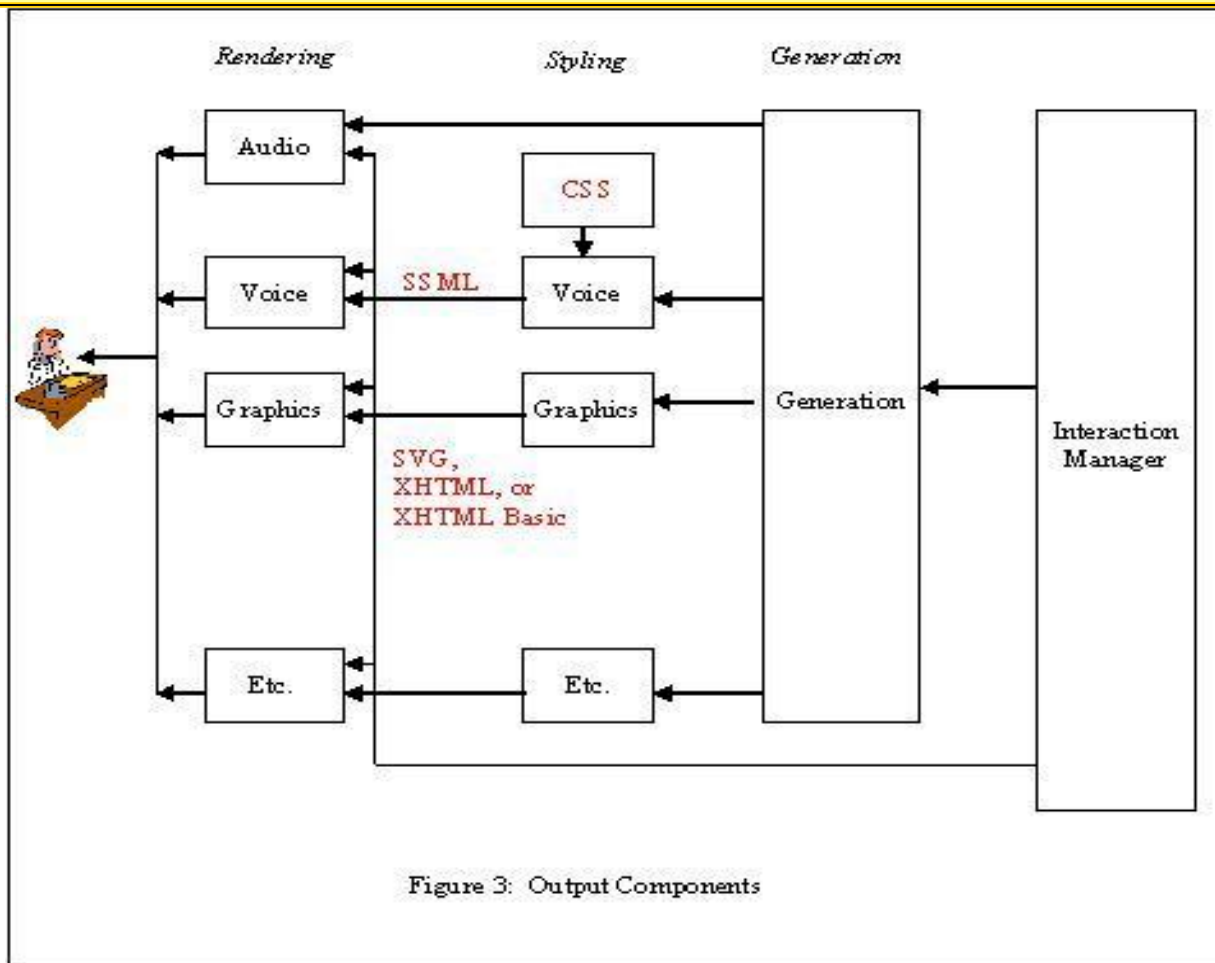


Figure 3: Output Components



# VoiceXML MM Driving Ex



## MultiModal

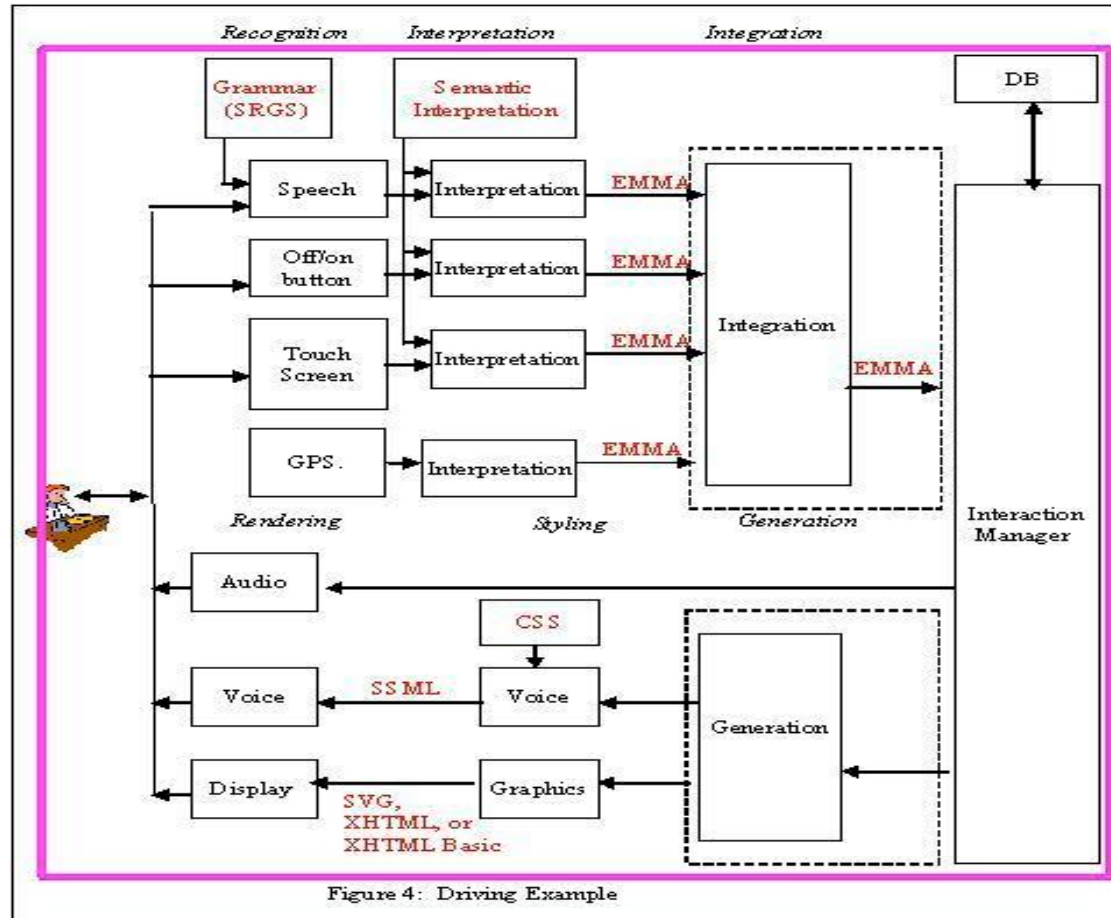


Figure 4: Driving Example