Deutsche Telekom Laboratories

The W3C Multimodal Architecture
W3C Multimodal Architecture.
Agenda.

- The W3C
- W3C Multimodal Interaction Working Group
- W3C Multimodal Architecture (overview)
- Upcoming web standards
  - SCXML (State Chart XML)
  - VoiceXML 3.0
  - EMMA (Extensible Multimodal Annotation Markup Language)
- Context awareness in multimodal applications
- A practical implementation of the W3C Multimodal Architecture
- Conclusion
The W3C.
W3C Multimodal Architecture.
The W3C – Mission and Goals.

- Mission: To lead the **World Wide Web** to its **full potential** by developing protocols and guidelines that ensure long-term growth for the Web.
- The main goal: To ensure **interoperability of web standards**, irrespective of special interests.
W3C Multimodal Architecture.
The W3C – Governance and Statistics.

- Web standards creation
  - Since 1994 some **110 Web standards** have been published.
- Infrastructure
  - **International** representation (ERCIM*, MIT**, Keio Uni.***) with
    - **400 member organizations** from 40 countries
- Accountability
  - **Vendor-neutral** forum
  - Public discussion with full consideration of **public feedback**
  - **Participation in working groups**, as member or invited expert
- Patent policy
  - Strong patent policy to **protect company and public interests**

---

*European Research Consortium for Informatics and Mathematics (ERCIM)  
**MIT Computer Science and Artificial Intelligence Laboratory (CSAIL)  
***Keio University

Deutsche Telekom Laboratories

Nov 19, 2008 5
W3C MMI Working Group.
W3C Multimodal Architecture.
Upcoming web standards.
W3C Multimodal Architecture.
W3C Multimodal Interaction Working Group.

- Mission: “... to extend the Web to allow users to dynamically select the most appropriate mode of interaction for their current needs ...“

- Participants
  - 40 participants from 22 organizations

- More information: http://www.w3.org/2002/ bmi
W3C Multimodal Architecture.
Multimodal Interaction Framework.

- W3C Multimodal Interaction Framework (published 2003)
"Multimodal Architecture and Interfaces" (working draft status)

- Runtime Framework provides the basic infrastructure and controls communication among the constituents.
- Interaction Manager (IM) coordinates modality components (MCs) by life-cycle events and contains the shared data (context).
- Event-based communication between IM and MCs.
“Multimodal Architecture and Interfaces“ (working draft status)

- MMI lifecycle events
  - newContextRequest/Response
  - startRequest/Response
  - cancelRequest/Response
  - prepareRequest/Response
  - statusRequest/Response
  - pauseRequest/Response
  - resumeRequest/Response
  - Done
  - extension
W3C Multimodal Architecture.
State Chart XML (SCXML).

“State Chart XML (SCXML): State Machine Notation for Control Abstraction”
(working draft status)

```xml
<scxml initial="foo" ...>
  <datamodel>
    <data id="x" expr="'1'"/>
  </datamodel>
  <state id="foo">
    <transition event="foo"
      target="bar">
      <send event="play" target="VUI"
        targettype="basichttp"/>
    </transition>
  </state>
  <state id="bar">
    ...
  </state>
</scxml>
```
W3C Multimodal Architecture.
Voice Modality Component (VoiceXML).

W3C VoiceBrowser Working Group

- Existing standards:
  - VoiceXML 2.1
  - CCXML 1.0
  - SRGS 1.0 and SISR 1.0
  - PLS 1.0 and SSML 1.0

- Work in progress:
  - SCXML 1.0
  - VoiceXML 3.0
    - External eventing
    - New features (e.g. speaker verification)
    - Alignment with MMI architecture

* Note: VoiceXML 3 syntax not yet finalized, example shows principle.
“EMMA: Extensible MultiModal Annotation markup language”

- EMMA 1.0 ([www.w3.org/TR/emma](http://www.w3.org/TR/emma)) currently in Candidate Rec. status
- Goal: Annotation/representation of user input
- Example: user utterance “flights from Boston to Denver“

```
<emma:emma version="1.0" xmlns:emma="http://www.w3.org/2003/04/emma" ...>
  <emma:one-of id="r1" emma:start="1087995961542" emma:end="1087995963542"
    emma:medium="acoustic" emma:mode="voice">
    <emma:interpretation id="int1" emma:confidence="0.75"
      emma:tokens="flights from boston to denver">
      <origin>Boston</origin>
      <destination>Denver</destination>
    </emma:interpretation>
  </emma:one-of>
</emma:emma>
```
Implementation of the W3C Multimodal Architecture using W3C markup languages.
Context Awareness in multimodal applications.
W3C Multimodal Architecture. Related W3C Standards.

- The W3C has developed other standards which are interesting in the context of multimodal applications. e.g.:

  - DCCI (Delivery Context Client Interfaces)
    - http://www.w3.org/TR/DPF
    - Accessing Static and Dynamic Delivery Context Properties

  - Device Description Repository
    - http://www.w3.org/TR/DDR-Simple-API
    - Simple API to access device information
Delivery Context Client Interfaces (DCCI): Client interface to access (static and dynamic) device information, e.g. loudspeaker volume.
Related activities with regards to Multimodal Interaction.
W3C Multimodal Architecture. Upcoming Activities.

- Emotion Incubator Group:
  - Annotation of Emotions ([http://www.w3.org/2005/incubator/emotion](http://www.w3.org/2005/incubator/emotion))
  - Goal: Standardization of Emotion Markup Language (EmotionML)

- Model based User Interfaces Incubator Group:
  - Evaluation of research on model-based user interface design ([http://www.w3.org/2005/incubator/model-based-ui](http://www.w3.org/2005/incubator/model-based-ui))
  - Source of further information: UsiXML ([http://www.usixml.org](http://www.usixml.org))
W3C Multimodal Architecture. Upcoming Activities.

- Geolocation Working Group:
  - Provisioning of location information within web browser environments
    (http://www.w3.org/2008/geolocation)

- Rich Web Application Backplane Incubator Group:
  - Rich application backplane
    (http://www.w3.org/2005/incubator/app-backplane)
  - Definition of a set of common building blocks for web applications
A practical implementation of the W3C Multimodal Architecture.
W3C Multimodal Architecture.
Proof of Concept Implementation.

- Distributed implementation using SCXML, HTML and CCXML/VXML
W3C Multimodal Architecture.
Proof of Concept Implementation.

- Apache commons SCXML interpreter + ECMAScript interpreter.
- VoiceXML/CCMXML platform and HTML browser with AJAX support.
Conclusion.
W3C Multimodal Architecture.
Conclusion.

- Multimodal architecture specification (work in progress)
  - Loosely coupled architecture
  - Allows for co-resident and distributed implementations
  - Leverages existing W3C standards (HTML, SVG, ...)

- EMMA 1.0 specification (annotation of user input) completed
- SCXML will be a candidate for Interaction Manager implementations
Thank you for your attention.

Ingmar Kliche
Deutsche Telekom Laboratories
10589 Berlin
Goslarer Ufer 35
Tel.: +49 30 3497 2348
Email: ingmar.kliche@telekom.de