Multimedia Content Representation and Adaptation for the Mobile Devices: a Survey of SMIL/MMS/PSS Technologies for 3GPP/3GPP2

Tayeb LEMLOUMA

November 2004

WAM Research Team, INRIA Rhône-Alpes

Plan

Introduction

- Multimedia systems become more and more heterogeneous
- Several heterogeneous devices are used today
- Different complex applications and content exist on the servers side
- Increasing need to use the content using small devices and in non classical situation (example in mobility)
- Problem: mobile devices are different and subject of many limitations
- How can we enable the use of the Web and multimedia applications on limited terminals?

Mobiles Generation, from 1G to 4G

- **First Generation** (1G) :
  - 70/80 years, first wireless terminals, analogical radio phones. Big size, security problems, techniques close to the those used in FM radios.
- **Second Generation** (2G) :
  - Beginning of 90 years, best quality, reduced size. Based on GSM (Global System for Mobile Communication). In France: frequency = 900-1800 MHz, speed = 9,6 Kbps
  - Several users, 70% population in France (2004)

Mobiles Generation, from 1G to 4G

- **Third Generation** (3G) :
  - GPRS (General Packet Radio Service, 2.5G) offers a speed of 20-30 Kbps. Allows to access to the network more easily (Internet, email without attachment, surfing..). Voice over GSM and Data over GPRS.
  - Amelioration: EDGE (Enhanced Data Rate for GSM Evolution). Speed: up to 250 Kbps.
  - **3G**: After several ameliorations, UMTS/W-CDMA is adopted. W-CDMA uses the frequency band 1900-2200 MHz.
  - Allows to send in the same time all the data in packets. A speed up to 2Mbps (from a fixed position) and 384 Kbps in mobility. Advantages in voice and data transfer. Best quality of services. Multimedia applications can be used.

Mobiles Generation, from 1G to 4G

- **Forth Generation** (4G) :
  - Will appear around 2007-2010 in Japan.
  - Some experimentations are already done on OFDM (Orthogonal Frequency Division Multiplexing), will allow to reach normally a speed of 300Mbps
  - For the near future, ameliorations of UMTS: HSDPA (High Speed Downlink Package Access) that allows to reach really the 2Mbps promised by UMTS
  - The adoption of the 3.5G will be done around 2005 in Japan and later in Europe

MMS : SMIL on Mobiles

- **MMS** (Multimedia Messaging Service) is **SMIL** (Synchronized Multimedia Integration Language)
- **SMIL** is a W3C standard that allows authors to write interactive multimedia presentations.
  - Using SMIL, an author can describe the temporal behavior of a multimedia presentation, associate hyperlinks with media objects and describe the layout of the presentation on a screen.
- SMIL can be viewed by the user as a live stream in PSS or stored locally on the mobile phone
- SMIL streaming protocols in PSS can be RTP, RTCP, RTSP, SDP

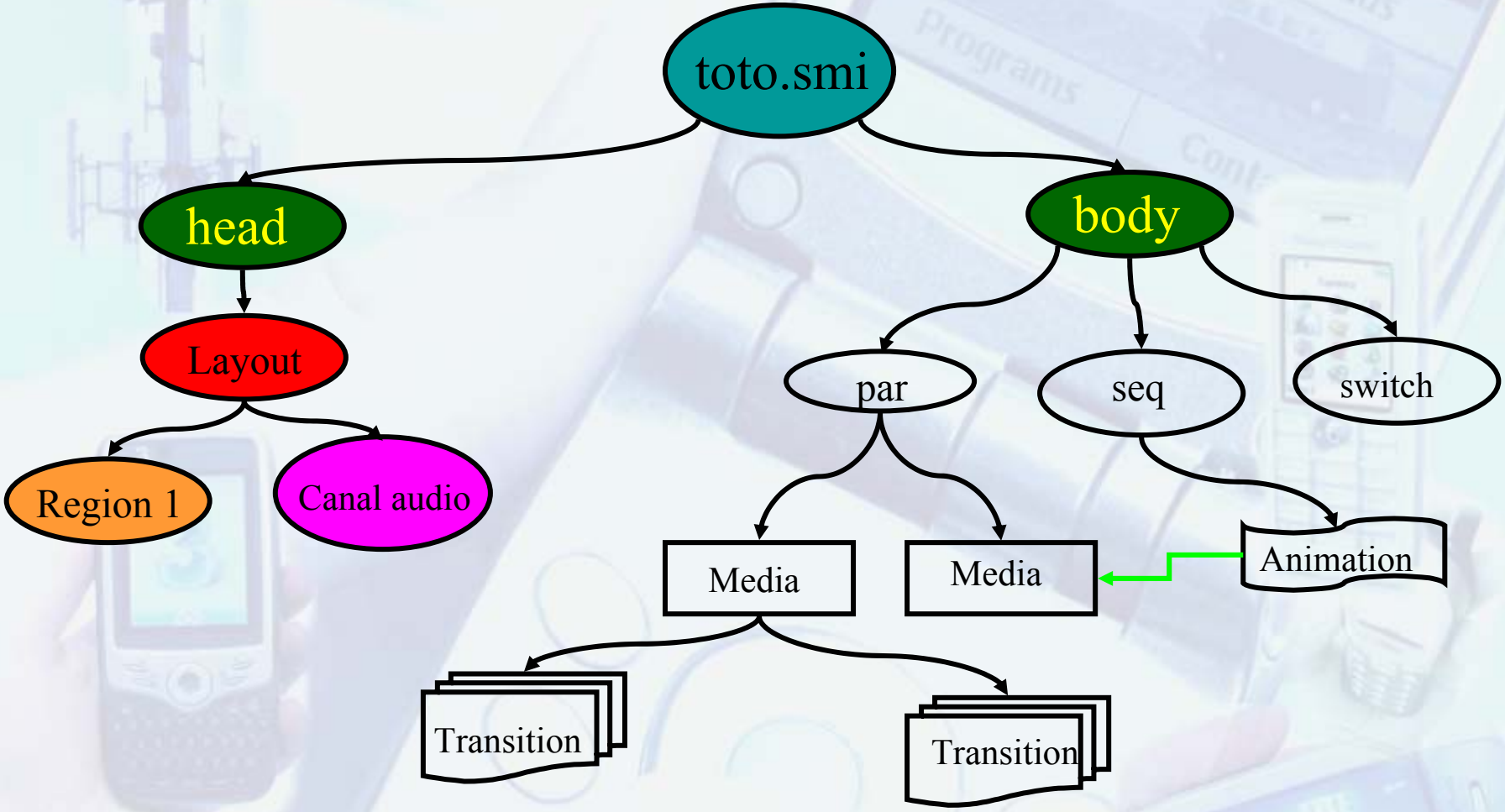
Examples :

<video src="**rtsp**://www.example.com/video1" begin="1" dur="120"/>
<video src="**http**://www.example.com/video2" begin="1" dur="120"/>

Organization of a SMIL document

- **Two parts**:
- **Head**: contains information (meta and layout information, author-defined content control) that is not related to the temporal behavior of the presentation.
- **Body**: contains information that is related to the temporal and linking behavior of the document

Document Structure



```
<smil xmlns="http://www/w3.org/2000/SMIL20/Language">
  <head>
   <layout type="text/smil-basic">
            <region id="left-video" left="20" top="50" z-index="1"/>
            <region id="left-text" left="20" top="120" z-index="1"/>
            <region id="right-text" left="150" top="120" z-index="1"/>
   </layout>
  </head>

  <body>
   <par>
            <seq>
                        <img src="graph" region="left-video" dur="45s"/>
                        <text src="graph-text" region="left-text"/>
            </seq>
            <par>
                        <a href="http://www.w3.org/People/Berners-Lee">
                                    <video src="tim-video" region="left-video"/>
                        </a>
                        <text src="tim-text" region="right-text"/>
            </par>
            <seq>
                        <audio src="joe-audio"/>
                        <video id="jv" src="joe-video" region="right-video"/>
            </seq>
     </par>
   </body>
</smil>
```

Head

Body = scenario

Packet-Switched Streaming Services (PSS)

- Standard defined by the Third Generation Partnership (3GPP)
- Main focus: communication between servers and terminals
- Covers protocols and codecs necessary for streaming content delivery to 3G multimedia terminals
- PSS: how the server and the terminal must behave in order to enable streaming services
- Application examples :
  - Using on-demand audiovisual content
  - Viewing content from live sources, example a live camera
  - Streaming content over Multimedia Messaging Service (MMS)

3GPP Profile of SMIL

- 3GPP SMIL profile is based on SMIL Basic profile and SMIL Scalability Framework

- The profile uses the following modules:

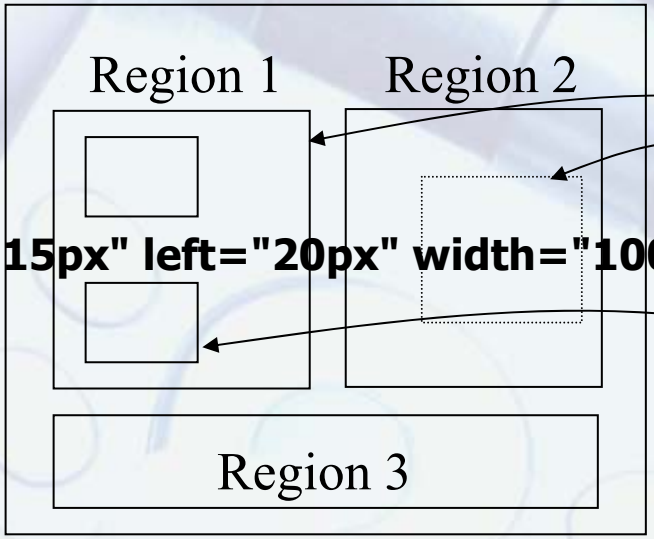
SMIL 2.0 Structure Module	Structure
SMIL 2.0 Layout Module	BasicLayout
SMIL 2.0 Metainformation Module	Metainformation
SMIL 2.0 Media Object Modules	BasicMedia, MediaClipping, MediaParameter, MediaAccessibility and MediaDescription
SMIL 2.0 Linking Module	BasicLinking, LinkingAttributes
SMIL 2.0 Timing and Synchronization Modules	BasicInlineTiming, MinMaxTiming, BasicTimeContainers, RepeatTiming and EventTiming
SMIL 2.0 Content Control Modules	BasicContentControl, SkipContentControl and PrefetchControl

Structure Module

- The Structure module defines the top-level structure of the document.
- The elements of this module are: *smil*, *head* and *body*
- *smil: acts as the root element for the SMIL document*
- *head*: contains information that is not related to the temporal behavior of the presentation.
- *body*: related to the temporal and linking behavior of the document. It acts as the root element of the timing tree.

Basic Layout Module

- Includes a layout model for organizing media elements into regions on the visual rendering surface.
- Media elements declare which region they are to be rendered into with the region attribute




```
<layout>
<region id="r" top="15px" left="20px" width="100px" height="50px"/>
</layout>
```

Metainformation

- Metadata is "data about data", for example, a library catalog is metadata, since it describes publications; data describing Web resources, ..

- This module contains elements and attributes that allow the description of SMIL documents.

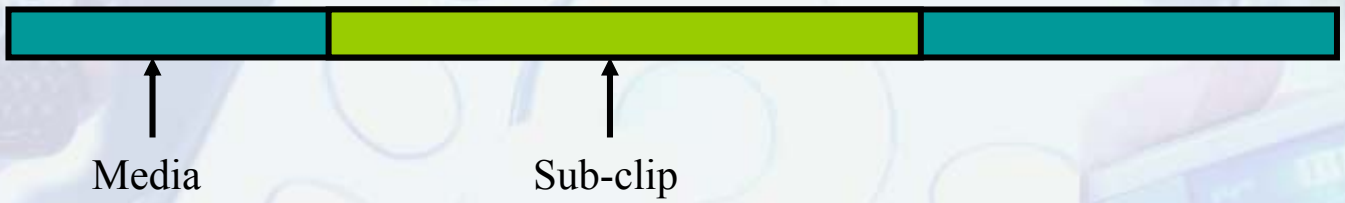
Example

```
<!-- Metadata about the SMIL presentation -->
<rdf:Description about="http://www.example.com/meta.smi" dc:Title="An
    Introduction to the Resource Description Framework" dc:Description="The
    Resource Description Framework (RDF) enables the encoding, exchange and
    reuse of structured metadata" dc:Publisher="W3C"
    dc:Date="1999-10-12" dc:Rights="Copyright 1999 John Smith"
    dc:Format="text/smil" >
<dc:Creator>
    <rdf:Seq ID="CreatorsAlphabeticalBySurname">
          <rdf:li>Mary Andrew</rdf:li>
          <rdf:li>Jacky Crystal</rdf:li>
  </rdf:Seq>
</dc:Creator>
<smilmetadata:ListOfVideoUsed>
    <rdf:Seq ID="VideoAlphabeticalByFormatname">
      <rdf:li Resource="http://www.example.com/videos/meta-1999.mpg"/>
      <rdf:li Resource="http://www.example.com/videos/meta2-1999.mpg"/>
     </rdf:Seq>
</smilmetadata:ListOfVideoUsed>
<smilmetadata:Access LevelAccessibilityGuidelines="AAA"/>
</rdf:Description>
```

Media Object Module

- Contain elements and attributes used to describe media objects and allow the inclusion of media objects into a SMIL presentation.
- 3GPP SMIL profile includes media elements from BasicMedia module and attributes from *MediaClipping*, *MediaAccessibility* and *MediaDescription*
- *MediaClipping:* *address sub-clips of continuous media*

```
<video id="a" src="vid1.mpg"
  clip-begin="smpte=00:01:45"
  clip-end="smpte=00:01:55"
  …
/>
```



Media Object Module

- *MediaAccessibility:* basic accessibility support for media elements using the attributes:
  - alt: short description of the media
  - longdesc: long description
  - readIndex: order in which *longdesc*, *title* and *alt* text are read aloud by assistive devices

```
<par>
<video id="carvideo" src="car.rm" region="videoregion" title="Car video" alt="Illustration of
relativistic time dilation and length contraction." longdesc="carvideodesc.html" readIndex="3"/>
<audio id="caraudio" src="caraudio.rm" region="videoregion" title="Car presentation voiceover"
begin="bar.begin"/>
<animation id="cardiagram" src="car.svg" region="animregion" title="Diagram of the car"
readIndex="2"/>
<img id="scvad" src="scv.png" region="videoregion" title="Advertisement for Sugar Coated
Vegetables" readIndex="1"/>
</par>
```

Media Object Module

- *MediaDescription* is included by the *MediaAccessibility* module and adds 'abstract', 'author' and 'copyright' attributes to media elements.
- *MediaParameter:* allows passing additional parameters to the rendering of a media object.
- 3GPP SMIL Profile (Release 6, june 2004) adds the following attributes: color, foreground-color, font-size and textsize
- Concerned medias are: application/text, application/xhtml+xml, application/vnd.wap.xhtml+xml

Linking

- Used to provide hyperlinks between documents and document fragments.
- BasicLinking: uses the elements *a* et *area*
  - ***a***: very similar to the functionality of the a element in HTML
  - ***area***: extends the syntax and semantics of the HTML area. It allows breaking up an object into temporal subparts (ex. using begin and end)

Example:

```
<video src="video" title="Interview" >
<area shape="rect" coords="5,5,50,50" title="Journalist"
    href="http://www.example.org/journalist"/>
<area shape="rect" coords="60,5,100,50" title="Subject"
    href="http://www.example.org/subject"/>
</video>
```

Timing and Synchronization

- ***BasicInlineTiming***: defines the attributes that make up basic timing support for adding timing to XML elements. (dur, begin and end)
- ***MinMaxTiming***: defines the attributes that allow setting minimum and maximum bounds on element active duration. (max and min)
- ***BasicTimeContainers***: defines basic time container elements, attributes that describe an element's display behavior within a time container, and end conditions for time containers. (par and seq)
- ***RepeatTiming***: defines the attributes that allow repeating an element for a given duration or number of iterations. (repeatDur, repeatCount and repeat)
- ***EventTiming***: defines the attribute value syntax for begin and end attributes that allow elements to begin and end in response to an event. (begin and end with event values)

Timing and Synchronization

- **Example of Simple timing within a Parallel time container**:

- All the medias children of a par begin by default when the par begins.

- <par>
-
-
-
- </par>

Timing and Synchronization

- **Example of exclusive time container with link-based activation:**

```
<par>
   <excl>
      <par id="p1">...</par>
      <par id="p2">...</par>
   </excl>
   <a href="p1"><img src="Button1.jpg"/></a>
   <a href="p2"><img src="Button2.jpg"/></a>
</par>
```

- Activating the first image hyperlink activates the media items of parallel container "p1". If the link on the second image is traversed, "p2" is started (thereby deactivating "p1" if it would still be active) from time 0.

Content Control Module

- 3GPP SMIL includes the content control functionality of the BasicContentControl, SkipContentControl and PrefetchControl modules of SMIL 2.0.

- BasicContentControl: contains content selection elements and predefined system test attributes (systemScreenSize, systemLanguage, systemCPU, etc.)

- SkipContentControl: containins attributes that support selective attribute evaluation (skip or evaluate an element)

- PrefetchControl: containing presentation optimization elements and attributes (control the fetching of content to improve the rendering performance)

Content Control Module and the Adaptation

- SMIL can be used to adapt the content for different devices
- The Content control modules allows to tailor the different parts of the presentation to different contexts
- Choose the best variant of the multimedia content or object on behalf of the user agent
- Based on:
  - Available variants (server)
  - Variants descriptions
  - Device capabilities and the user preferences
- Selection criteria may include the language, the media type, the char-set, etc.
- SMIL 2.0 allows using the *switch* element to specify inside the document a collection of alternative elements

SMIL Adaptability

Adaptation for the user preferences

Example using the SMIL Switch:

```
- <switch>
    <audio id="sound" src="../media/audioVersions/russian.mp3" begin="1" dur="30" systemLanguage="ru" />
    <audio id="sound" src="../media/audioVersions/japanese.mp3" begin="1" dur="30" systemLanguage="jp" />
    <audio id="sound" src="../media/audioVersions/english.mp3" begin="1" dur="30" />
  </switch>
- <switch>
    <text src="Privyet. Menya zovout Dominique." region="lyrics1" dur="30" systemLanguage="ru" />
    <text src="Konnichi-wa! Boku wa Dominiku desu." region="lyrics1" dur="30" systemLanguage="jp" />
    <text src="Hello, my name is Dominique." region="lyrics1" dur="30" />
  </switch>
- <switch>
    <text src="Ya zhivou v Kanade." region="lyrics2" dur="30" systemLanguage="ru" />
    <text src="Kanada ni sunde-imasu." region="lyrics2" dur="30" systemLanguage="jp" />
    <text src="I live in Canada." region="lyrics2" dur="30" />
  </switch>
```

SMIL Adaptability

SMIL content is adapted according to the user preferences

SystemLanguage = en

SystemLanguage = ru

Content Control Module and the Adaptation

SMIL + contextual attributes = Adapted Multimedia

- The content selection can be expressed using the SMIL system test attributes
- The player (or the adaptation mecanism) evaluates the test using the information of the context

Content Control Module and the Adaptation

A 'switch' example

SMIL:

```
<par>
  <audio src="welcome_to_inria.wav" . . . />
  <switch>
      <img src="inria_1024_1280.gif"
       systemScreenSize="1024X1280"/>
      <img src="inria_480_640.gif" systemScreenSize="480X640"/>
      <img src="inria_240_320.gif" systemScreenSize="240X320"/>
      <img src="inria_default.gif"/>
  </switch>
 </par>
```

Content Control Module and the Adaptation

An example of the SMIL SWITCH

Device: …, *Device screen: (240, 320), …*

Adapted SMIL:

```
<par>
  <audio src="welcome_to_inria.wav" . . . />
  <img src="inria_240_320.gif" systemScreenSize="240X320"/>
</par>
```

Content Control Module and the Adaptation

A *in-line* test attribute example

SMIL:

```
<par>
    <textstream src="presentation_speech_translation.rt"
     systemLanguage="fr" />
    <audio src="presentation_speech.mp3" . . . />
    <video src="presentation.mpg" . . . />
</par>
```

Content Control Module and the Adaptation

A *in-line* test attribute example

User:

language: English

SMIL:

```
<par>
        <audio src="presentation_speech.mp3" . . . />
        <video src="presentation.mpg" . . . />
 </par>
```

SMIL Adaptability

- The limitations of target devices and user preferences should be considered

➡ **Best use of the content on mobiles**

- Content authors should make their best effort to be device independent to a particular context or device

- Adaptation mechanisms should generate a content that meets the characteristics of the targeted device

SMIL Adaptability

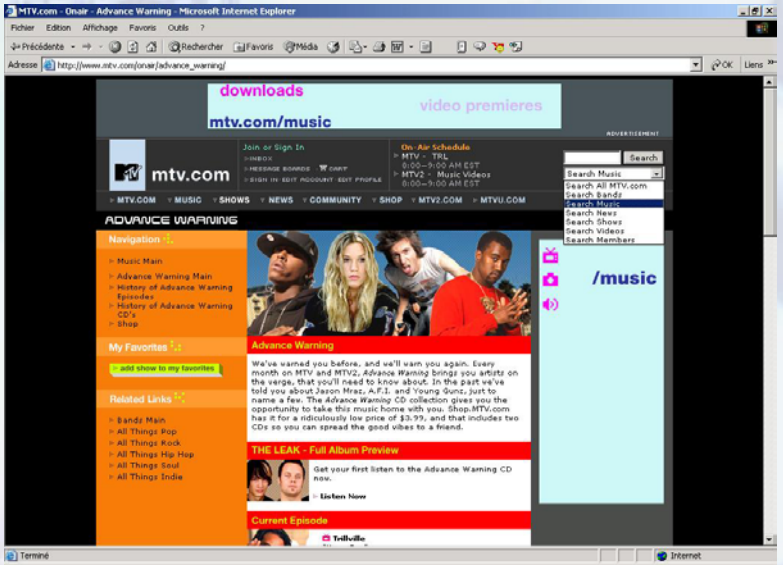
- Avoid *area* BasicLinking if the device doesn't have the appropriate user interface
- Consider display sizes of terminals for defining the layout of the SMIL presentation
- Don't be dependent to a given resolution of a rendering area
- Use the SMIL ContentControl for defining multiples layouts and media objects for different contexts
- Define a default layout in case when none of the other layouts can be used
- Use relative positions and size attributes
- Media objects (images, videos..) should be delivered in a size suitable to the targeted device
- Avoid some ways of fitting the media objects (SMIL fit attribute) if the target devices is unable to apply this functionality (example: *fit = scroll*)

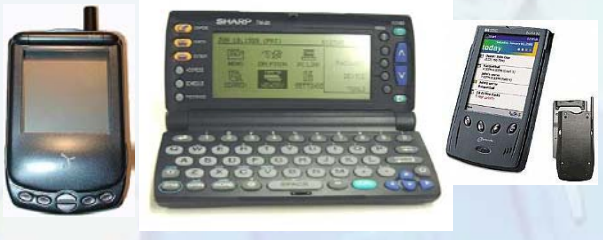
Device Independence

- A wide range of heterogeneous devices are used in multimedia systems
- Heterogeneity

- Displaying capabilities
- Access methods
- Media support (images, video, text, etc.)
- Languages (SMIL, MMS, SVG, XHTML, etc.)
- Protocols: HTTP, RTP, RTSP, etc.

Device Independence





Device Independence

Problem : Original content that exist in the server side can not be used directly by all the clients

Need : Adapted content must be delivered according to end user preferences and device capabilities

Device Independence WG

- **The mission of the Device Independence Activity is to avoid fragmentation of the Web into spaces that are accessible only from subsets of devices. In particular the Device Independence Working Group will**
  - collect requirements for Web access via various kinds of presentation device.
  - review related specifications within and outside of W3C
  - provide use cases and requirements to related activities within W3C
  - describe techniques which allow authors to improve management of device dependencies
  - in some specific areas not covered by other groups, to propose recommendations that will lead to enhanced device independence

Device Independence WG

- CC/PP
- Glossary for DI
- Authoring techniques for DI
- DI principles
- Content Selection
- Content Presentation Characteristics

Conclusions

- Several technologies are ready to be applied in order to use multimedia content on mobile devices
- Actually, many multimedia applications start to be used on mobile devices
- Problem of content adaptation and customization to meet devices limitations and user preferences
- Efforts: Continuous work of Device Independence Working Group (W3C)

Thank you

Tayeb LEMLOUMA

Tayeb.Lemlouma@inrialpes.fr