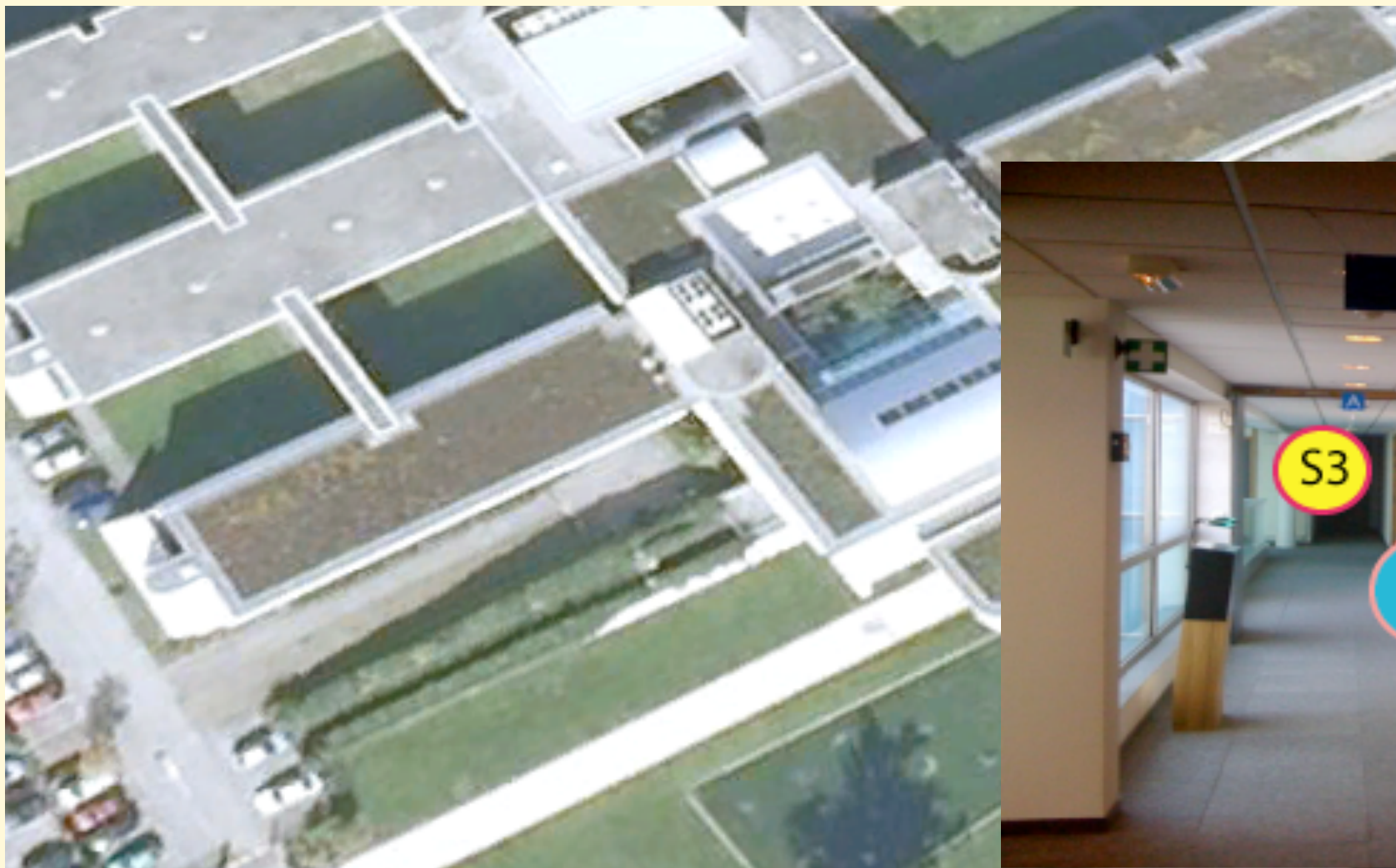


Project WAM

Web, XML and Multimedia

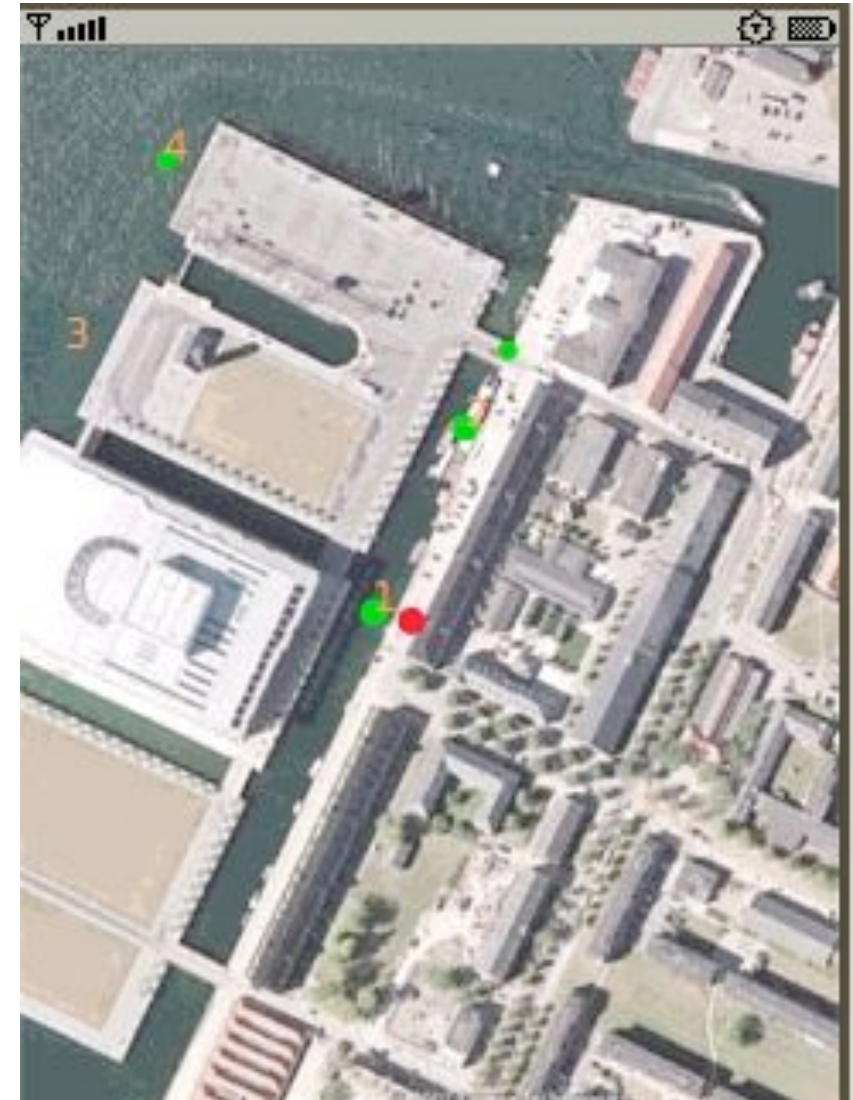
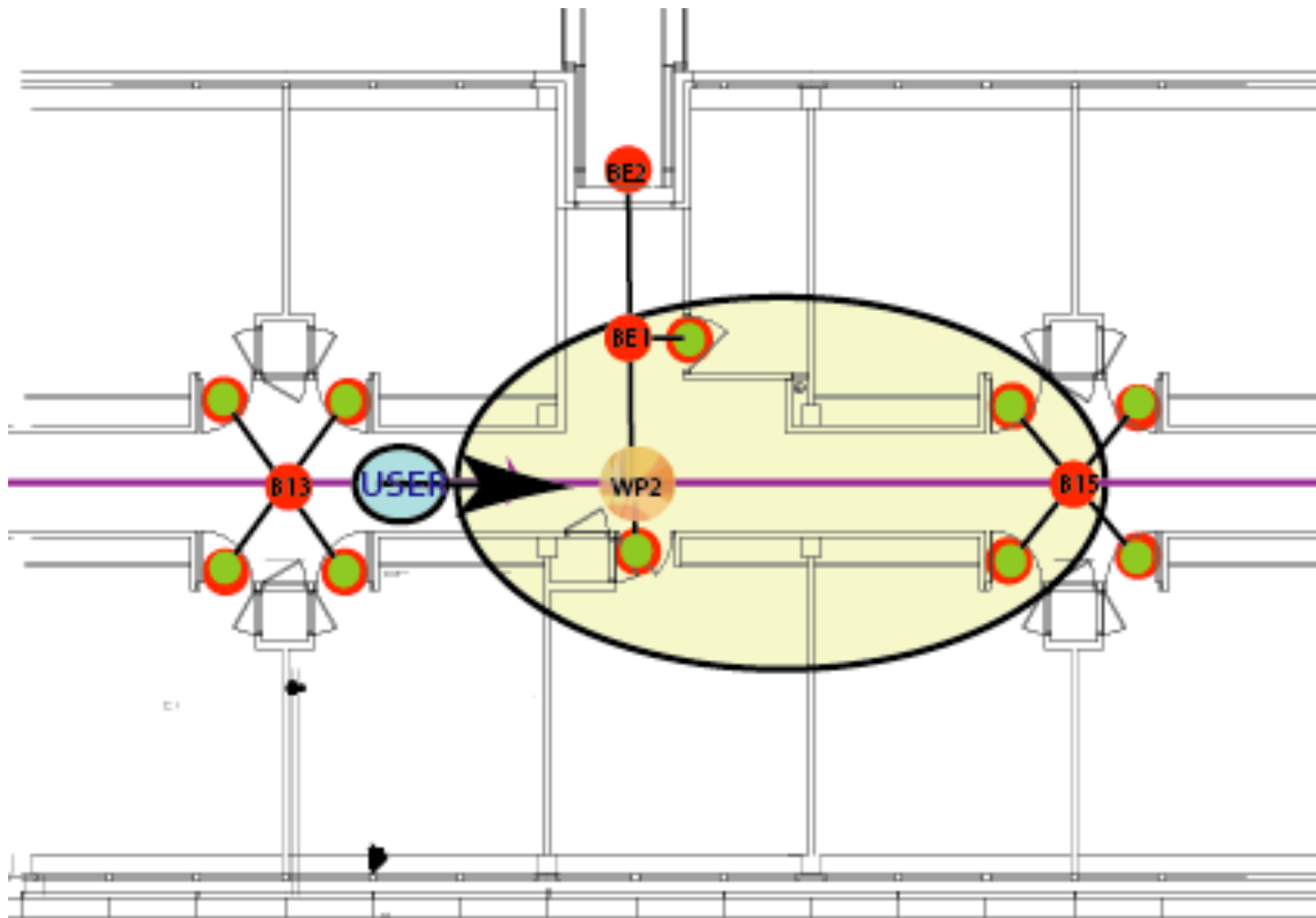
jacques.lemordant@inria.fr

MW4D and Interactive Audio on Mobiles



ARA Applications

- Audio walking tours
- Geolocalized games
- Navigation systems
- SoundScapes







AR Systems

- Two kinds of AR systems
 - POI (audio, video, images, 2D & 3D Objects) Browsers
 - Navigation Systems (indoor & outdoor)
- Convergence very soon or complementarity?
- Who will win?
- The main difference
 - Location data based for browsers
 - GIS data based for navigation systems

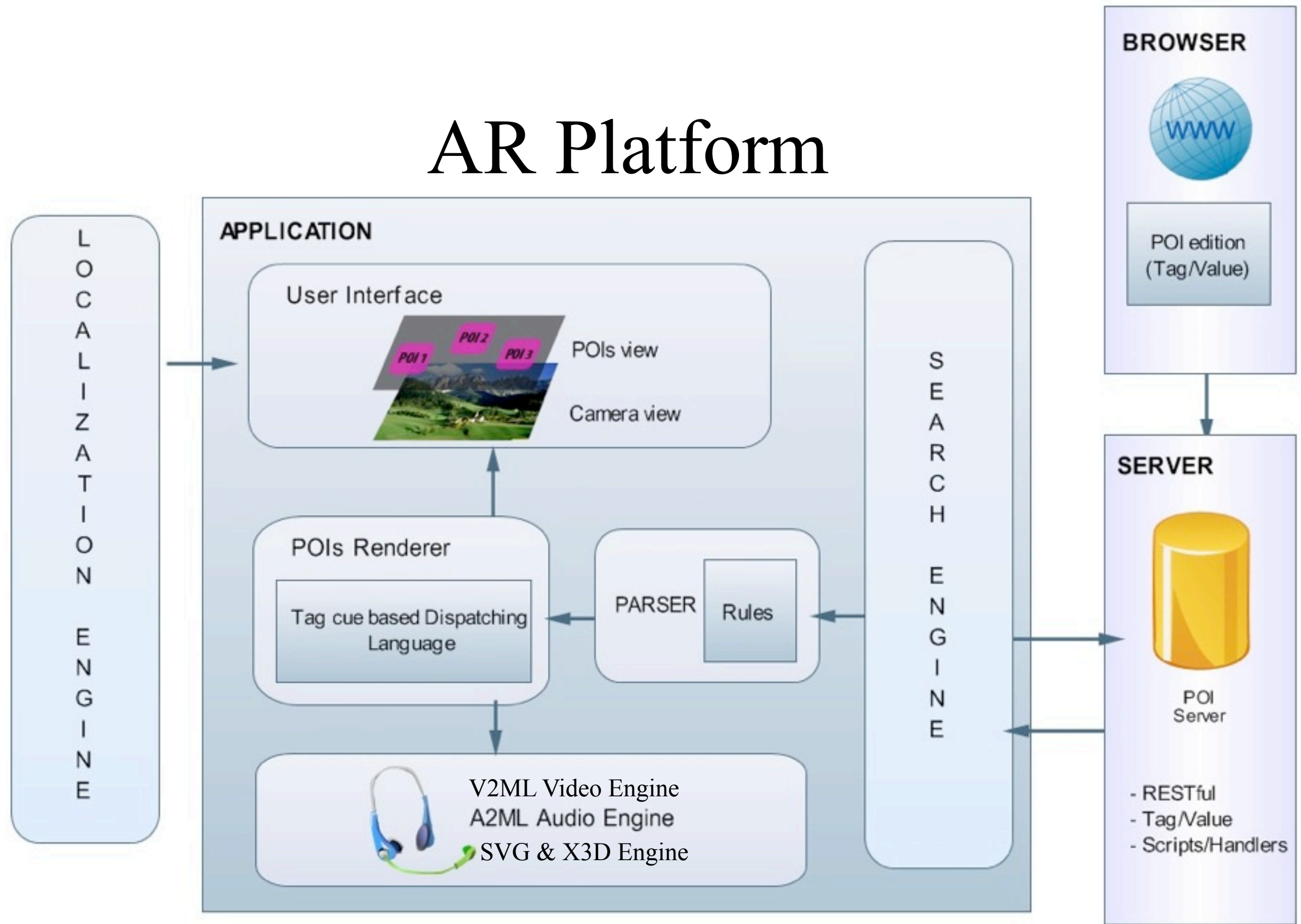
Augmented Reality

- Can be defined as a binding between Geolocation Data and Location Information

Or better (?) as

- Real-time Rendering of Geographical Information through a tag-based dispatching language with instantiation of graphics, images, audio and video objects
 - ---> XML formats for these objects
 - ---> Players/Schedulers for these object
- This will allow merging of geolocation data with local data

AR Platform



Tag-Based Dispatching Language

```
<!-- Standard Audio Style Sheet -->
<rules name="standard">
  <rule e="node" k="anemity" v="*">
    <layer name="anemity" active="no">
      <rule k="anemity" v="toilets">
        <cue name="details.toilets" />
      </rule>
      <rule k="anemity" v="bar">
        <cue name="details.bar"/>
      </rule>
    </layer>
  </rule>
  <rule e="way" k="floortype" v="carpet">
    <cue name=" environment.surface.carpet"/>
  </rule>
  ...
</rules>
```


Augmented Reality

An AR system may be defined as a system having the three following characteristics

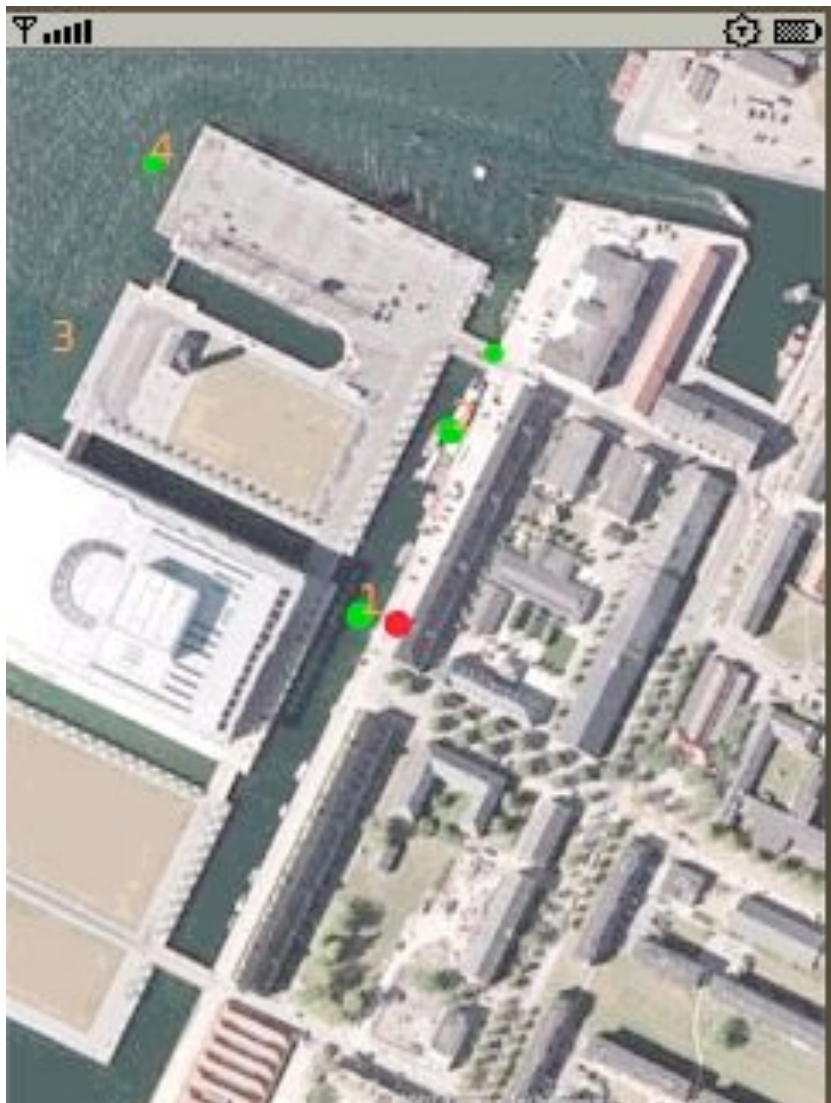
- combines real world objects and virtual objects
- is reactive or better interactive (objects with behavior)
- uses 3D positioning of virtual objects

This definition is clearly applicable to Augmented Reality Audio (ARA) systems where the virtual objects are sound objects.

WAM background (2007)

Copenhagen Underwater Soundscape

First step towards an XML format for IA and its link with an XML format (SVG) for the real world

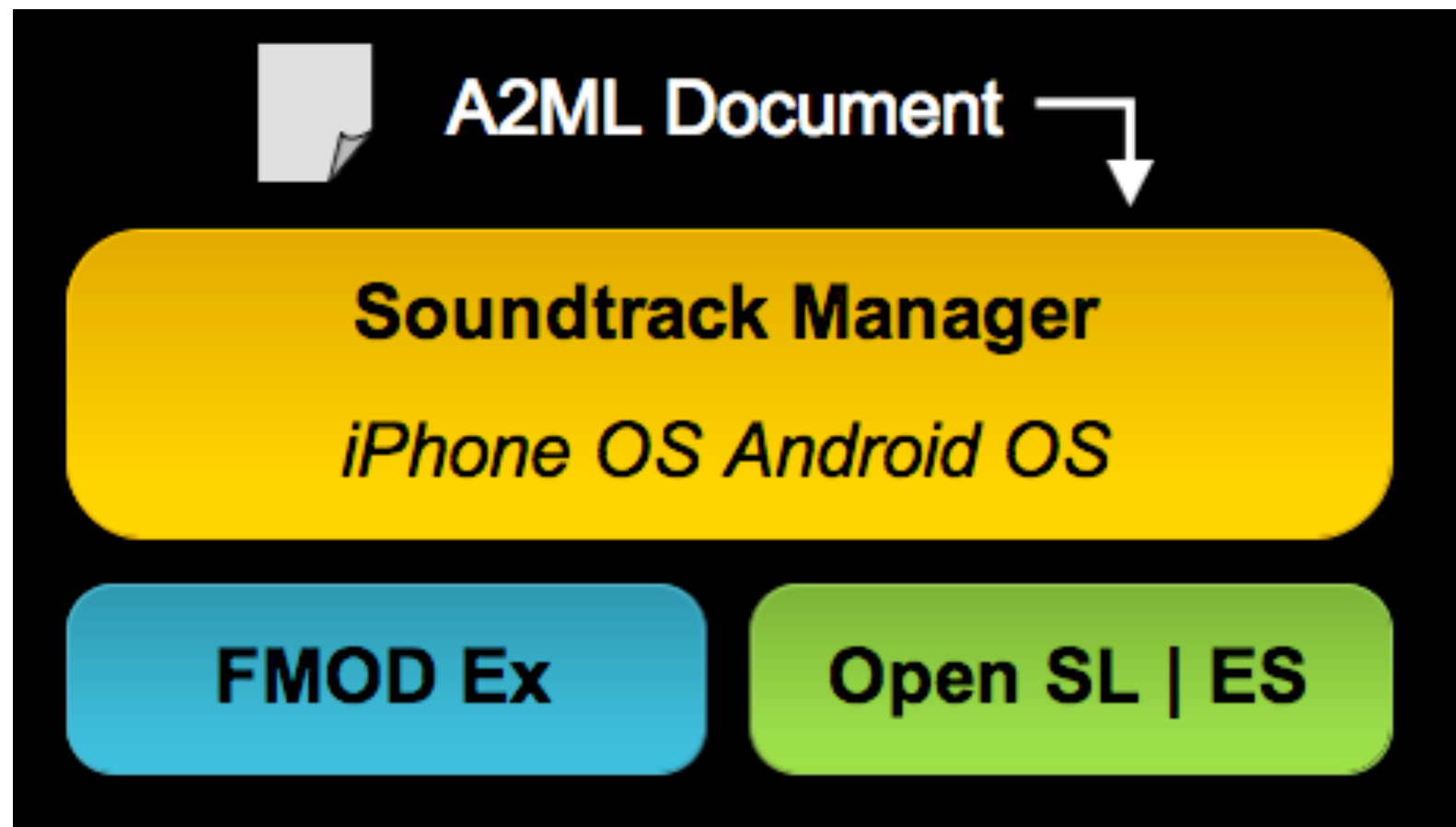


- 3D sound sources in green
- user with a mobile in red

The audio XML format was a serialisation of JSR 234 sound objects

From 2007 to 2010

A2ML: an XML Interactive Audio format with sound object models (cues): instantiation of models through an event API



iPhone Sound Manager (downloadable from WAM)
Symbian and android planned for 2011

ARA Systems

- Hardware for binaural rendering
- Formats for authoring
- Interactive versus reactive
- Applications

FULL ARA Rendering

The rendering of an ARA scene can be experimented through the use of:

- Bone conduction headsets
- Headphones with integrated microphones
- Earphones with acoustically transparent earpieces, with the audio being played by a mobile phone.

why? to combine real world sounds with virtual sounds (needed for safe navigation)



Formats for ARA Editing

An ARA scene has to be authored through the join use of two XML languages:

- one for the representation of the real world
- one for the representation of the 3D interactive virtual audio scene

The link between the two is done through a tag-based dispatching language (audio stylesheet): **don't put sound objects in a geographical language**

Real World : SVG, openStreetMap, KML, ...

Virtual Sound Objects: A2ML, HTML5 audio?, W3C standardization through the audio incubator group?

Reactive vs Interactive Audio

An XML Format for Interactive Audio and its associated DOM/event API have to be used to describe the 3D virtual audio scene.

The format A2ML describes models of sound objects whose instantiation is done through implicit events

A2ML is related to iXMF from IAsig and on the web side to SVG and SMIL

Interactive Audio

- Sound Objects
- Events
- Sound Manager

Sound Objects:
Where are they
coming from?

Closed Groove & Cut Bell

- The “closed groove” and the “cut bell” are the two “experiments in interruption” which were at the origins of musique concrète and certain discoveries in experimental theory.

With the arrival of the tape recorder, the tape loop replaced the closed groove

- Pierre Schaeffer

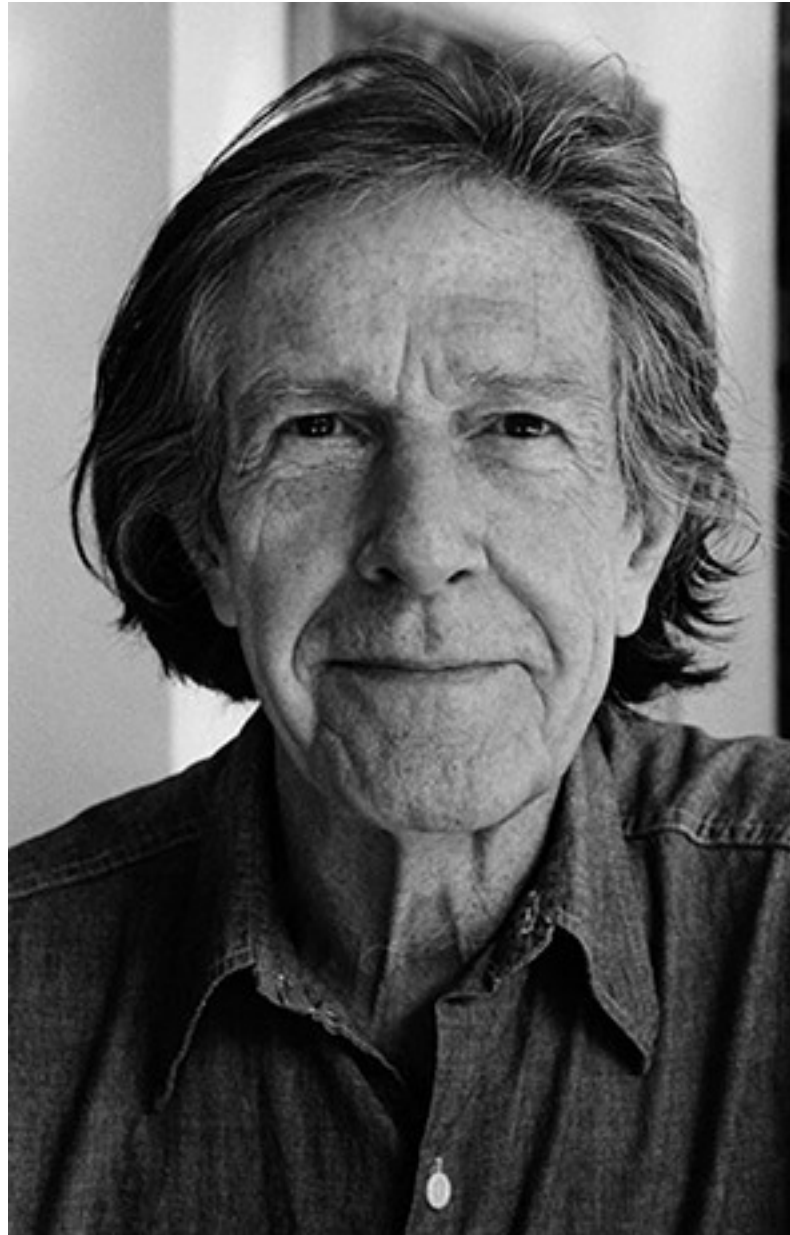


Concatenative Sound Synthesis (CSS)

- The groundwork for concatenative synthesis was laid in 1948 by the *Groupe de Recherche Musicale* (GRM) of Pierre Schaeffer, using for the first time recorded segments of sound to create their pieces of *Musique Concrète*.
 - Schaeffer defines the notion of *sound object*
- ➔ A composite sound object is a sequential container for chunks (50ms to 2s)

Indeterminacy

- John Cage



Ipods in shuffle mode

- Kenneth Kirschner



Models and Instances

- Phasing: a compositional technique popularized by Steve Reich (It's Gonna Rain)



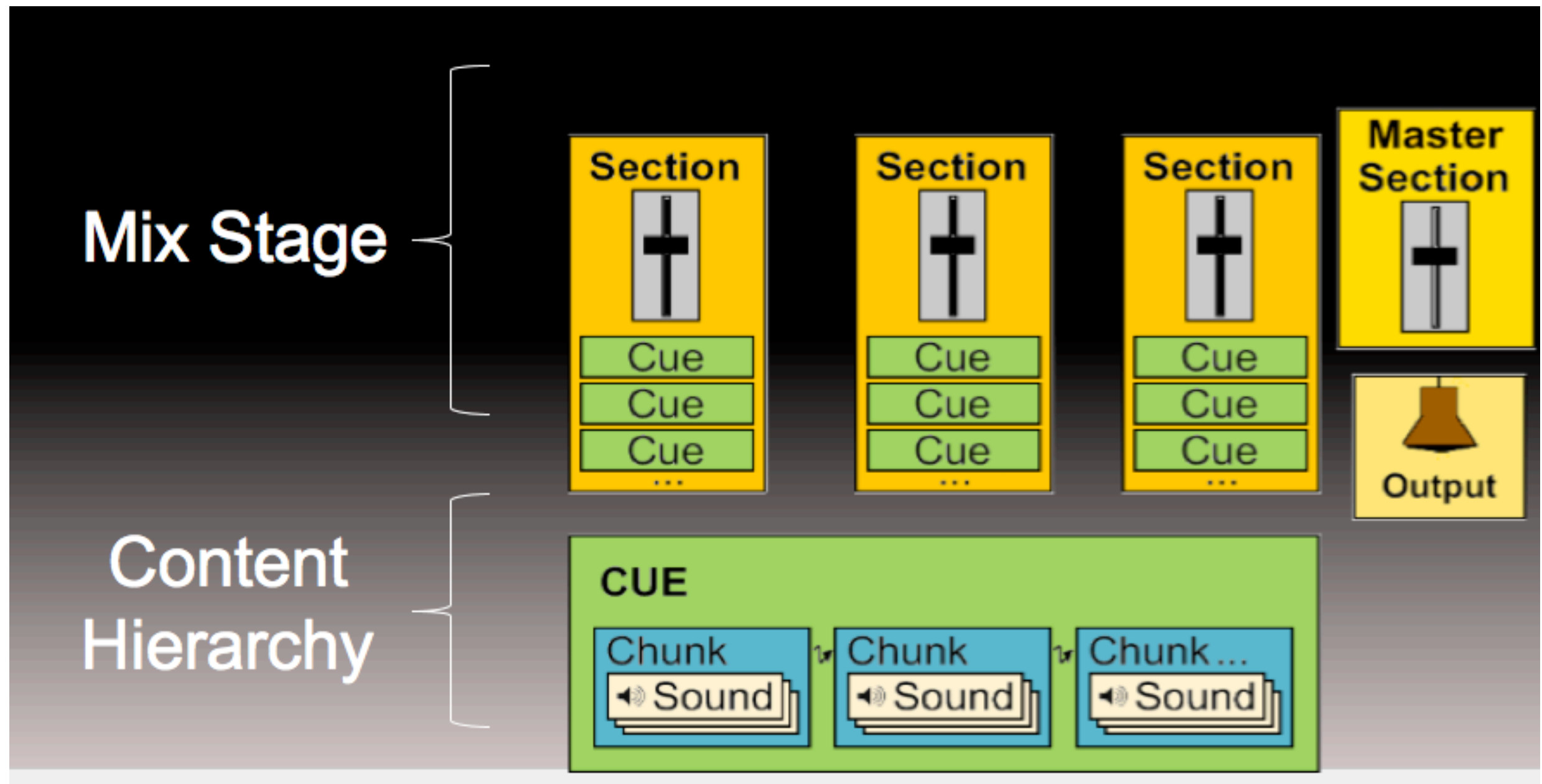
Repetitive Structures

- Philip Glass (Einstein on the beach)



XML Format for Interactive 3D Audio

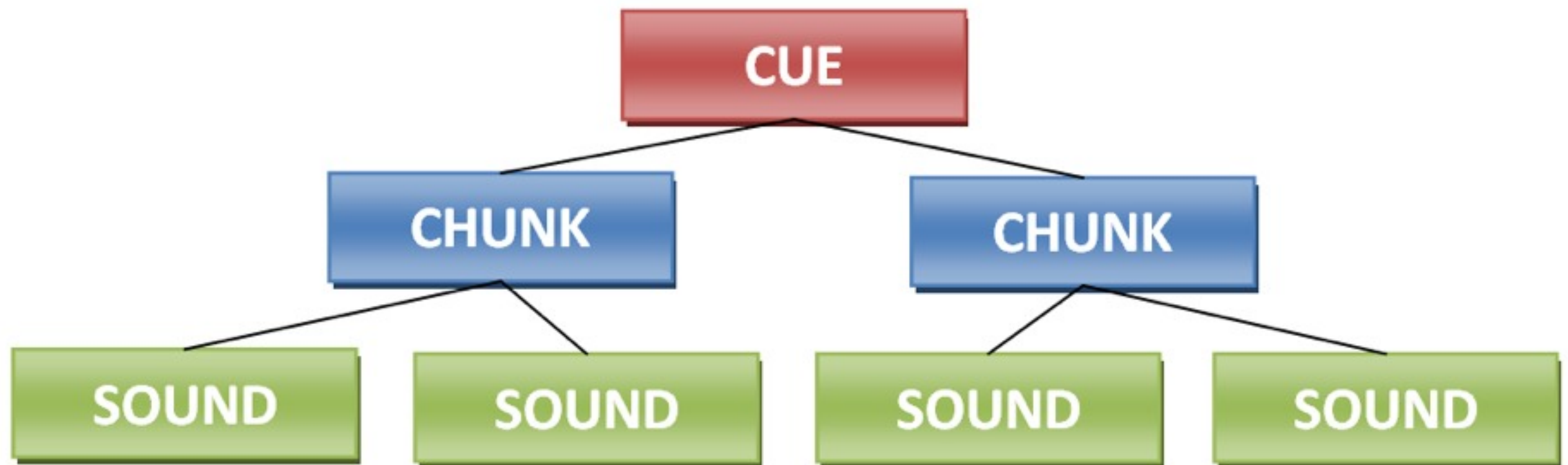
XML language inspired by iXMF from IASIG:
Cue based (a cue is a sound model with parameters)



Content Hierarchy

Why?

- Better organization (sound classification)
- Easy non-linear audio cues creation / randomization
- Better memory usage by the use of small audio chunks (common parts of audio phrases can be shared)
- Allows separate mixing of cues to deal with priority constraints easily
- Reusability



Content Hierarchy

```
<cue id="ambiance" loopCount="-1" begin="environment.ambiance">
  <chunk pick="fixed">
    <sound src="/environment/ambiance_office.wav" setActive="environment.set_ambiance.office"/>
    <sound src="/environment/ambiance_hall.wav" setActive="environment.set_ambiance.hall"/>
  </chunk>
</cue>
<cue id="distance_to_wp" loopCount="1" begin="audioguide.get_distance_to_wp">
  <chunk>
    <sound src="/guide/voice_about.wav"/>
  </chunk>
  <chunk pick="fixed">
    <sound src="/guide/voice_1.wav" setActive="audioguide.set_distance_to_wp.1_step"/>
    <sound src="/guide/voice_2.wav" setActive="audioguide.set_distance_to_wp.2_step"/>
    <sound src="/guide/voice_3.wav" setActive="audioguide.set_distance_to_wp.3_step"/>
    <sound src="/guide/voice_4.wav" setActive="audioguide.set_distance_to_wp.4_step"/>
    <sound src="/guide/voice_5.wav" setActive="audioguide.set_distance_to_wp.5_step"/>
    <sound src="/guide/voice_6.wav" setActive="audioguide.set_distance_to_wp.6_step"/>
    <sound src="/guide/voice_7.wav" setActive="audioguide.set_distance_to_wp.7_step"/>
    <sound src="/guide/voice_8.wav" setActive="audioguide.set_distance_to_wp.8_step"/>
  </chunk>
  <chunk>
    <sound src="/guide/voice_steps_to_next_wp.wav"/>
  </chunk>
</cue>
```

Sections

```
<sections>
  <!-- Mix group for the global audio guide. Use the reverb as a way to notify room size changes. -->
  <section id="audioguide" cues="next_wp door stairway elevator elevator_button ambiance
floor_surface way">
    <dspControl dspName="reverb">
      <parameter name="preset" value="default"/>
      <animate id="preset_change" attribute="preset" values="env.change_reverb_preset"/>
    </dspControl>
    <volumeControl level="70"/> </section>

  <!-- Activates 3D positioning for the object that need it. Position of the objects is controlled by the
guidance application. -->
  <section id="objects3D" cues="next_wp door stairway elevator floor_surface">
    <mix3D> <distanceAttenuationControl attenuation="2"></mix3D>
  </section>

  <!-- Submix group for the environment details. -->
  <section id="details" cues="atrium_door_number">
    <mix3D>
      <distanceAttenuationControl attenuation="5 </mix3D>
      <volumeControl level="100"/>
    </section>
  </sections>
```


Grain-based music

- Demo

```
<cue id="coconut_music" loopCount="-1">  
  <chunk pick="exclusiveRandom">  
    <sound src="/loop_1-1.wav" pickPriority="2"/>  
    <sound src="/loop_1-2.wav" pickPriority="2"/>  
    <sound src="/loop_1-3.wav" pickPriority="1"/>  
  </chunk>  
  <chunk pick="exclusiveRandom">  
    <sound src="/loop_2-1.wav" pickPriority="2"/>  
    <sound src="/loop_2-2.wav" pickPriority="2"/>  
    <sound src="/loop_2-3.wav" pickPriority="1"/>  
  </chunk>  
</cue>
```


Interactive jungle

- Demo

```
<cue id="river" begin="0s" ...>
<cue id="bird" end="shot.start" ...>
<cue id="monkeys" end="shot.start" ...>
<cue id="shot" begin="player.shot_event"
restartOnRetrigger="always">
  <chunk pick="random" fadeOutType="simplefade"
fadeOutDur="1s">
    <sound src="/shot_1.wav"/>
    <sound src="/shot_2.wav"/>
    <sound src="/shot_3.wav"/>
  </chunk>
</cue>
```


Sound variation

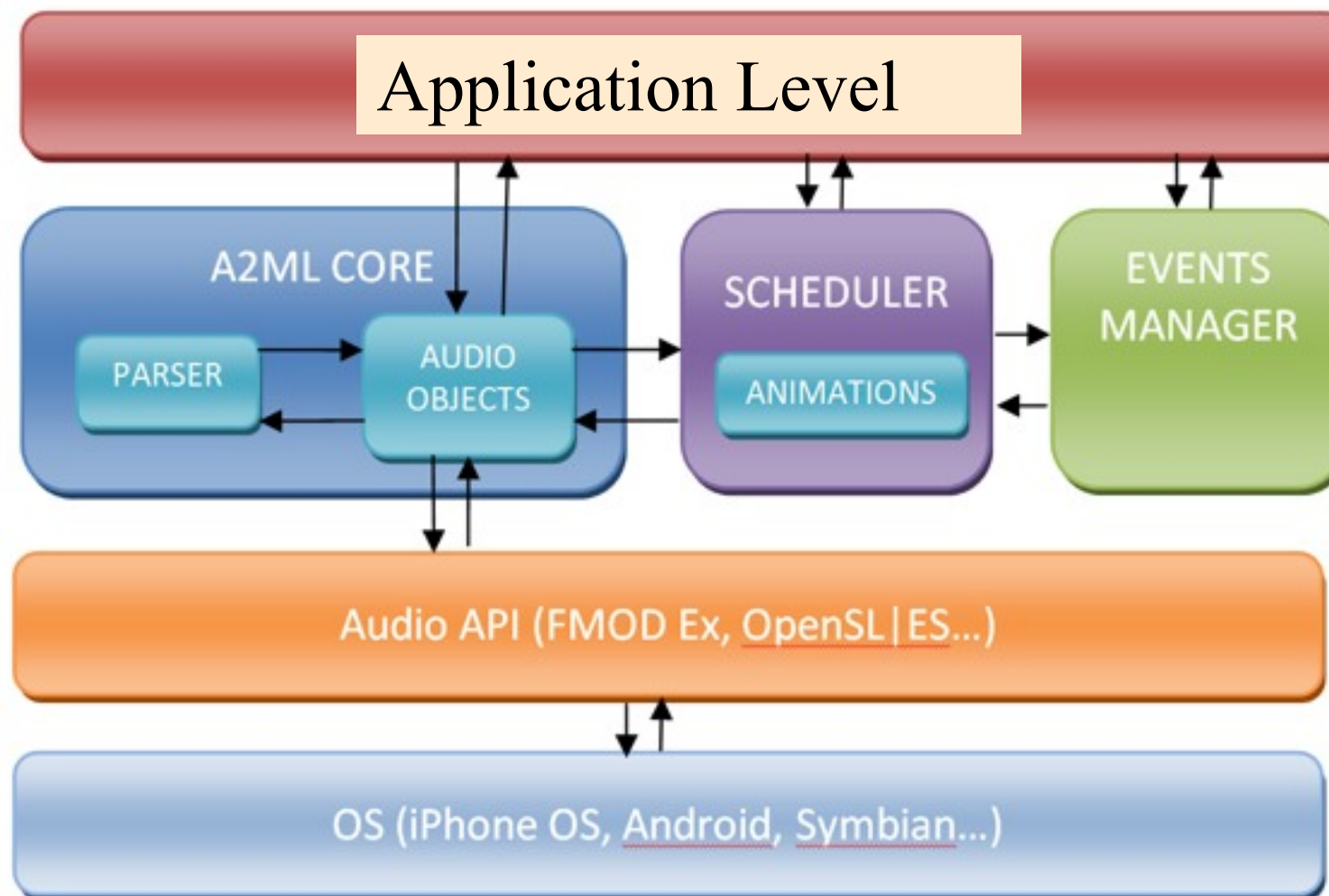
- Demo

```
<cue id="laser">  
  <chunk><sound src="/space_laser.wav"/></chunk>  
  <panControl>  
    <animate id="slide" attribute="pan" begin="laser.start"  
from="-100" to="100"/>  
  </panControl>  
</cue>  
  
<section id="dsp_fx" cues="laser">  
  <dspControl dspName="flanger">  
    <animate id="sweep" attribute="rate" begin="laser.start"  
from="0" to="rand(5000, 10000)"/>  
  </dspControl>  
</section>
```

A2ML API

Provides 2 freely mixable levels of access :

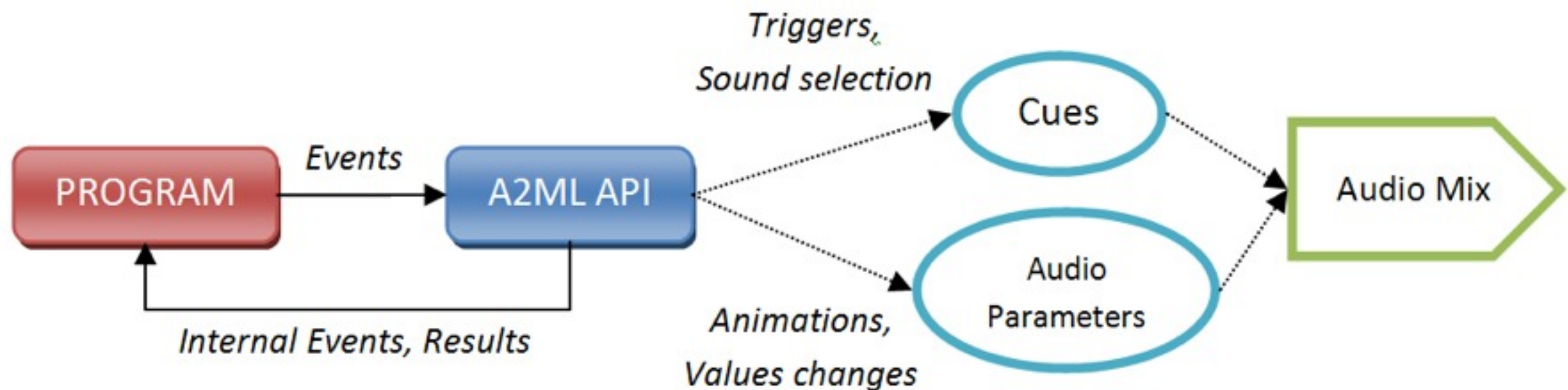
- High level : using A2ML events and scheduler timing
- Low level : provides core access to A2ML objects (Cues, Sections, audio objects instances...)



Events

Can be used for :

- Cues management (start / stop)
- Sound selection
- Parameter animation (3D positioning, mixing, reverb changes...)



iPhone A2ML Engine

- Based on **FMOD Ex** low-level API for mixing and 3D audio
- Native Objective-C / Cocoa implementation
- Low CPU / RAM overhead
- Easy integration with other XML format, for example SVG documents in a WebView

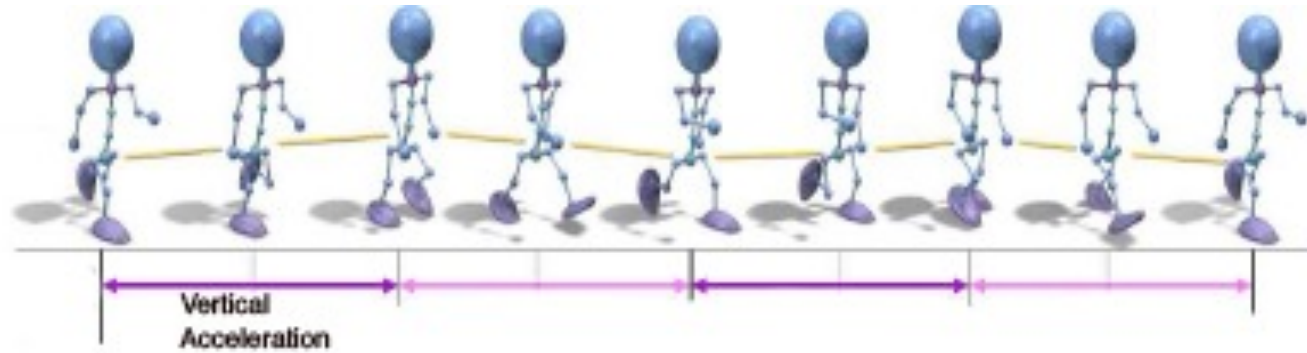


Localization

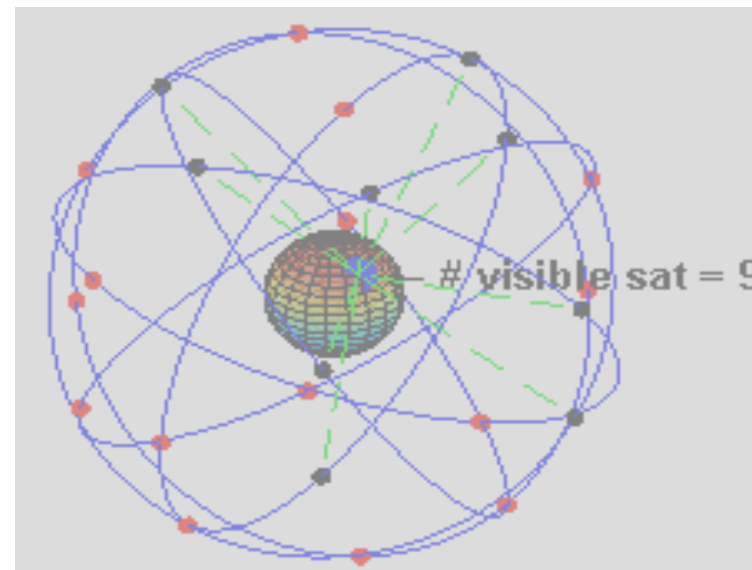
- Indoor/outdoor localization
- Geographical format for indoor/outdoor

Localization

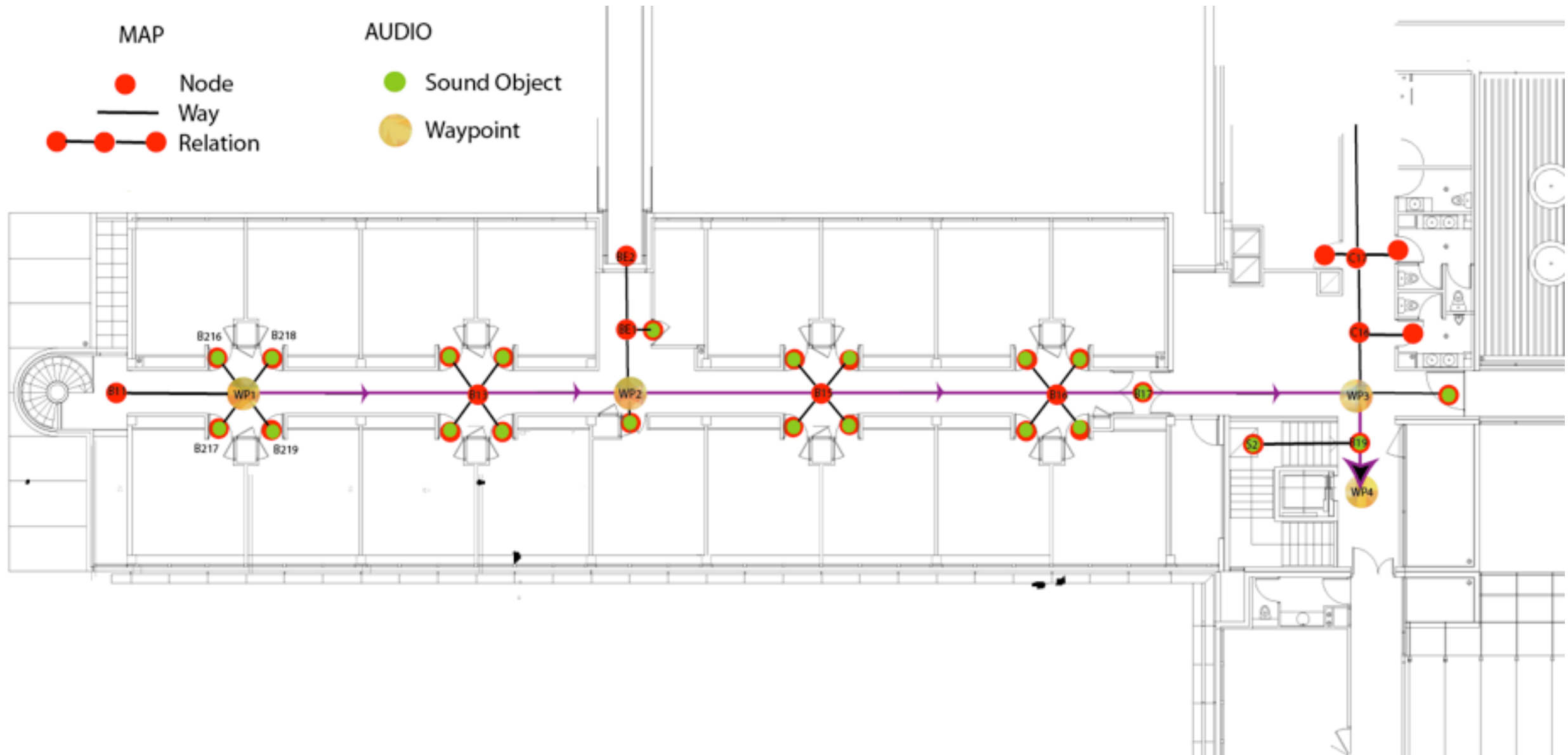
- Indoor: Accelerometers + Compass+ Giroscopes + zigbee network of sensors (time of flight)



- Outdoor: GPS



Map-Aided Positioning



OpenStreetMap (wiki format)

```
<osm version="0.6">
  <node id="S2" x="42.00" lon="3.00">
    <tag k="floor_access" v="stairs"/>
    <tag k="ele" v="14"/>
  </node>
  <node id="C17" x="45.2143789" y="7.00">
    <tag k="anemity" v="toilets"/>
    <tag k="addr:roomnumber" v="B207"/>
  </node>
  <relation id="R1">
    <member type='node' ref="B12"/>
    <member type='way' ref="B217" role="door"/>
    <tag k="type" v="junction"/>
  </member>
  ...
</relation>
  <way id="B1">
    <nd ref="B11"/>
    <nd ref="B12"/>
    ...
  </way>
  ...
</osm>
```

Audio Stylesheet and Mobile Mixing

- Link between the audio and geographical format
- See-through interface for mobile mixing

Cue Dispatching Language

```
<!-- Standard Audio Style Sheet -->
<rules name="standard">
  <rule e="node" k="anemity" v="*">
    <layer name="anemity" active="no">
      <rule k="anemity" v="toilets">
        <cue name="details.toilets" />
      </rule>
      <rule k="anemity" v="bar">
        <cue name="details.bar"/>
      </rule>
    </layer>
  </rule>
  <rule e="way" k="floortype" v="carpet">
    <cue name=" environment.surface.carpet"/>
  </rule>
  ...
</rules>
```


See-through Mobile Mixing



ARA Editing

- Alternate between two modes:



- Static mode: using an XML editor, create, modify or rearrange cues quickly and intuitively.
- Mobile mode: test the soundscape through a visual AR interface and adjust parameters

Links

- These slides
 - <http://wam.inrialpes.fr/>
- Software and specifications
 - <http://gforge.inria.fr/projects/iaudio>
 - scheduler for iPhone, A2ML specifications