

Estándar SMIL y aplicaciones

Raquel Sánchez Paniagua
Departamento de Informática. Universidad de Oviedo

Resumen

La mejora en las redes de acceso de los usuarios y el incremento de ancho de banda disponible, ha provocado que la cantidad de servicios que los usuarios tienen a su disposición se dispare. Ahora los usuarios ya no cuentan solamente con servicios basados en texto e imágenes, si no que disponen de capacidad suficiente para acceder a servicios de audio, vídeo, compartir ficheros,... Toda esta riqueza de medios, no tiene porque ofrecerse de forma independiente, si no que puede sincronizarse, ofreciendo al usuario un servicio con múltiples medios audio visuales coordinados. En este paper, se analizan las principales características del lenguaje SMIL y se presenta un ejemplo de aplicación en el que, mediante el uso de SMIL se añade valor a un servicio de radio a través de Internet.

Palabras Clave

SMIL
PRESENTACIONES MULTIMEDIA
SINCRONIZACIÓN

1 Introducción

La indudable mejora en las redes de comunicación está contribuyendo a la aparición de una serie de servicios que hace unos años era totalmente impensable. Prensa digital, servicios multimedia, comercio electrónico, transferencia de ficheros...son servicios que están al alcance de la mayor parte de los usuarios.

Estos nuevos servicios hacen más atractivo el uso de Internet llegando algunos de ellos a tener una enorme difusión. Este es el caso de los servicios multimedia, que en los últimos años están teniendo una acogida sorprendente por parte de los usuarios. La aparición de todo tipo de información, texto, imagen, animación, vídeo, sonido en los contenidos disponibles en la web hicieron necesario lenguajes que permitieran integrar y sincronizar los diferentes contenidos.

SMIL (Synchronized Multimedia Integration Lenguaje) es un estándar del W3C basado en XML. Permite sincronizar e integrar audio, vídeo, imágenes, texto... a presentaciones multimedia. Además el desarrollador puede realizar múltiples acciones en las presentaciones multimedia tales como:

- Describir el comportamiento temporal de la presentación.
- Describir la distribución de los elementos de la presentación en la pantalla.
- Crear hiperenlaces a otros objetos multimedia, otras presentaciones SMIL o a otras páginas.
- Utilizar botones para controlar el flujo de la presentación.
- Utilizar funciones que permiten definir diferentes secuencias y duración de los elementos de la presentación.

2 SMIL

2.1 Breve historia de SMIL en el W3C

En junio de 1998 aparece la primera recomendación del W3C de SMIL. Este primer estándar es denominado SMIL1.0 [SMIL 1.0] permitiendo al autor las siguientes operaciones.

- Describir el comportamiento temporal de la presentación.
- Describir la disposición de la presentación en una pantalla.
- Asociar los hyperlinks a los objetos medios.

En agosto del 2001 aparece la recomendación de SMIL 2.0 [SMIL 2.0] cuya principal diferencia respecto a SMIL 1.0 es la división en módulos. SMIL 2.0 introduce nuevas y potentes facilidades de navegación, animación y soporte para realizar broadcast.

SMIL 2.1 se convierte en recomendación del W3C en diciembre del 2005. Esta nueva recomendación [SMIL 2.1] incrementa las herramientas disponibles por el programador incluyendo mejoras de las funciones en el formato visual. Otro factor importante es que puede mejorar el rendimiento, característica de gran importancia en especial en dispositivos móviles.

2.2 Perfiles

Un “perfil” de SMIL 2.1 es un conjunto de módulos. La recomendación SMIL 2.1 define tres nuevos perfiles

- Perfil de la lengua de SMIL 2.1.
- Perfil del móvil de SMIL 2.1.
- Perfil del móvil de SMIL 2.1 Extended.

Los perfiles Móvil y Extendido han sido creados para adaptarse a las diferentes capacidades de los dispositivos móviles.

2.3 Estructura

SMIL define una serie de elementos predefinidos, los más importantes son:

1. **Smil:** contiene todos los elementos de un documento smil, es decir, es el elemento raíz del documento. Dentro del ámbito de este elemento podemos incluir un elemento head y debemos incluir un elemento body.
2. **head:** Contiene información sobre el interfaz de la presentación y otra meta-información.
3. **body:** Contiene los elementos propios de la presentación, tales como los contenidos, los enlaces o la temporización.
4. **layout:** Contiene todo lo relativo a la distribución, incluidos los atributos de una ventana. Esto se realiza mediante el uso de la marca layout incluida en la sección head.
5. **region:** Para introducir un objeto multimedia en una presentación, se debe especificar la región donde se debe mostrar. Esto se realiza mediante el uso del elemento incluido en la sección head.

2.4 Contenidos de la recomendación SMIL 2.1

A continuación se describen los principales contenidos de esta recomendación [Smil1].

- Módulo de integración.

- Módulo de disposición.
- Módulo de sincronización.
- Modulo de transiciones.
- Módulos de animación
- Módulos de control de contenidos

2.4.1 Módulo de integración

En una presentación SMIL se pueden incluir diferentes elementos indicándolo con su URL correspondiente. Deben ser especificados por las siguientes etiquetas.

- **ref**: referencia para cualquier tipo de contenidos.
- **animation**: para animaciones.
- **audio**: fichero de audio.
- **Img**: Imagen como JPEG, png..
- **Text**: Referecia a un texto.
- **Textstream**: Texto con información sincronizada.
- **Video**: Contenido de video

Brush: Dado que no es un elemento propiamente dicho, se considera como que permite definir el color de relleno para una región.

Para integrar los contenidos se emplea el atributo "scr" que almacena la URI del fichero. Por ejemplo, si se desea integrar la pagina html "texto.htm" a la presentación multimedia se expresará de la siguiente manera:

```
<text scr="texto.html"... />
```

A continuación se puede observar un sencillo ejemplo (ejemplo1) del uso de SMIL. En él se definen dos regiones una para el audio y otra para la imagen. Además también se le añade un texto explicativo a la imagen. El código fuente sería:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE smil PUBLIC "-//W3C//DTD SMIL 2.0//EN"
"http://www.w3.org/2001/SMIL20/SMIL20.dtd">
<smil xmlns="http://www.w3.org/2001/SMIL20/Language">
  <head>
    <layout>
      <root-layout title=" RADIO IMAGEN " id="layout1" width="300" height="400"/>
      <region id="R1" left="0%" right="0%" width="300" height="300" fit="fill"/>
      <region id="R2" left="0%" right="0%" width="300" height="100" top="300"
fit="fill"/>
      <region id="R3" left="0%" right="0%" fit="fill"/>
    </layout>
  </head>
  <body>
    <par>
      
      <textstream src="./texto1.rt" region="R2" />
      <audio src="./musica.mp3" region="R3"/>
    </par>
  </body>
</smil>
```

Ejemplo1 Ejemplo de smil con diferentes regiones

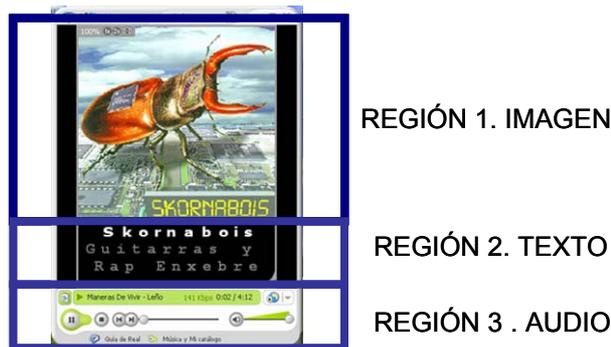


Ilustración 2-1. Resultado de ejemplo1

El resultado de este SMIL es una pantalla como la que se muestra en la ilustración2-1 a continuación con el audio de fondo.

2.4.2 Módulo de disposición

En este módulo contiene los elementos y atributos que permiten posicionar localizar los diferentes contenidos en las presentaciones multimedia. Para ello se definen los siguientes elementos.

- **Layout:** Define una colección de regiones para la presentación.
- **Root-layout:** Define la pantalla principal de la presentación.
- **Region:** Es necesario definir una región para cada objeto multimedia.

Una presentación puede tener cero, uno o múltiples elementos layout.

A las regiones se le pueden añadir atributos que ayuden a definir más concretamente cada región.

- **id:** Identificador para posteriormente referirse a la región.
- **left,right,top,botton:** Son atributos que definen la ubicación de la región en la pantalla.
- **Width, height:** atributos que define la anchura y altura de la región.
- **Z-index:** Sirve para definir qué región se colocará por encima cuando más de una región coinciden en el espacio.
- El atributo **fit** permite ajustar los objetos a sus regiones contenedoras.
- **backgroundColor:** Esta cualidad especifica el color del fondo de la región.

2.4.3 Módulo de sincronización.

La sincronización define los elementos y cualidades para coordinar y sincronizar las presentaciones multimedia. Se define tres elementos para realizarla:

- **Seq:** cada hijo comienza cuando finaliza el anterior.
- **Par:** Todos los hijos son iniciados a la vez.
- **Excl:** No impone orden. No se puede reproducir más de un hijo a la vez.

Los atributos que se admiten son los siguientes:

- **Begin:** Comienzo por defecto.
- **End:** fin por defecto.
- **Dur:** Duración explícita que sustituye a la implícita.
- **repeatDur:** Iterar durante un periodo concreto.
- **repeatCout:** Iterar un nº de veces concreto.
- **Min:** Valor de tiempo mínimo de presentación
- **Max:** Valor de tiempo máximo de presentación.
- **Fill:**

1. **freeze**: (por defecto) mantiene visible un video hasta que expira la vida de su padre. No tiene efecto con audio.
2. **remove** : Hace desaparecer un medio visual (comportamiento por defecto si se define duración).
3. **restart**: Reinicio de una reproducción en curso (“always” “whenNotActive” o “never”).

2.4.4 Módulo de transiciones

El elemento *transition* permite realizar transiciones en la presentación mediante los atributos oportunos. En el ejemplo2 (véase apéndice) se puede observar un ejemplo de uso de las transiciones combinado con elementos de sincronización. En las Ilustraciones 2-2 y 2-3 se muestran dos instantes diferentes en la transición de las imágenes.

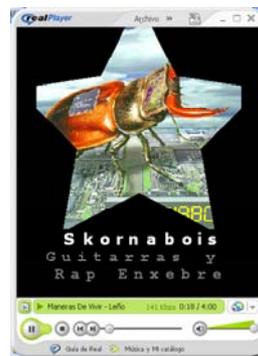
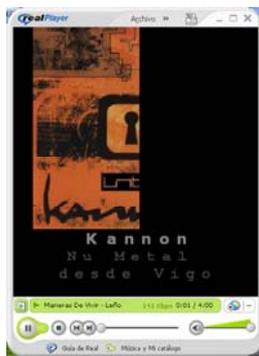


Ilustración 2-2. Transición de la primera imagen Ilustración 2-2transición de la segunda imagen

2.4.5 Módulo de control de contenidos

El principal elemento de este módulo es switch:

- **Switch**: Permite elegir entre un grupo de elementos.

```

...
<par>
  <video src="anchor.mpg" ... />
  <switch>
    <audio src="dutchHQ.aiff" systemBitrate="56000" ... />
    <audio src="dutchMQ.aiff" systemBitrate="28800" ... />
    <audio src="dutchLQ.aiff" ... />
  </switch>
</par>

```

Si el bitrate del sistema es 56000 o más alto, se selecciona el objeto *dutchHQ.aiff*. Si el bitrate del sistema es por lo menos 28800 pero menos de 56000, se selecciona el objeto *dutchMQ.aiff*. Si no se selecciona ningunos otros objetos, se selecciona el alternativa *dutchLQ.aiff*.

3 Ejemplo de aplicación de SMIL

Las aplicaciones que se abren en el campo de los servicios multimedia a la sincronización de diversos contenidos, es enorme. En esta sección presentaremos un sistema de radio con información adicional sobre los contenidos que se están reproduciendo en cada momento. La información adicional puede contemplar desde

imágenes de los grupos, información en forma de texto sobre los componentes, historia del grupo, discos, enlaces a tiendas virtuales con las canciones reproducidas, etc,... De esta forma, añadiremos valor al servicio, haciéndolo más atractivo para los usuarios y abriendo nuevas opciones relacionadas con la financiación del servicio; publicidad, tiendas virtuales promocionadas,...

3.1 Sistema de radio con metadatos

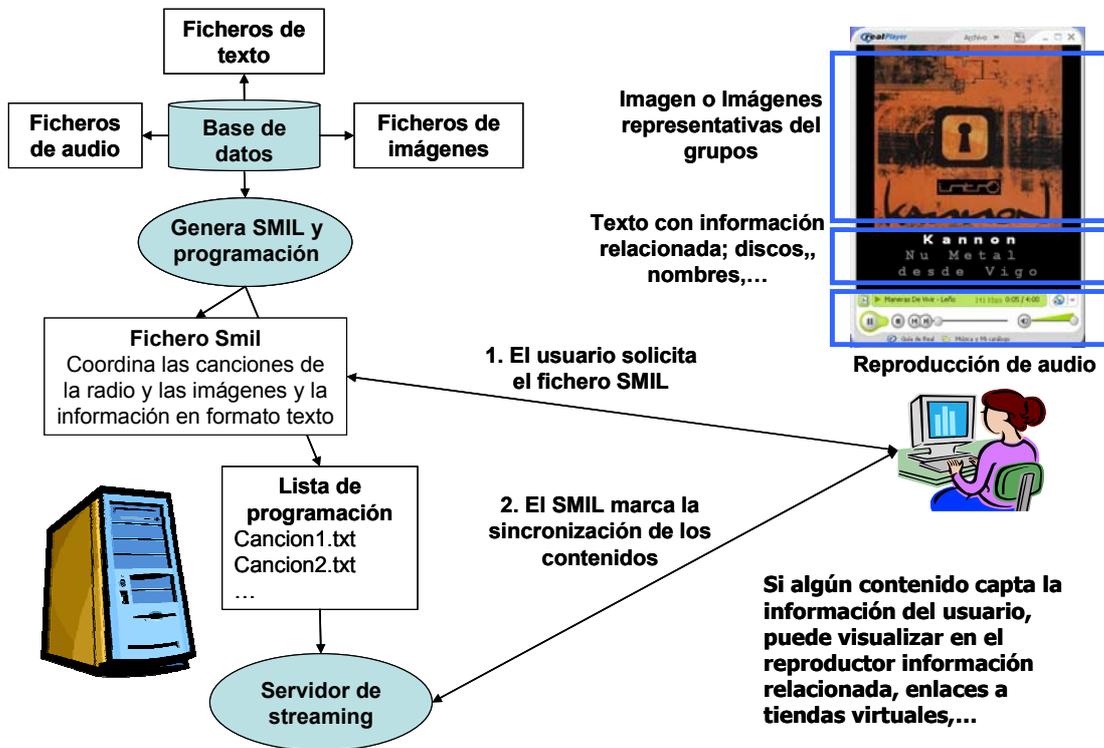


Ilustración 3-1 Arquitectura del sistema.

En la Figura 3-1, podemos ver la estructura del sistema. Este sistema está formado por los siguientes módulos:

- Servidor de streaming. Será el encargado de entregar los distintos flujos multimedia a los usuarios. Ya sea el propio servidor de streaming [WindowsMedia] o una aplicación adicional incluida en la distribución [RealNetworks], este elemento leerá una lista de reproducción que marque el orden de los elementos a presentar a los usuarios.
- Generador de lista de reproducción. Este módulo generará el fichero que marcará los contenidos del flujo del radio en cada momento. Definirá el orden de reproducción de los contenidos.
- Generador de SMIL. Este módulo, basándose en la programación diaria, generará un fichero SMIL en el que se coordinen las imágenes y la información de texto de cada grupo con el instante en el que se reproduce la canción.
- Base de datos. Esta base de datos se encargará de almacenar las relaciones entre grupos, canciones, imágenes y contenidos de texto. Además, almacenará la duración de las canciones y su situación.

Como se puede ver en la figura3-1, el resultado final es un conjunto de elementos multimedia sincronizados, que añaden riqueza a un sencillo servicio de radio basado en contenidos prealmacenados.

3.1.1 Base de datos

Esta base de datos almacenará información relacionada con los contenidos almacenados en el sistema y las relaciones entre ellos. Esta base de datos almacenará información sobre:

- Grupos dados de alta en el sistema:
 - ID grupo.
 - Nombre del grupo.
 - Miembros.
 - Estilo musical.
 - Discografía.
- Canciones dadas de alta en el sistema:
 - ID canción
 - Ruta y fichero de la canción.
 - Duración de la canción.
 - ID grupo al que pertenece la canción.
- Imágenes dadas de alta en el sistema:
 - ID de la imagen.
 - Ruta y fichero de la imagen.
 - ID del grupo al que pertenece la canción.

Toda esta información puede incorporarse durante la fase de alta de los contenidos con herramientas como las presentadas en [Pañeda2004].

3.1.2 Generación de lista de reproducción

Basándose en la información almacenada en la base de datos sobre los contenidos musicales almacenados en el servidor, la herramienta generará una lista de reproducción con los contenidos de la emisión diaria. Por ejemplo, la información necesaria para construir esta lista en los sistemas Real Networks, se limita a la ruta y el nombre de cada uno de los archivos en un fichero de texto.

Basándose en la información almacenada en estas tablas sobre los contenidos del sistema y aplicando un proceso de selección aleatorio, se puede construir una lista de reproducción diferente para cada día. Otros algoritmos más complejos para la selección de contenidos basados en información relacionada, por ejemplo, con promociones de discográficas podría aplicarse en este proceso de selección de contenidos.

3.1.3 Generación de SMIL

El generador de SMIL leerá el contenido de la programación diaria, obteniendo el ID de cada una de las canciones incluidas en el sistema. Basándose en este ID, el generador de SMIL, consultará la duración de la canción. Esta duración es la base de la sincronización de los contenidos.

Para cada una de las canciones, se consultará el grupo relacionado y se obtendrán las imágenes relacionadas. Con estas imágenes relacionadas, se preparará una presentación cuya duración será la misma que la de la canción presentada. Por ejemplo, se podrá cambiar de imagen cada 10 o 20 segundos mientras la canción no finalice.

Además, para cada canción se consultará el grupo relacionado y se obtendrán los metadatos relacionados con el ID del grupo. Estos metadatos se utilizarán para crear los contenidos de texto relacionados con el grupo. La duración de estos contenidos de texto, nuevamente, serán iguales a la duración de los contenidos del sistema de radio. A la hora de dotar de dinamismo al texto presentado por el grupo, se puede recurrir a la definición de ficheros de RealText [RealText].

A modo de ejemplo, un SMIL que coordinase los diferentes contenidos del sistema de radio, podría tener la estructura presentada a continuación. Como podemos ver, se trata de una versión simplificada que solo presenta una imagen por cada canción y un fichero Real Text por canción. Sin demasiadas dificultades, este ejemplo podría extenderse a la inclusión de múltiples imágenes por cada canción.

```
<smil xmlns="http://www.w3.org/2001/SMIL20/Language">
...
<body>
  <seq repeatCount="indefinite">
    <par>
      <audio src="rtsp://Servidor/Cancion1.rm" " region="R3"/>
      
      <textstream src="/Grupo1.rt" region="R2" dur="DuracionCancion1"/>
    </par>
    <par>
      <audio src="rtsp://Servidor/Cancion2.rm" " region="R3"/>
      
      <textstream src="/Grupo2.rt" region="R2" dur="DuracionCancion2"/>
    </par>
    ...
  </seq>
</body>
</smil>
```

Ejemplo3. Ejemplo de SMIL de coordinación de contenidos.

```
<window bgcolor="black">
  <center><b>
    <font size="10" color="white" face="verdana">Nombre grupo</font>
  </b></center>
  <center><b>
    <font size="8" color="gray" face="courier">Descripción grupo</font>
  </b></center>
</window>
```

Ejemplo4. Ejemplo de fichero RealText.

3.1.4 Funcionamiento del sistema

En el momento en el que el usuario accede al sistema y solicita los contenidos del portal de radio, se generará una programación personalizada y se añadirán las referencias necesarias en el fichero SMIL. Este fichero sincronizará los diferentes contenidos multimedia. El servidor multimedia devolverá este fichero al cliente de streaming y el cliente de streaming lo interpretará, es decir, establecerá las conexiones necesarias con el servidor para descargar los contenidos hasta el cliente. Además, la descarga y la

presentación de contenidos se realizarán en el orden y forma definidos en el fichero SMIL.

4 Conclusiones

La sincronización de contenidos multimedia abre un campo de aplicaciones enorme a los servicios de audio/vídeo en Internet. Ahora no solo estamos dotando a nuestros servicios de características multimedia (texto, audio, video, imágenes,...) si no que coordinamos todos estos contenidos de forma dinámica. El dinamismo llegó no solo a fijar los patrones de evolución del SMIL si no a poder particularizar estos en función de la interactividad mantenida por el usuario.

Todo ello convierte a SMIL en uno de los estándares de presente y futuro de la web, siendo soportado actualmente por multitud de reproductores multimedia e incluido en numerosos servicios de streaming comerciales.

5 Trabajos futuros

Como trabajos futuros queda la extensión de funcionamiento del sistema presentado a un sistema de radio en directo, formado por una emisión continua de contenidos y común para todos los usuarios. En este caso surge el problema de la sincronización, ya que, cada usuario deberá recibir un SMIL particularizado para los contenidos que se están reproduciendo en ese momento.

Así mismo, la exploración de las funcionalidades del SMIL en la generación de subtítulos o información relacionada con los videos presentados al usuario, queda abierta como trabajo futuro.

Finalmente, la exploración de las posibilidades que se abren en la interacción de SMIL con otros estándares del consorcio W3C es otra alternativa enormemente interesante.

6 Apéndice.

Código fuente del ejemplo 2

```
<?xml version="1.0" encoding="UTF-8"?>
...
<head>
  <layout>
    <root-layout title=" RADIO IMAGEN " id="layout1" width="300" height="400"/>
    <region id="R1" left="0%" right="0%" width="300" height="300" fit="fill"/>
    <region id="R2" left="0%" right="0%" width="300" height="100" top="300"
      fit="fill"/>
    <region id="R3" left="0%" right="0%" fit="fill"/>
  </layout>
  <transition id="tran1" type="barWipe" subtype="leftToRight" dur="4s" />
  <transition id="tran2" type="irisWipe" subtype="rectangle" dur="4s"
    direction="reverse"/>
  <transition id="star1" type="starWipe" subtype="fivePoint" dur="4s" />
  <transition id="star2" type="starWipe" subtype="fivePoint" dur="4s"
    direction="reverse"/>
  <transition id="fade1" type="fade" subtype="crossfade" dur="4s"/>
  <transition id="fade2" type="fade" subtype="crossfade" dur="4s"
    direction="reverse"/>
```

```

</head>
<body>
  <par>
    <audio src="./ManerasDeVivir.mp3" id="audio1" region="R3" end="240s"/>
    <seq repeatCount="12">
      <par>
        
        <textstream src="./kannon.rt" region="R2" dur="10s"
          transIn="fade1" transOut="fade1"/>
      </par>
      <par>
        
        <textstream src="./skornabois.rt" region="R2" dur="10s"
          transIn="fade1" transOut="fade1"/>
      </par>
    </seq>
  </par>
</body>
</smil>

```

7 Referencias

- [Pañeda2004] Xabiel Garcia Pañeda, David Melendi, Manuel Vilas, Isabel Rodríguez, Ricardo Bonis. Fesoria: An integrated tool for performance and content analysis, SLA evaluation, management and smart presentation for video-on-demand services. International Conference on E-Business and Telecommunication Networks (ICETE). Setubal, Portugal 2004.
- [RealNetworks] Plataforma de streaming Real Networks. Accedido por última vez el 11-11-2006 desde <http://www.realnworks.com>
- [RealText] Real Text. Accedido por última vez el 11-11-2006 desde: <http://service.real.com/help/library/guides/realtext/realtext.htm>
- [SMIL 1.0] <http://www.w3.org/TR/REC-smil/>
- [SMIL 2.0] <http://www.w3.org/TR/2005/REC-SMIL2-20050107/>
- [SMIL 2.1] http://www.w3c.es/prensa/2005/nota051213_smil
- [Smil 1] <http://www.w3.org/TR/SMIL2/>
- [WindowsMedia] Plataforma de streaming Windows Media. Accedido por última vez el 11-11-2006 desde: <http://www.windowsmedia.com/>