UNIVERSIDAD CARLOS III DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

INGENIERÍA TÉCNICA EN INFORMÁTICA DE GESTIÓN

PROYECTO FIN DE CARRERA

EVALUACIÓN DE ESTÁNDARES SOBRE MATERIALES AUDIOVISUALES SUBTITULADOS Y AUDIODESCRITOS: ISAN, TV-ANYTIME y ESCORT

Autora: Lourdes Crespo Fuentes
Tutora: Ana María Iglesias Maqueda

Noviembre 2009

A mis padres.

"Le secret de mon bonheur est de ne pas lutter pour le plaisir,mais à trouver du plaisir dans l'effort*", André Gidé.*

AGRADECIMIENTOS

A mis padres por el apoyo prestado en todo momento, no solo a la hora de realizar este trabajo, sino a lo largo de mi vida.

Sin ellos no estaría donde estoy ni sería lo que soy.

A mi familia por su ayuda, paciencia y confianza.

A Elodie por su apoyo, consejo, cariño y largos días de estudio juntas, además de la gran amistad que nos une.

A mi tutora Ana porque su dirección ha sido fundamental a la hora de encauzar y dar formato a este proyecto.

A todos vosotros, GRACIAS.

ÍNDICE

1. INTRODUCCIÓN 8
1.1. Objetivos 9
1.2. Estructura del documento 9
2. GESTIÓN DEL PROYECTO 10
2.1. Gestión software 11
2.1.1. Metodología 12
2.1.2. Ciclo de Vida 12
2.2. Organización del trabajo 14
2.3. Planificación inicial 14
2.3.1. Costes 18
3. ESTADO DEL ARTE 19
3.1. Dublin Core 19
3.2. Digital Video Broadcasting (DVB) 20
3.3. Escort 21
3.4. ISAN y V-ISAN 22
3.5. MPEG-7 22
3.6. P/META 23
3.7. SMEF 24
3.8. SMPTE 24
3.9. TV-AnyTime 25
3.10. Discusión 25
4. ISAN y V-ISAN 26
4.1. Introducción 26
4.2. Definición de obras audiovisuales según estándar ISAN 27
4.3. Supuestos semánticos complementarios a la especificación de ISAN 31
4.4. Esquema Entidad/Interrelación 32
4.5. Explicación del Esquema E/R 34
4.6. Supuestos semánticos adicionales al diagrama E/R 36
4.6.1. Dominios 36
4.6.2. Otros 37
4.7. Análisis Crítico 37
5. TV-ANYTIME 43
5.1. Introducción 43
5.2. Supuestos semánticos complementarios a la especificación de TV-AnyTime 46
5.3. Esquema Entidad/Interrelación 47
5.4. Explicación del Esquema E/R 49
5.5. Supuestos semánticos adicionales al diagrama E/R 52
5.5.1. Dominios 52
5.5.2. Otros 55
5.6. Análisis Crítico 55
6. ESCORT 57
6.1. Introducción 57
6.2. Estructura Escort 2007 58
6.3. Supuestos semánticos complementarios a la especificación de Escort 60

6.4. Esquema Entidad/Interrelación 60
6.5. Explicación del Esquema E/R 62
6.6. Supuestos semánticos adicionales al diagrama E/R 65
6.6.1. Dominios 65
6.6.2. Otros 71
6.7. Análisis Crítico 71
7. PROPUESTA DE NUEVO META-MODELO 73
7.1. Introducción 73
7.2. Supuestos semánticos complementarios a la especificación de la Propuesta 75
7.3. Esquema Entidad/Interrelación 77
7.4. Explicación del Esquema E/R 77
7.5. Supuestos semánticos adicionales al diagrama E/R 84
7.5.1. Dominios 84
7.5.2. Otros 86
7.6. Discusión 87
8. PLANIFICACIÓN FINAL Y ANÁLISIS DE COSTES 92
8.1. Planificación final 92
8.2. Costes 93
9. CONCLUSIONES Y TRABAJOS FUTUROS 96
9.1. Conclusiones 96
9.2. Trabajos futuros 97
10. GLOSARIO DE TÉRMINOS 98
11. BIBLIOGRAFÍA 100
11.1. Convencional 100
11.2. En Web 100

ÍNDICE DE FIGURAS

Figura 1 Gestión de un proyecto de Ingeniería Software 11
Figura 2 Ciclo de vida del proyecto 13
Figura 3 Diagrama WBS 15
Figura 4 Planificación Inicial - Diagrama Gantt 16
Figura 5 Tareas y Tiempos de Planificación Inicial 17
Figura 6 Ejemplo de Metadato Dublin Core 20
Figura 7 Ejemplo de DVB-S 21
Figura 8 Ejemplo de MPEG-7 23
Figura 9 Barra de colores de SMPTE 25
Figura 10 Código ISAN 27
Figura 11 Ejemplos de representación numérica de códigos ISAN 29
Figura 12 Diagrama E/R estándar ISAN y V-ISAN 33
Figura 13 Propuesta de identificación en entidad PARTICIPANT 38
Figura 14 Propuesta de identificación en entidad COUNTRY 38
Figura 15 Propuesta de relación de países e idiomas 39
Figura 16 Propuesta de almacenamiento de fechas en entidad AW 40
Figura 17 Propuesta de almacenamiento de actores principales 40
Figura 18 Propuesta de almacenamiento de episodios de una serie 41
Figura 19 Propuesta de identificación en entidad COMPANY 42
Figura 20 Diagrama E/R estándar TV-AnyTime 48
Figura 21 Formato de clasificación con herencia (géneros y subgéneros) 59
Figura 22 Ejemplo de clasificación con herencia de la dimensión CONTENT 59
Figura 23 Diagrama E/R estándar Escort 61
Figura 24 Diagrama E/R (1) propuesta de meta-modelo 78
Figura 25 Diagrama E/R (2) propuesta de meta-modelo 79
Figura 26 Ventajas y desventajas de ISAN, TV-AnyTime y Escort 87
Figura 27 Variación de costes Inicial-Final 93
Figura 28 Planificación Final - Diagrama Gantt 94
Figura 29 Tareas y Tiempos de Planificación Final 95

1. INTRODUCCIÓN

En 1966 se creó en el Reino Unido un identificador único para libros denominado ISBN (*International Standard Book Number*), únicamente con el fin de un uso comercial; pero rápidamente se convirtió en un estándar internacional, ISO 2108 por los años 70. ¿Por qué lo que comenzó como un simple método de organización de libros para unas librerías inglesas, llegó a convertirse en un estándar internacional totalmente oficial? Podemos responder rápidamente a esta pregunta si nos paramos a pensar unos segundos en la cantidad de escritos que pueden archivarse en una biblioteca y la dificultad de poder encontrar un determinado libro en ella. Con este sistema de clasificación y los sistemas informáticos, se pueden tener almacenados en una BD (base de datos) cualquier tipo de libro que creamos oportuno y con ello facilitar el acceso a dicha información.

Ahora pensemos detenidamente en el día a día.

Poco a poco los libros están perdiendo protagonismo y cobrando mayor importancia los fondos audiovisuales, pero el problema se plantea a la hora de realizar una clasificación normalizada y estandarizada de dichos ítems. No existe normativa alguna que nos ayude a esta tarea y por lo tanto se agrava el hecho de poder localizar una determinada obra audiovisual, dicha tarea de localización se vuelve compleja y duradera. Además con el avance que ha tenido la informática en estos años atrás se podría conseguir un sistema de clasificación que asigne un código identificativo a cada uno de ellos, con el cual se podría almacenar toda la información de un determinado objeto en la base de datos y así acceder en cuestión de segundos a todas las características de un determinado ítem. Es decir, generar un código parecido al ISBN que anteriormente hemos citado, pero en este caso que dicho identificador sea para programas de televisión, películas de cine, ...

Existen distintos estándares que definen normas de identificación para fondos audiovisuales, pero cada uno de estos se centra en un diferente sector de las obras audiovisuales, por lo tanto no podemos decir que son genéricos. Entre dichos estándares, dedicados al almacenamiento de información y la estructuración de metadatos, podemos citar los siguientes como los más conocidos y utilizados: ISAN [ISAN, 2002], V-ISAN [VISAN, 2004], SMEF [SMEF, 2004], P/META [P/META,

2005], TV-AnyTime [TV-AnyTime, 2000], DVB [7], SMPTE [3], MPEG [MPEG-7, 2003], ESCORT [Escort, 1995], Dublin Core [6].

En este proyecto se realiza un riguroso análisis de tres estándares de los anteriormente citados, concretamente de ISAN y V-ISAN, TV-AnyTime y Escort. Además de generar una propuesta de sistema de clasificación y numeración de fondos audiovisuales.

1.1. Objetivos

El principal objetivo de este proyecto de fin de carrera es crear un meta-modelo de identificación y clasificación para cualquier tipo de fondo audiovisual, de manera que trate las necesidades que un ítem de ese tipo solicite y que se identifiquen todas aquellas características que desean ser almacenadas.

Se ha de generar un esquema conceptual (Esquema Entidad/Interrelación-E/R) en el cual se plasmen todas las necesidades de información a almacenar, evitando en este la redundancia de datos, exigiendo un identificador único para poder enumerar de manera unívoca, generando un modelo de datos fuerte, eficaz y eficiente.

Para poder llevar a cabo este proyecto de fin de carrera ha sido inevitable el documentarse sobre la numerosa variedad de estándares de este tipo que existen en la actualidad, entre otro podemos hablar de Dublin Core, Tv-AnyTime, Escort, etc. Eligiendo tres de los más utilizados para realizar un análisis crítico de los mismos y teniendo éste como base de la definición de una propuesta de metamodelo de identificación y clasificación de fondos audiovisuales que almacenan características accesibles como subtitulado o audiodescripción de los mismos.

1.2. Estructura del documento

En este apartado del proyecto de fin de carrera se va a detallar la estructura del documento realizado a la vez que se describe brevemente el contenido del propio proyecto.

En el capítulo 2 se realiza una selección del tipo de metodología a seguir y el tipo de ciclo de vida del proyecto. Además se realiza una organización del trabajo, definiendo las tareas fundamentales para llevarlo a cabo junto con una planificación y estimación de recursos y tiempo necesarios para realizar éste.

En el capítulo 3 se presenta el estado del arte donde se resumen los estándares que a día de hoy se encuentran definidos como sistemas de clasificación de materiales audiovisuales.

En el capítulo 4, 5 y 6 se realiza un análisis crítico de los estándares ISAN y V-ISAN, TV-AnyTime y Escort respectivamente incluyendo en dicho análisis sus esquemas entidad/interrelación.

En el capítulo 7 se realiza una propuesta de metamodelo, además de realizar una comparación de dicho estándar con los anteriormente citados.

En el capítulo 8 se realiza una comparación de la planificación inicial del proyecto con la real. Además se realiza un estudio de la variación de costes sufridos con respecto a los que se habían propuesto.

En el capítulo 9 se realiza una breve conclusión del trabajo elaborado incluyendo diferentes mejoras que se podrán realizar en un futuro para reforzar más todos los puntos estudiados en este proyecto.

Por último, en los capítulos 10 y 11 se encuentran los epígrafes de Glosario de términos y la bibliografía, respectivamente.

2. GESTIÓN DEL PROYECTO

Al igual que cualquier proyecto de ingeniería es necesario realizar una buena gestión del mismo, de esta manera se podrá lograr la satisfacción de haber cumplir los objetivos definidos.

Dicha gestión es realizada con el fin de poder culminar todo el trabajo requerido en el proyecto dentro del alcance, coste y recursos definidos. Estos tres puntos son la clave de un proyecto, que si han sido correctamente planificadas el éxito estará asegurado.

En caso de que se realice una mala gestión del proyecto, los resultados serán nulos. En la Figura 1 se puede observar de manera irónica los resultados de una mala gestión de cualquier proyecto de ingeniería de software.

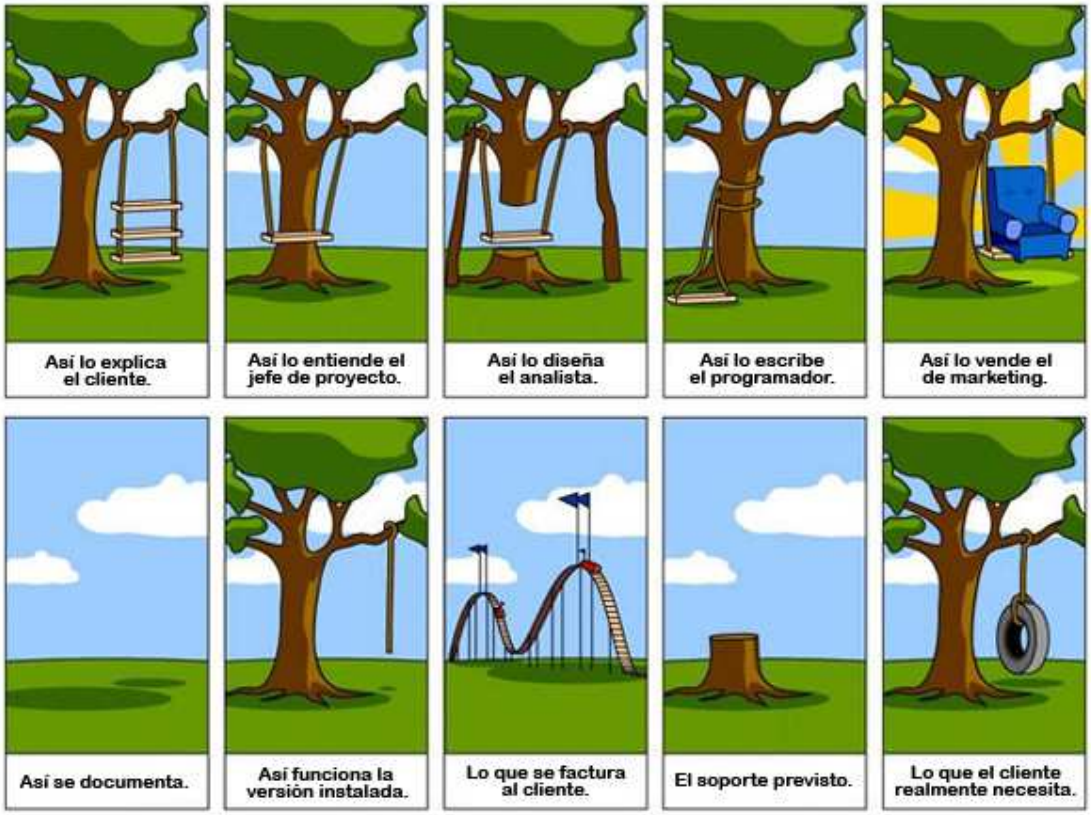

Figura 1 Gestión de un proyecto de Ingeniería Software

En este apartado del proyecto se plasman los conocimientos adquiridos durante las asignaturas de Ingeniería del software I –II y Gestión de Proyectos cursadas a lo largo de la carrera.

2.1. Gestión software

Para el desarrollo de todo proyecto de ingeniería de software debemos seguir una metodología de trabajo que nos garantice con un determinado grado de certeza de que el producto software resultante va a contar con una serie de características que definen lo que denominamos como calidad de software.

- En primer lugar debemos definir la "eficacia". La eficacia es la característica que determina si el producto cumple todos los requisitos que se establecieron inicialmente.
- En segundo lugar definimos la "eficiencia", que es una medida que establece, si los requisitos se cumplen de una manera equilibrada en coste de recursos y tiempo.

- En tercer lugar definimos la "corrección" como una medida de la cantidad de errores como las divergencias en el resultado esperado.
- Por último definimos la "satisfacción" como el grado en el que el producto se adapta con exactitud a lo que pidió el cliente o demandaban los usuarios.

Para alcanzar estos objetivos y lograr así la calidad en el producto final necesitaremos seguir una serie de pasos en orden, integrándolos en un marco de trabajo adecuado a las tareas que se realizarán, que es lo que conocemos como "proceso".

Para hacer esto, lo más recomendable y lo que más garantías de calidad nos ofrece es seguir una metodología de ingeniería de software.

2.1.1. Metodología

Una metodología de ingeniería de software es una estrategia de desarrollo que acompaña al proceso, así como un conjunto de métodos técnicos para completar las etapas a través de las cuales el proyecto va creciendo para llegar a convertirse en el producto final. Un proceso fuertemente estructurado, guiado por una metodología bien elegida nos ayudará a obtener un producto de calidad.

Los procedimientos empleados a lo largo de este estudio se basan en las fases de modelados de datos, consiguiendo como objetivo los esquemas Entidad-Interrelación de cada estándar analizado.

2.1.2. Ciclo de Vida

Un modelo de ciclo de vida define el estado de las fases a través de las cuales se mueve un proyecto de desarrollo software.

En este proyecto se optó por un ciclo de vida en cascada con retroalimentación; dicho tipo de ciclo de vida se caracteriza por ofrecer la posibilidad de según se va avanzando etapa por etapa, dar marcha atrás, es decir, volver a la etapa anterior. Se puede observar dicho funcionamiento en la figura. El fin de esta "marcha atrás" es poder ir mejorando el análisis de los estándares que se van tratando.

Se va analizando estándar por estándar para poder generar una propuesta de estándar.

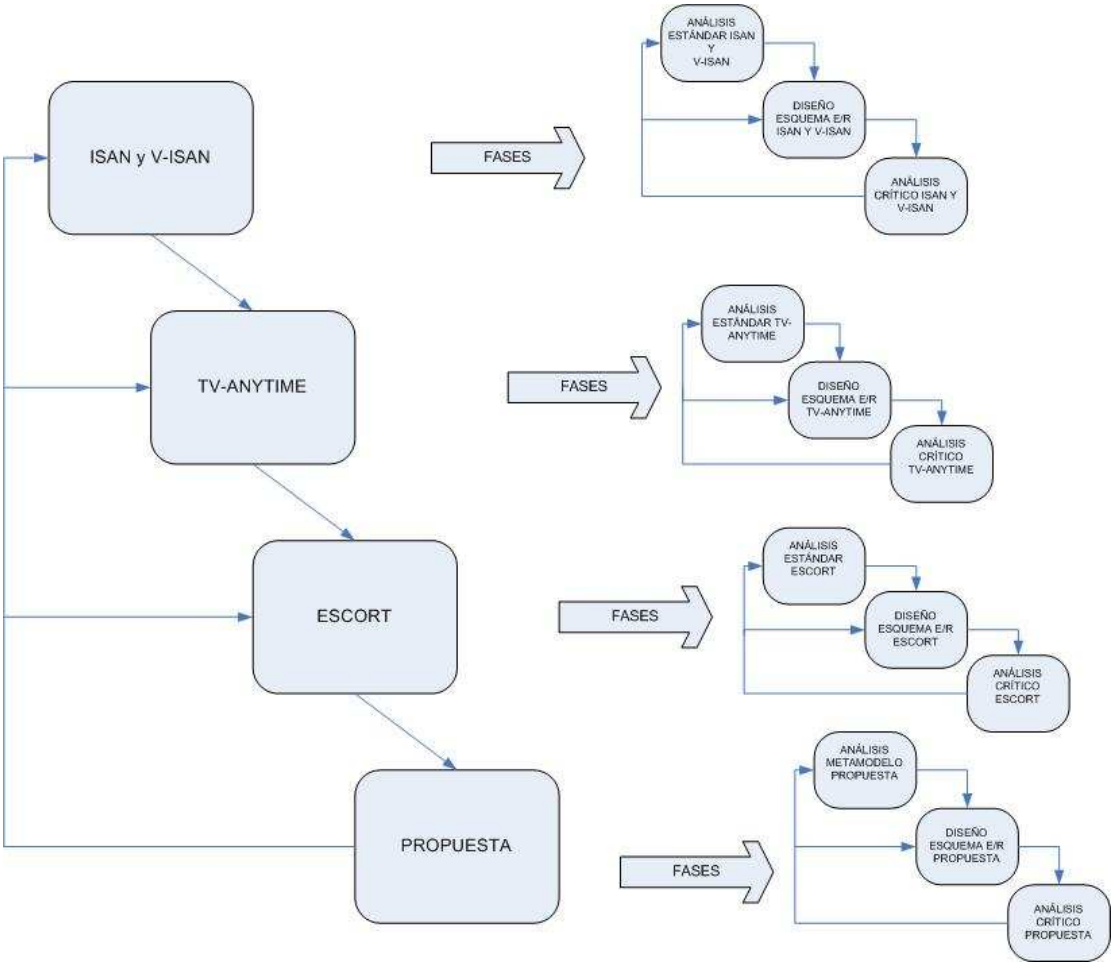

Figura 2 Ciclo de vida del proyecto

Cada fase del ciclo de vida se compone de varias tareas que a su vez son otro ciclo de vida en cascada con retroalimentación. Como se observa en la Figura 2 por cada estándar se realizan tres fases diferentes denominadas Análisis, Diseño de Esquema E/R y Análisis Crítico del estándar que en dicha fase se esté tratando. Es de gran necesidad un ciclo de vida en cascada con retroalimentación en las subfases ya que pueden ir surgiendo ideas, característica según se va realizando el estudio de dicho estándar, lo que supone una marcha atrás para la actualización de las diferentes fases anteriores que se vean involucradas en dichos cambios. De cada fase se obtiene un resultado:

- En la etapa de Análisis de estándar se genera un estudio de la definición que existe del estándar a estudiar.
- En la etapa de Diseño de Esquema E/R se genera el Diagrama Entidad-Interrelación junto con sus supuestos semánticos complementarios a éste.

- En la última fase se realiza un Análisis Crítico de las ventajas y desventajas que dicho estándar posee, además de generar propuestas de mejoras del estándar en estudio.

Todas estas fases se repiten a lo largo del ciclo de vida del proyecto, son fases que pertenecen a los tres estándares estudiados junto con la propuesta de meta-modelo que se genera en este proyecto de fin de carrera.

2.2. Organización del trabajo

Para llevar a cabo la organización de este trabajo, se han creado unas actividades elementales que surgen tras descomponer el proyecto en diferentes tareas. Para ello se ha generado un WBS (*Work Breakdown Structure*) en el cual se muestran dichas actividades (Figura 3).

2.3. Planificación inicial

Al igual que en cualquier proyecto, después de haber organizado claramente las tareas que se deben de llevar a cabo para poder sacar adelante el trabajo, se debe planificar el tiempo en el cual se va a realizar dicho proyecto.

Dicha planificación no es más que una estimación del tiempo que conllevará cada tarea en ser resuelta y de los recursos necesarios para poder realizarla.

La planificación temporal de proyectos se guía por unos principios básicos:

- División en módulos. El proyecto debe dividirse en un número manejable de tareas y actividades, descomponiendo tanto el producto como el proceso.
- Interdependencia. Se deben determinar las interdependencias entre actividades para su ordenación temporal en un gráfico de tiempos.
- Asignación de costes a las actividades: en tiempo, personal, trabajo y otros recursos.
- Distribuir tareas entre los miembros del grupo de trabajo y definir hitos y fechas de entrega para los resultados parciales.

La planificación realizada puede observarse en la Figura 4 a través de un diagrama de Gantt en el que se incluyen las fases a realizar, las fechas estimadas para elaborarlas y los recursos necesarios para cada tarea.

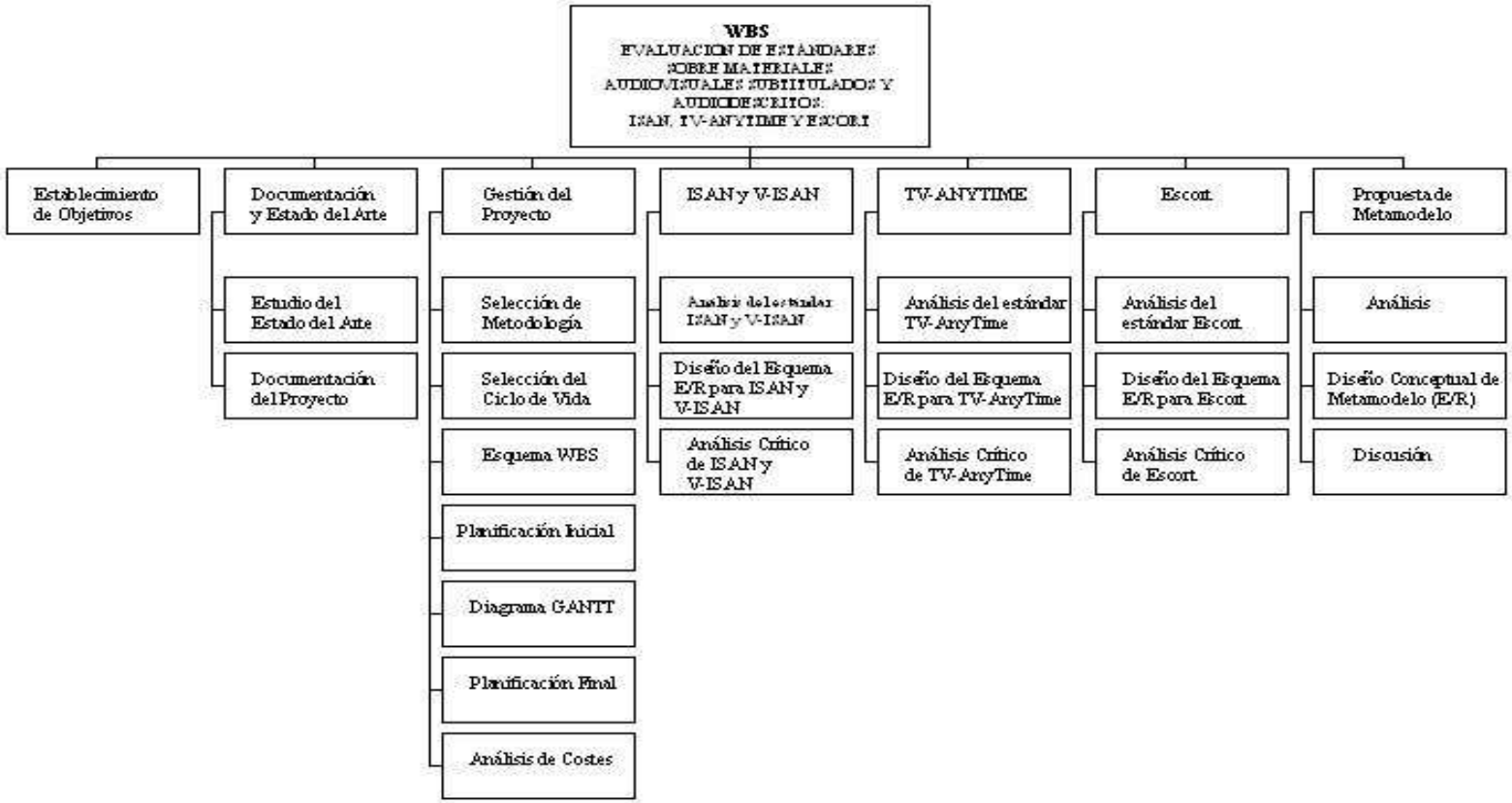

Figura 3 Diagrama WBS

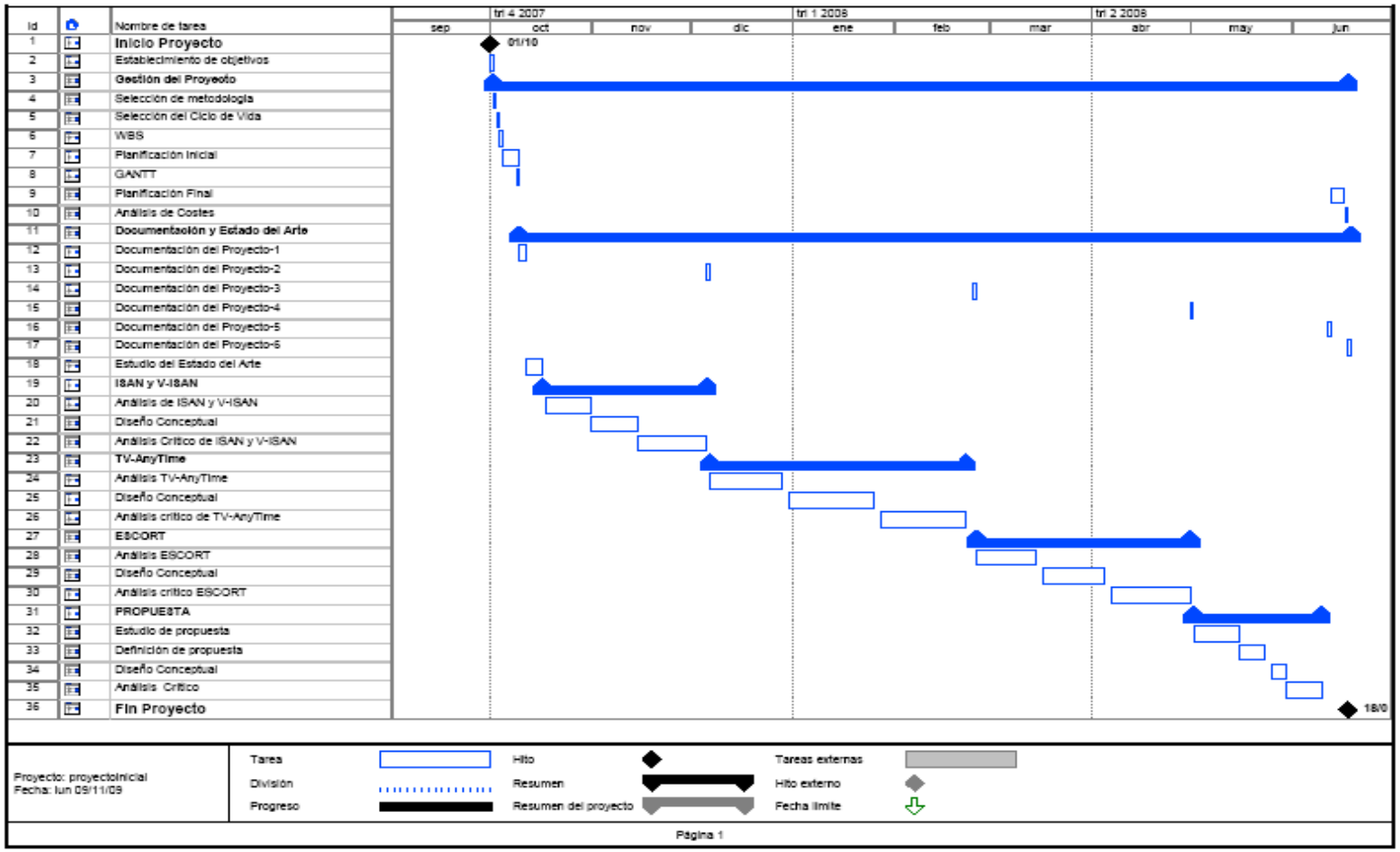

Figura 4 Planificación Inicial - Diagrama Gantt

Id	Nombre de tarea	Duración	Comienzo	Fin
1	***Inicio Proyecto***	0 días	lun 01/10/07	lun 01/10/07
2	**Establecimiento de objetivos**	1 día	lun 01/10/07	lun 01/10/07
3	**Gestión del Proyecto**	186 días	lun 01/10/07	mar 17/06/08
4	Selección de Metodología	1 día	mar 02/10/07	mar 02/10/07
5	Selección de Ciclo de Vida	1 día	mié 03/10/07	mié 03/10/07
6	WBS	1 día	jue 04/10/07	jue 04/10/07
7	Planificación Inicial	3 días	vie 05/10/07	mar 09/10/07
8	GANTT	1 día	mar 09/10/07	mar 09/10/07
9	Planificación Final	2 días	vie 13/06/08	lun 16/06/08
10	Análisis de Costes	1 día	mar 17/06/08	mar 17/06/08
11	**Documentación y Estado del Arte**	181 días	mar 09/10/07	mié 18/06/08
12	Documentación del proyecto-1	2 días	mié 10/10/07	jue 11/10/07
13	Documentación del proyecto-2	1 día	jue 06/12/07	jue 06/12/07
14	Documentación del proyecto-3	1 día	lun 25/02/08	lun 25/02/08
15	Documentación del proyecto-4	1 día	jue 01/05/08	jue 01/05/08
16	Documentación del proyecto-5	1 día	jue 12/06/08	Jue 12/06/08
17	Documentación del proyecto-6	1 día	mié 18/06/08	mié 18/06/08
18	Estudio del Estado del Arte	3 días	vie 12/10/07	mar 16/10/07
19	**ISAN y V-ISAN**	36 días	mié 17/10/07	mié 05/12/07
20	Análisis de ISAN y V-ISAN	10 días	mié 17/10/07	mié 31/10/07
21	Diseño Conceptual	10 días	jue 01/11/07	mié 14/11/07
22	Análisis crítico de ISAN y V-ISAN	15 días	jue 15/11/07	mié 05/12/07
23	**TV-AnyTime**	56 días	vie 07/12/07	vie 22/02/08
24	Análisis de TV-AnyTime	16 días	vie 07/12/07	vie 28/12/07
25	Diseño Conceptual	20 días	lun 31/12/07	vie 25/01/08
26	Análisis crítico de TV-AnyTime	20 días	lun 28/01/08	vie 22/02/08
27	**ESCORT**	47 días	mar 26/02/08	mié 30/04/08
28	Análisis de Escort	14 días	mar 26/02/08	vie 14/03/08
29	Diseño Conceptual	15 días	lun 17/03/08	vie 04/04/08
30	Análisis crítico de Escort	28 días	lun 07/04/08	mié 30/04/08
31	**PROPUESTA**	27 días	vie 02/05/08	lun 09/06/08
32	Estudio de Propuesta	10 días	vie 02/05/08	jue 15/05/08
33	Definición de Propuesta	6 días	vie 16/05/08	vie 23/05/08
34	Diseño Conceptual	4 días	lun 26/05/08	jue 29/05/08
35	Análisis Crítico	7 días	jue 29/05/08	lun 09/06/08
36	***Fin de Proyecto***	0 días	mié 18/06/08	mié 18/06/08

Figura 5 Tareas y Tiempos de Planificación Inicial

En la Figura 5 se detallan los períodos de tiempo para cada tarea. Hay que destacar que el horario laboral es de 8:00 a 16:00 y que la jornada es de Lunes a Viernes.

2.3.1. Costes

Una vez llevada a cabo la estimación de tiempos y recursos necesarios para el desarrollo del proyecto, debemos calcular el coste del mismo con objetivo de presentarle al cliente un plan de actuación, unas fechas de entrega y una valoración económica del coste que le va a suponer.

Para realizar la estimación se va a utilizar el método de valorar el precio por hora de cada recurso y multiplicarlo por el tiempo en horas que está previsto que invertirá en el proyecto.

Se han estimado los costes por recurso en relación al precio ofertado a los clientes por empresas consultoras, por lo que se ha establecido un salario de 30 euros por hora por el único necesario parar todas las fases del proyecto, cubriendo éste el papel de analista programador junior y siendo el propio autor de este proyecto de fin de carrera.

Hay que indicar que las jornadas laborales son de 8 horas diarias de Lunes a Viernes (inclusive).

El presupuesto económico del proyecto se valorará de la siguiente manera:

- Establecimiento de Objetivos: 1 día (8 horas).
- Gestión del Proyecto: 9 días (72 horas).
- Documentación y Estado del Arte: 10 días (80 horas).
- ISAN y V-ISAN: 36 días (288 horas).
- TV-AnyTime: 56 días (448 horas).
- ESCORT: 47 días (376 horas).
- Propuesta de metamodelo: 29 días (232 horas)

Es decir, se ha planificado que el analista tardará 188 jornadas laborales (1504 horas) en realizar dicho trabajo.

El presupuesto económico del proyecto se alza a un coste total de 45120 euros.

3. ESTADO DEL ARTE

Hoy en día existe un gran número de estándares de numeración y clasificación de medios audiovisuales, pero cada uno de ellos centrado en un problema y entorno concreto, no existiendo ninguno general.

En este capítulo del proyecto se hace una breve introducción de los estándares más conocidos de los cuales algunos serán estudiados de manera más profunda en este trabajo, en posteriores capítulos.

3.1. Dublin Core

Es un esquema de metadatos elaborado y auspiciado por la DCMI (*Dublin Core Metadata Initiative*). Se creó con el propósito de crear estándares que facilitarán la descripción y recuperación de recursos de información.

Las implementaciones de Dublin Core usan generalmente XML [10] y se basan en el Resource Description Framework (Framework para metadatos en la *World Wide Web (www)*). Dublin Core se define por ISO en la norma NISO Z39.85-2007 [13] y su norma ISO 15836 del año 2003 [14].

Es un sistema de 15 definiciones semánticas descriptivas que pretenden transmitir un significado semántico a las mismas. Dichas definiciones son opcionales, se pueden repetir y pueden aparecer en cualquier orden. Este sistema de definiciones fue diseñado para proporcionar un vocabulario de características “base”, capaces de proporcionar la información descriptiva básica sobre cualquier recurso, sin que importe el formato de origen, el origen cultural o el área de especialización.

Las características más destacables son la simplicidad, la flexibilidad, la independencia sintáctica, interoperabilidad semántica, alto nivel de normalización formal, el crecimiento y evolución del estándar a través de una institución formal consorciada (DCMI), consenso internacional, etc.

Los principales recursos que utilizan las descripciones Dublin Core son las páginas web pero se pueden destacar también las bibliotecas, organizaciones educativas, instituciones del gobierno, etc.

Como curiosidad cabe destacar que en septiembre del 2005 se celebró en la Universidad Carlos III de Madrid la Conferencia Internacional sobre Dublin Core y Aplicaciones de Metadatos [12].

Para completar la información sobre este estándar se puede consultar la página Web de Dublin Core [6].

En la Figura 6 se puede ver el registro de un fondo audiovisual en Dublin Core.

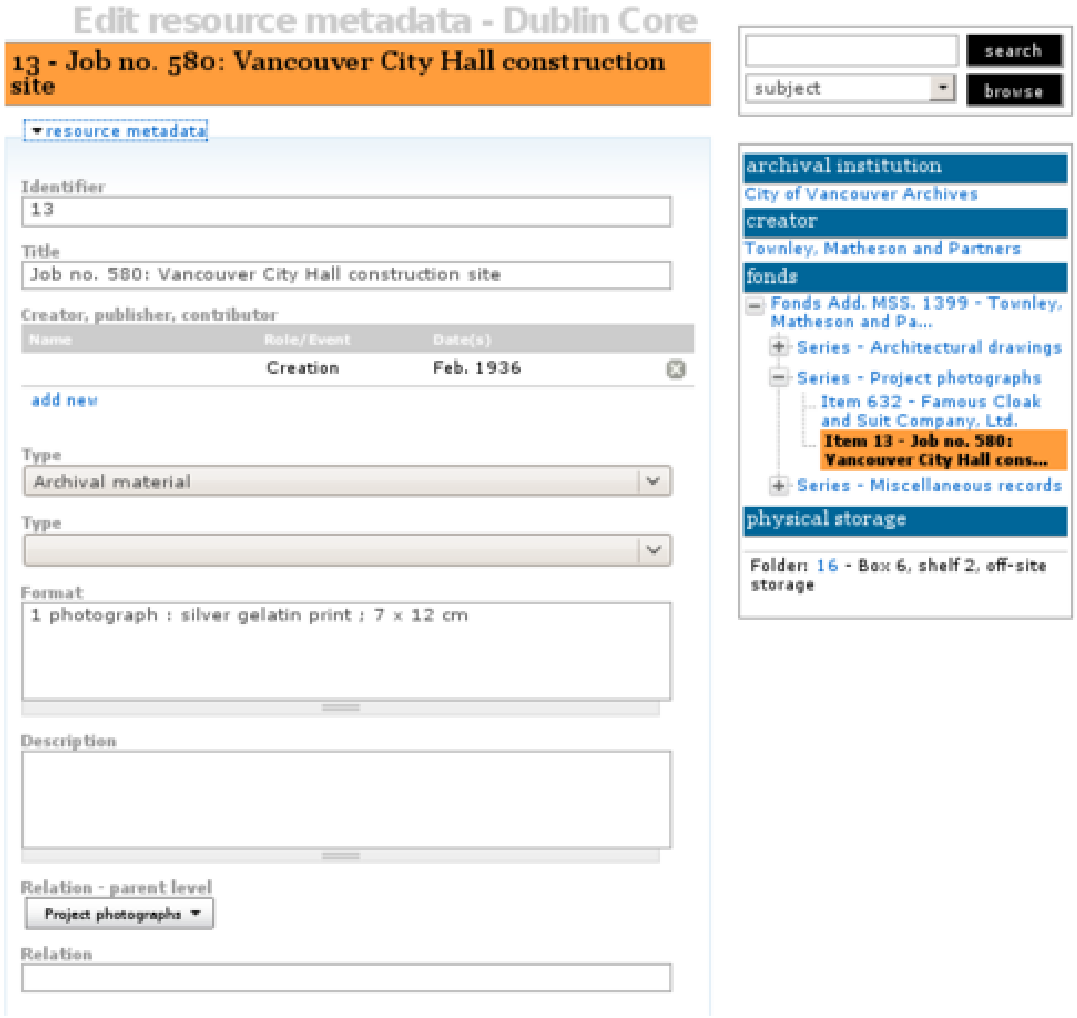

Figura 6 Ejemplo de Metadato Dublin Core

3.2. Digital Video Broadcasting (DVB)

Digital Video Broadcasting es una organización cuya función principal es el diseño y desarrollo de estándares globales orientados a la transmisión de televisión digital, incluyendo transmisión de datos.

Se creó en 1992 como iniciativa europea, pero en la actualidad se encuentra constituido por más de 300 instituciones y empresas de todo el mundo, entre ellos: fabricantes de equipos, *broadcasters*, operadoras de red, desarrolladores de software, etc.

DVB ha elaborado distintos estándares en función de las características del sistema de radiofusión, si es por satélite DVB-S (ver Figura 7), por cable DVB-C, por televisión terrestre DVB-T, por televisión terrestre para dispositivos portátiles DVB-H.

Se sitúa con una alta aceptación a nivel mundial, destacando prácticamente en toda la Unión Europea, gran parte de Asia, África y Oceanía.

En la página Web de DVB [7], viene detallado el conjunto de elementos que forman dicho estándar.

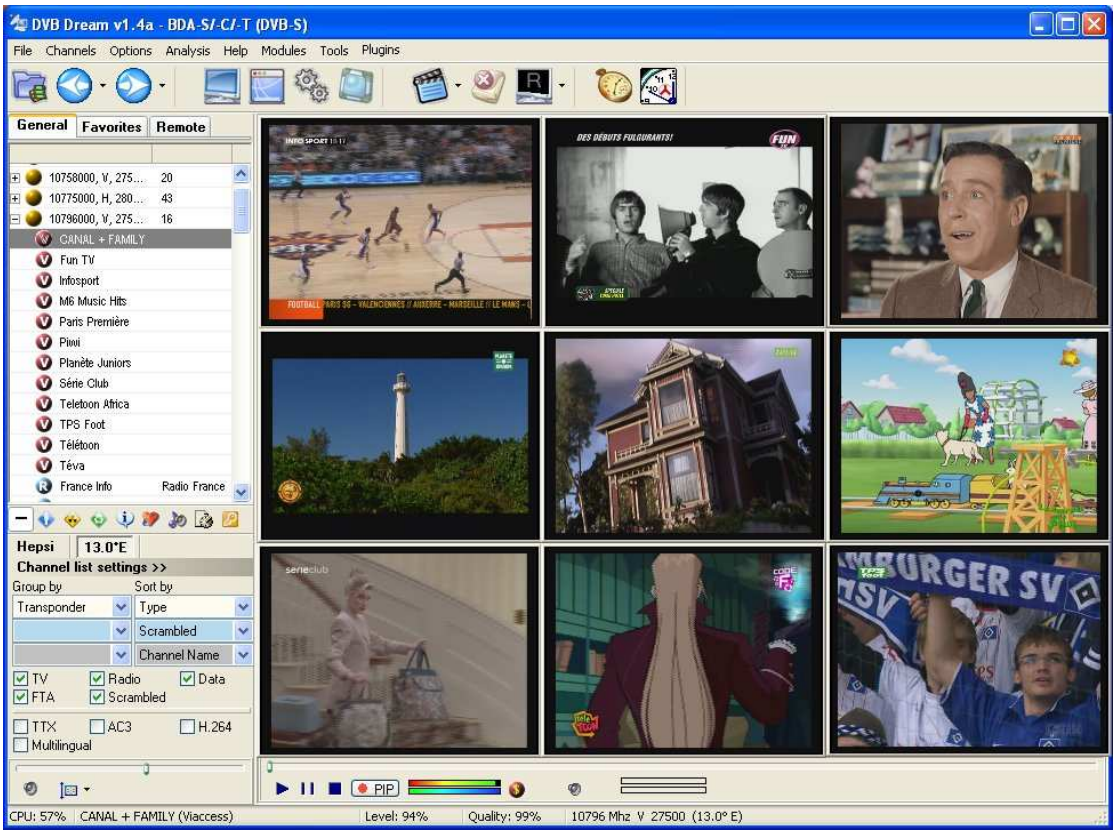

Figura 7 Ejemplo de DVB-S

3.3. Escort

Este estándar [Escort, 1995] es un sistema de clasificación de programas tanto de radio como de televisión. Fue diseñado en 2004 por EBU (*European Broadcasting Union*) y es de tipo multidimensional, cuyas dimensiones realizan una separación de géneros en categorías.

Se caracteriza por su utilidad en temas de financiación, contabilidad, comercialización y generación de informes de cumplimiento.

3.4. ISAN y V-ISAN

Es un sistema de numeración que identifica de forma unívoca las obras audiovisuales y sus versiones relativas (V-ISAN). Fue desarrollado por el Organismo Internacional para la Normalización ISO; es dirigido y gestionado por ISAN-IA (*ISAN International Agency*).

ISAN identifica las obras por todo el ciclo de vida, de la proyección a la producción hasta llegar a la distribución y el goce de dicha obra. Su característica clave es que tiene una estabilidad duradera; además puede emplearse tanto en medios físicos como digitales.

El código V-ISAN [V-ISAN, 2004] es una ampliación del código ISAN además de emplearse para las versiones de los trabajos audiovisuales.

El identificador ISAN está incorporado en muchos estándares como AACCS, DCI, MPEG, DVB y ATSC (muchos de ellos mencionados en este apartado).

3.5. MPEG-7

MPEG-7 [4] es un estándar de la Organización Internacional para la Estandarización ISO/IEC y desarrollado por el grupo MPEG. La primera versión se aprobó en julio del 2001 (ISO/IEC 15938 [15]) y la última versión publicada y aprobada por la ISO data de octubre del 2004. También es conocido por Interfaz de Descripción de Contenido Multimedia. Cabe destacar que MPEG-7 es uno de los diversos estándares creados por el grupo MPEG, entre los cuales podemos citar MPEG-1 (estándar de compresión de datos de video y de canales de audio asociados), MPEG-2 (dedicado originalmente a la televisión digital), etc. El grupo MPEG se creó en 1988 con el fin de desarrollar estándares internacionales de compresión, descompresión, procesamiento y codificación de imágenes animadas y datos de audio.

MPEG-7 consiste en una representación estándar de la información audiovisual que permite la descripción de contenidos (metadatos) para palabras clave, significado semántico (quién, qué, cuándo, dónde) y significado estructural (formas, colores, texturas, movimientos, sonidos).

Se basa en el lenguaje XML de metadatos y trata información de audio, voz, vídeo, imágenes, gráficos y modelos 3D.

Dicho estándar se aplica en bibliotecas digitales (almacenaje y búsqueda de bases de datos audiovisuales), vigilancia (control de tráfico, cadenas de producción), entretenimiento (búsqueda de juegos, karaokes), etc.

En la Figura 8 se puede ver un ejemplo de MPEG-7.

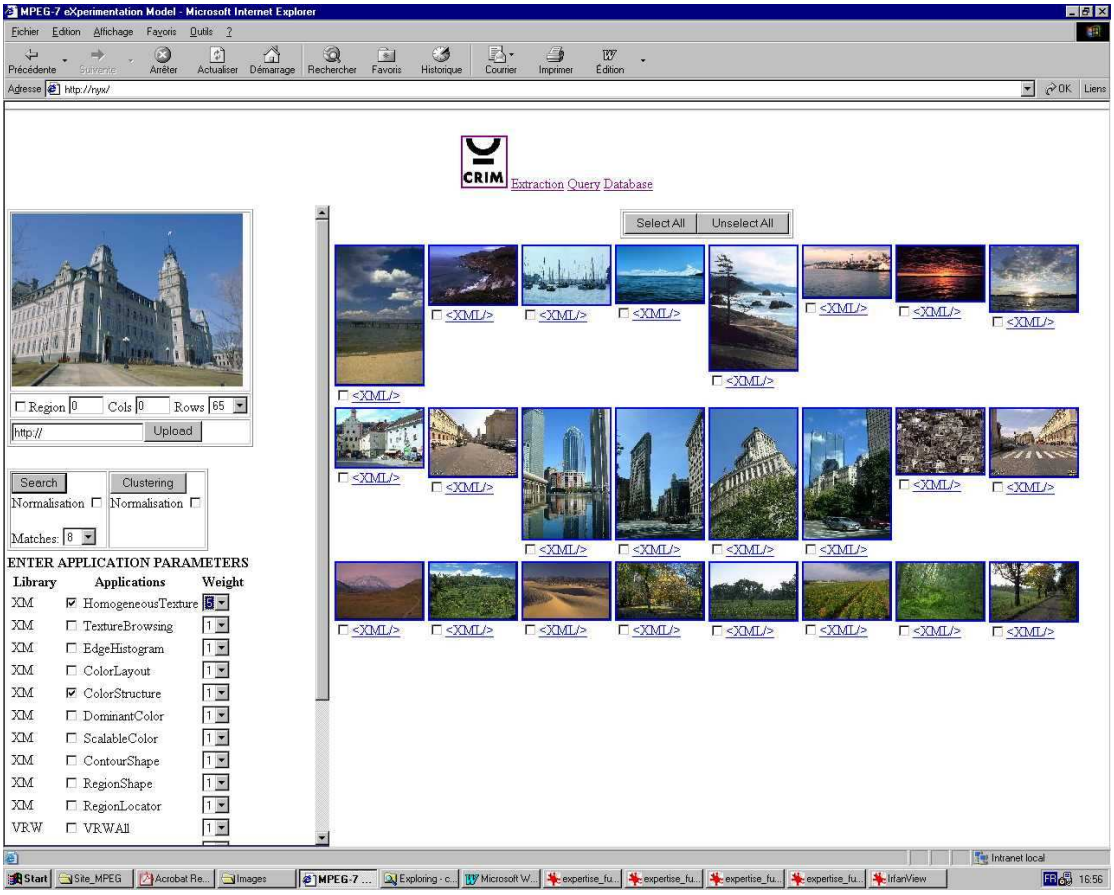

Figura 8 Ejemplo de MPEG-7

3.6. P/META

Es un estándar universal de intercambio de metadatos entre la industria del cine y la televisión, desarrollado por miembros del grupo EBU.

En realidad, define los metadatos para la identificación, descripción y uso del contenido en transacciones *Business-to-Business* (B2B).

Este estándar se caracteriza por tratar de manera independiente los ítems de idioma y sistema. Además es un sistema flexible ya que abarca un gran número de tipos de obras audiovisuales y por lo tanto, puede emplearse en una gran variedad de dichos fondos. Soporta cualquier implementación que esté desarrollada en XML o KLV.

P/META es un estándar que comparte cierta relación con el estándar TV-AnyTime analizado a lo largo del presente documento.

3.7. SMEF

El estándar SMEF fue definido por el grupo BBC y se constituye por un marco de intercambio de medios audiovisuales. Dicho estándar contempla información tanto de concepto como de contenido de los materiales.

El modelo de datos de SMEF proporciona un conjunto de definiciones de información requerida tanto en producción como en distribución como en administración de recursos de medios de comunicación, definidos mediante un diccionario de datos y un conjunto de diagramas E/R.

SMEF ganó el prestigioso premio *Royal Television Society Technology Innovations* en la categoría de I+D (Investigación y Desarrollo).

3.8. SMPTE

Society of Motion Picture and Television Engineers, fundada en 1916 en Nueva York como SMPE (*Society of Motion Picture Engineers*), es la organización norteamericana encargada de crear los estándares de la Industria Audiovisual. Tiene alrededor de 400 estándares los cuales incluyen:

- La barra de colores de SMPTE (Figura 9) empleada como patrón de pruebas y otras herramientas de diagnóstico.
- El formato MXF (*Material Exchange Format*), desarrollado por SMPTE. Es un formato de vídeo digital y de audio.
- Todos los formatos de cine y televisión.
- Interfaces físicas para la transmisión de señales de televisión y datos relacionados como el código de tiempo SMPTE y el Serial Digital Interface.

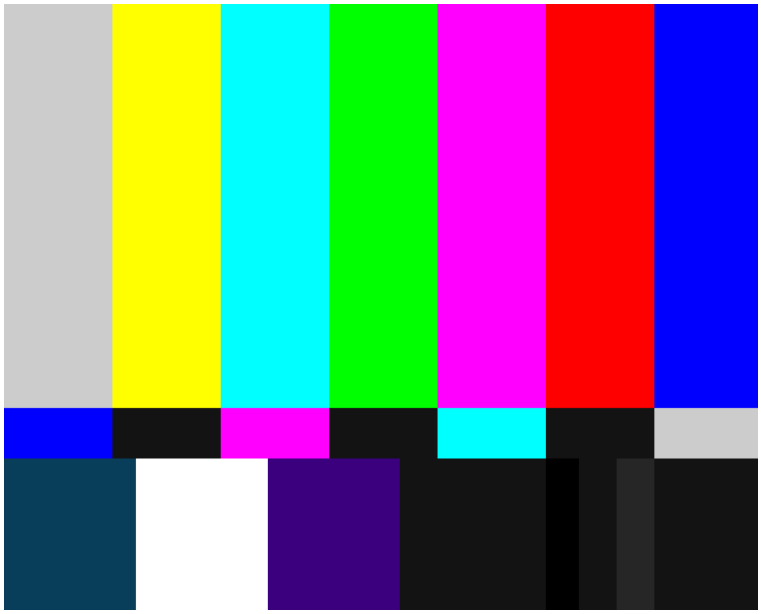

Figura 9 Barra de colores de SMPTE

3.9. TV-AnyTime

Es un estándar que describe la estructura de categorías de programas televisivos y de radio. Fue creado por Tv-AnyTime Forum y fue especificado en dos partes, la primera en agosto del 2003 y en la segunda parte en el año 2005.

TV-AnyTime identifica el contenido audiovisual a través del CRID ubicando así dicho contenido tanto en tiempo como en localización y basándose en todo momento en el amplio mundo de los metadatos.

3.10. Discusión

La motivación de este proyecto de fin de carrera se basa en la necesidad de un estándar que englobe todos los materiales audiovisuales y que éste facilite las tareas de encontrar, seleccionar y filtrar los contenidos de dichos materiales.

Tras realizar un breve análisis de los estándares más conocidos a día de hoy, se observa que ninguno de ellos resuelve dicha necesidad porque ninguno es general, es decir, todos dependen de un dominio, presentan redundancias de datos además de no tener mecanismos de almacenamiento eficientes de recursos accesibles (subtitulado y audiodescripción).

Se ha realizado un estudio de varios estándares importantes: ISAN y V-ISAN, TV-AnyTime y Escort para poder elaborar una propuesta que resuelva la necesidad existente citada anteriormente.

4. ISAN y V-ISAN

4.1. Introducción

El estándar ISAN (*Internacional Standard Audiovisual Number*), es un sistema internacional de numeración de trabajos audiovisuales (tratando las propias obras como sus diferentes versiones) identificándolos de manera unívoca mediante una clave única.

Dicho estándar fue creado por ISO (*Internacional Organization for Standarization*), concretamente por el grupo de trabajo TC46/SC9 (*Technical Comitee 46, Sub-Comitee9*).

La primera publicación del estándar se realizó en el año 2002 y fue denominado como ISO 15706. En el año 2006, se creó una extensión del estándar ISAN (V-ISAN) denominado como ISO 15706-2, el cual identifica las versiones asociadas a un determinado fondo audiovisual. Dichas publicaciones se encuentran tanto en Inglés como en Francés.

En Enero del 2004 se creó una página web específica para este estándar [1] en la cual aparece toda la información sobre el funcionamiento del sistema, la creación de un nuevo código ISAN, al igual que trata el registro on-line de un fondo. Además en Junio del mismo año se creó el sistema de ISAN con su propio repositorio central de datos.

Actualmente el código ISAN se emplea en un gran número de estándares, entre ellos están AACS (*Advanced Access Content System*) y MPEG, siendo gestionado y dirigido por ISAN-IA. ISAN-IA es una asociación, formada por AGICOA (*Association of Internacional Collective Management of Audiovisual Works*), CISAC (*Internacional Confederation of Societies*) y FIAPF (*Internacional Confederation of Societies*), fundada en 2003 y ubicada en Ginebra cuyo fin es promover el estándar entre la industria audiovisual, además de realizar el mantenimiento de la base de datos, gestionar, implementar y dirigir el sistema, así como acreditar y autorizar a las agencias la posibilidad de registrarse a nivel mundial.

Los usuarios de este sistema son tanto autores como productores y empresas. El objetivo es que todos los usuarios empleen el mismo código para evitar errores en el ciclo de vida del fondo audiovisual.

4.2. Definición de obras audiovisuales según estándar ISAN

Dicho estándar almacena fondos audiovisuales. Como trabajo audiovisual a todo aquel trabajo que está formado por una secuencia de imágenes relacionadas (con o sin sonido), con el fin de que en todo su conjunto se vea como una única imagen en movimiento a través del empleo de diferentes dispositivos tecnológicos.

El fin de este estándar es permitir la identificación, mediante la asignación de un número de referencia unívoco y reconocido a nivel mundial de las obras audiovisuales y sus versiones relativas, ya registradas en el sistema ISAN. El formato del número de referencia aparece ilustrado en la Figura 10.

Figura 10 Código ISAN

Se encuentra formado por 24 dígitos hexadecimales (96 bits), de los cuales 16 dígitos son obligatorios y los 8 restantes son opcionales, ya que éstos nos indican la versión.

Primeramente se inserta la palabra ISAN (con un espacio detrás de éste) y a continuación se indican los 24 dígitos que constituyen el número de referencia.
Los primeros 16 dígitos (divididos en 4 partes de 4 dígitos cada uno) se dividen en dos campos:

- 12 dígitos para identificar al programa audiovisual original.
- 4 siguientes dígitos para identificar el número de episodio que dicho T.A. obtiene dentro de la serie a la que pertenece. Si estamos tratando un T.A. que no forma parte de una serie, el campo episodio contendrá el valor 0000.

A continuación el código puede contener 8 dígitos de tipo hexadecimal (divididos en 2 partes de 4 dígitos cada una), los cuales indican la versión de dicho trabajo audiovisual. Tenemos que destacar que si un trabajo audiovisual no tiene partes o episodios sucesivos asociados, dicho campo no adquirirá el valor 00000000, directamente no se indica en el código.

Existen dos caracteres de control para verificar posibles errores de trascripción, vienen definidos por una letra comprendida entre la A y la Z (asignada

automáticamente por un algoritmo matemático, el cual viene definido en el Anexo A del documento ISAN), y aparecen tras cada parte anteriormente definida; es decir, 16 dígitos + 1 dígito de control (para el programa), y 8 dígitos + 1 dígito de control (para la versión).

Hay que destacar que todo aquel código ISAN que comience por 1111, indica que es parte de las aplicaciones internas del propio estándar, por lo que nunca se podrá registrar un T.A. con esa raíz en el código.

En la Figura 11 se presentan algunos ejemplos de códigos ISAN desglosados por sus diferentes campos. Además se indica un ejemplo con y sin versión asociada.

Un código ISAN será válido siempre y cuando el fondo audiovisual esté terminado, no se encuentre en desarrollo; en caso contrario, durante el desarrollo de éste se le podrá asignar un ISAN en desarrollo pero ese mismo carece de validez en el sistema.

Los trabajos audiovisuales, a los cuales se les pueden asignar un código ISAN, tienen que estar entre los siguientes tipos:

- Anuncio
- Película
- Programa de televisión
- Documental
- Programa educativo
- Serie, telenovela ...
- Cortometraje
- Programa multimedia
- Telediario
- Evento en vivo
- Representación
- Evento deportivo
- Programa de entretenimiento
- Videoclip

Nunca podrá ser un trabajo audiovisual aquel programa que sea del tipo:

- Grabación de audio
- Banda sonora
- Partitura

- Guión
- Relato
- Imagen de un T.A.
- Fotograma
- Cualquier otro tipo de audiovisual no definido anteriormente.

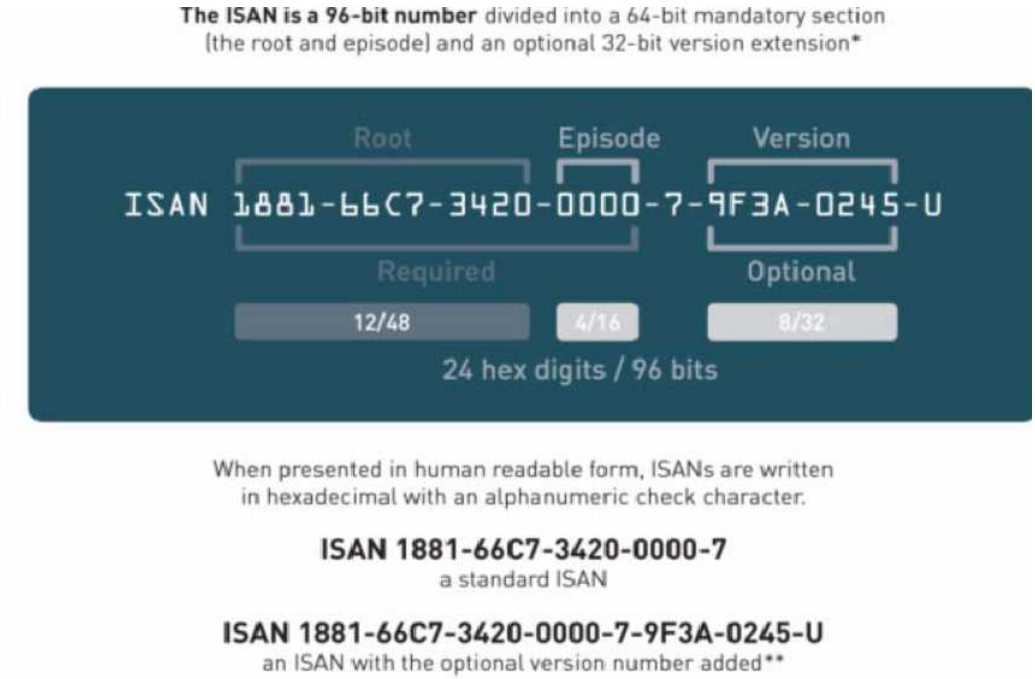

Figura 11 Ejemplos de representación numérica de códigos ISAN

En este estándar se clasifican las obras audiovisuales en dos tipos: Compuesto y No compuesto.

Un trabajo compuesto lo denomina como todo aquel T.A. que está formado por varios trabajos audiovisuales y un trabajo no compuesto, es todo aquel trabajo que no contiene partes de otros.

Un trabajo audiovisual nunca podrá ser compuesto si es:

- Un montaje de imágenes
- Una escena específica de una película

Todo aquel trabajo que sea de los siguientes tipos se tratará como trabajo audiovisual compuesto:

- Programa de televisión que contenga secuencias de episodios de anteriores películas
- Episodio que contenga secuencias de episodios de otras series televisivas

Como hemos indicado anteriormente en el documento, el estándar ISAN trata a los episodios o capítulos de un trabajo como otro trabajo audiovisual propio, heredando algunas características del trabajo al que están asociados. Pero existen obras audiovisuales, que según se encuentra especificado el estándar no admiten ser parte de una serie:

- Una nueva parte de una película
- Retransmisiones de partes separadas de un único trabajo
- Un trabajo que se retransmite una vez al año

La segunda publicación del estándar ISAN, añade el concepto de versión de un trabajo audiovisual, se genera una versión cuando se produce un cambio en las propiedades del trabajo. Pero al igual que hay trabajos que nunca puede ser compuestos o nunca pueden ser episodios/series, también ocurre con las versiones. Los cambios que dan lugar a la creación de una nueva versión son:

- Cambios en el idioma
- Cambios en el subtitulado
- Cambios en la imagen o en el formato del sonido
- Cambios en la duración del trabajo
- Cambios en la publicidad
- Cambios técnicos que afectan al contenido del propio trabajo
- Cambios en los derechos o propietarios del T.A.
- Cambios en el tipo de grabación

Existen otros cambios relevantes, que afectan al trabajo en sí, pero dichos cambios no generan una nueva versión, en cambio si generan un nuevo código ISAN. Dichos cambios son:

- Creación de una versión reeditada y resumida.
- Creación de una nueva secuencia de dicho trabajo.

4.3. Supuestos semánticos complementarios a la especificación de ISAN

Este apartado explica los conceptos que no han quedado claros, que son ambiguos o contradictorios en la especificación de ISAN que estamos tratando.

- Un participante puede ser director de varios trabajos audiovisuales pero en un trabajo audiovisual nada más que puede haber un único director. Así lo define ISAN.
- Un participante puede ser actor en distintos trabajos audiovisuales pero en un T.A. como mínimo debe haber 3 actores (dichos actores serán los principales), para los tres se almacenarán los apellidos como clave unívoca y el nombre como dato opcional.
- En un trabajo audiovisual puede haber diferentes participantes con el mismo o diferente rol.
- En un mismo trabajo audiovisual pueden colaborar distintas empresas, pero una misma empresa puede colaborar en varios T.A.
- Un mismo trabajo audiovisual se ha podido producir en varios países, y en un mismo país se han podido producir diferentes trabajos audiovisuales.
- Un trabajo audiovisual puede ser traducido en varios idiomas y por cada traducción tendrá un nuevo título. Se requiere almacenar por cada idioma, el código de éste (especificado según IS0 392-3). Además, se debe quedar registrado si ese título es el original.
- Existen tres tipos de trabajos audiovisuales, aquellos que son compuestos, los que no lo son y los trabajos denominados series. Aunque realmente un trabajo audiovisual únicamente puede ser compuesto o no compuesto, además en ningún momento un trabajo compuesto podrá ser a la vez trabajo no compuesto.
- Un trabajo audiovisual compuesto estará formado por otros dos trabajos audiovisuales como mínimo.
- Un trabajo audiovisual no compuesto puede contener diferentes episodios (por lo que este trabajo audiovisual será una serie), o simplemente ser un único trabajo, como por ejemplo una película.

- Un trabajo audiovisual puede tener más de una versión especificada, pero no es condición necesaria que tenga al menos una versión cada fondo audiovisual.
- Para añadir un nuevo episodio a un trabajo audiovisual de tipo serie, se elige el título de la serie a la que pertenece y automáticamente se carga la información principal de dicho trabajo, para la posterior inserción o modificación de datos correspondientes a ese episodio específico.
- Si un trabajo audiovisual no tiene versión, no se almacena ninguna información para dicho campo.

4.4. Esquema Entidad/Interrelación

En la Figura 12 se presenta un diagrama conceptual definiendo los conceptos y relaciones más importantes a tener en cuenta en la bases de datos que da lugar al sistema del estándar ISAN. Es importante tener en cuenta que en este diagrama se ha plasmado exactamente la información que venía dada por las especificaciones de ISAN, no se han tenido en cuenta las posibles mejoras que han surgido tras un profundo análisis del problema. Las críticas a este esquema E/R se especifica en la sección 4.7 (Análisis Crítico).

Para la generación del esquema conceptual se ha utilizado el modelo Entidad/Interrelación [de Miguel et al., 1999] [Cuadra et al., 2007].

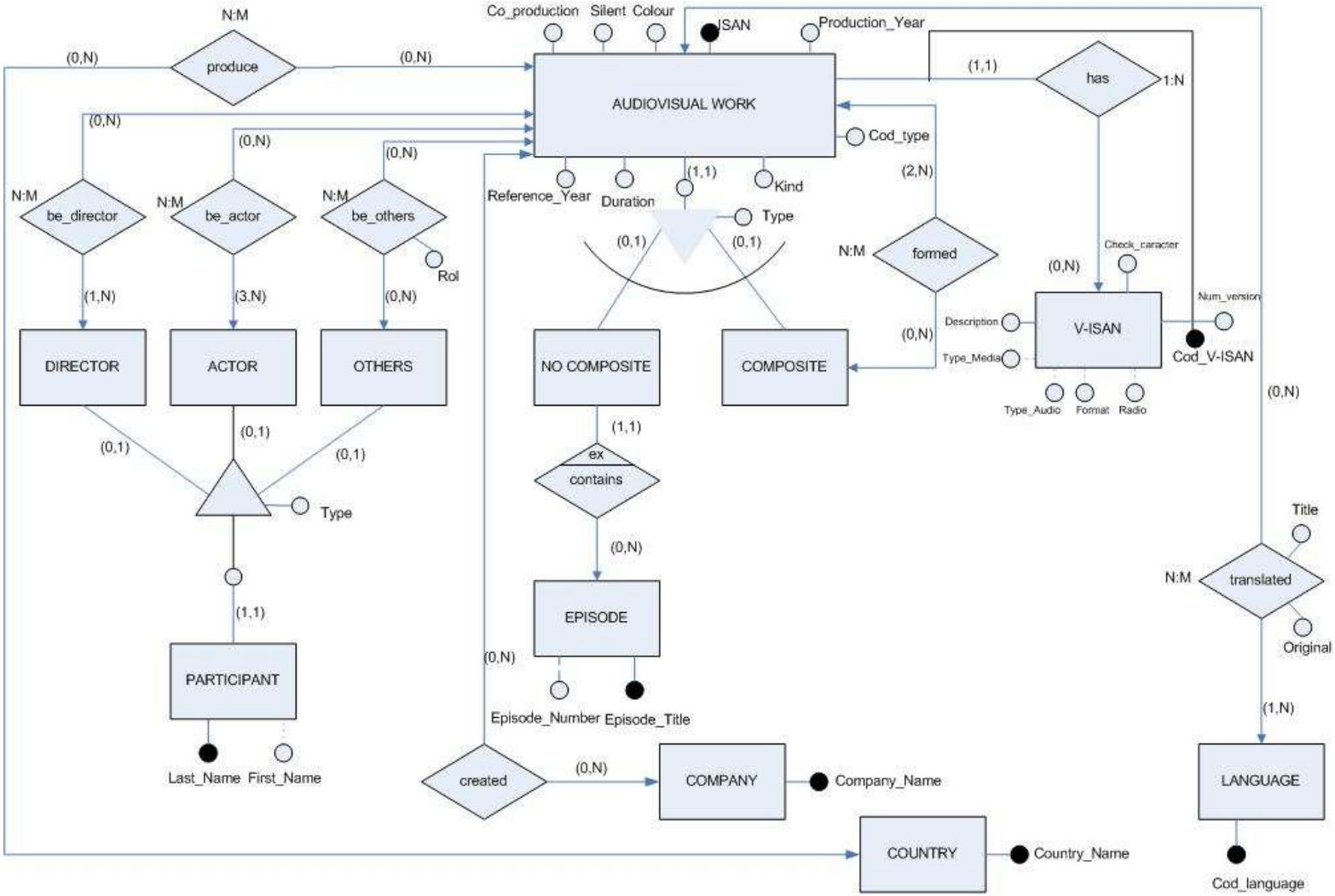

Figura 12 Diagrama E/R estándar ISAN y V-ISAN

4.5. Explicación del Esquema E/R

En este diagrama existe una entidad principal, denominada [*AUDIOVISUAL WORK*] en la cual se almacenan los datos principales de un fondo audiovisual. Dichos datos son los siguientes:

- Co_production: Indica si tiene o no co-producción dicho trabajo. Es un campo opcional.
- Silent: Indica si el trabajo audiovisual está grabado con sonido o sin éste. También es un campo opcional.
- Colour: Indica si el trabajo audiovisual está creado en color, en blanco y negro, o si es una mezcla de ambos; es un campo opcional.
- Production_Year: Indica el primer año en el que fue publicado el T.A.
- Reference_Year: Indica cualquier fecha clave del trabajo audiovisual, por ejemplo el año de producción del T.A.
- Duration: Indica el tiempo en minutos que dura el trabajo.
- Kind: Indica si el trabajo emitido es en vivo, si es animado, o simplemente si es una combinación de ambos.
- Cod_type: Indica el tipo de trabajo audiovisual que es, si es noticia, anuncio, serie, etc.

Se debe destacar la diferencia que existe entre el año de referencia y el año de producción. El año de referencia indica cualquier año identificativos para el trabajo audiovisual en cambio el año de producción indica únicamente el año en el que realmente se ha realizado la primera publicación de éste al público.

Esta entidad tiene como clave primaria el código ISAN y se relaciona con otras entidades necesarias para el sistema, a través de interrelaciones y diversos atributos que se especifican a continuación.

Entre las entidades [*COUNTRY*] y [*AUDIOVISUAL WORK*] se encuentra definida una interrelación de tipo N:M, llamada *produce*. El objetivo de esta interrelación es el de poder tener registrado qué países se han visto involucrados en el desarrollo de dicho fondo. La interrelación *created* de tipo N:M, cuyas entidades participantes son [*AUDIOVISUAL WORK*] y [*COMPANY*] Su fin es poder tratar las compañías que han colaborado en un mismo trabajo audiovisual. Existe otra

interrelación del mismo tipo N:M entre las entidades [*AUDIOVISUAL WORK*] y [*LANGUAGE*] llamada *translated*. Dicha relación tiene los atributos Title y Original. Cada trabajo audiovisual puede estar traducido en varios idiomas, además por cada traducción tendremos la información del título que obtiene dicho fondo en esa traducción y de si es o no el título original.

A la hora de tratar los participantes de un T.A. se ha optado por una jerarquía parcial con solapamiento, con supertipo [*PARTICIPANT*] y como subtipos los tres tipos de participantes que pueden trabajar en un fondo audiovisual, es decir, [*DIRECTOR*], [*ACTOR*] y [*OTROS*]. La entidad [*PARTICIPANT*] tiene como clave primaria Last_Name y como atributo opcional almacena el First_Name para cada participante. Además existen tres interrelaciones N:M (*Be_director, Be_actor, Be_others*), las cuales relacionan los directores, actores y otros tipos de participantes que se encuentran involucrados en un determinado fondo audiovisual, aunque debemos destacar que en la interrelación *Be_others* existe el atributo Rol, para saber qué rol tiene ese participante en ese fondo audiovisual.

Existe otra jerarquía parcial, pero en este caso es una jerarquía sin solapamiento, teniendo como supertipo [*AUDIOVISUAL WORK*] y como subtipos [*NO_COMPOSITE*] y [*COMPOSITE*]. Con esta generalización dividimos los fondos audiovisuales por trabajos simples o compuestos y así estarán diferenciados ambos tipos de trabajos. Aunque para los no compuestos, hemos definido una interrelación en existencia, denominada *Contains*, en la cual se relaciona con los episodios, ya que un episodio en sí, es un trabajo audiovisual simple. La interrelación *Contains* es de cardinalidad 1:N y de cada episodio se almacena su Episode_Number como atributo opcional y se identifica por su Episode_Title.

La interrelación denominada *formed*, cuyas entidades participantes son [*AUDIOVISUAL WORK*] y [*COMPOSITE*], tienen la particularidad de que aunque su cardinalidad es de tipo N:M, se especifica que un T.A. compuesto estará formado como mínimo por dos trabajos audiovisuales.

Por último, existe en el diagrama una entidad de gran importancia llamada V-ISAN, la cual tiene como identificador el Cod_V-ISAN, con una dependencia en identificación de la clave ISAN de la entidad [*AUDIOVISUAL WORK*].

4.6. Supuestos semánticos adicionales al diagrama E/R

En este apartado definimos la semántica que incorpora el estándar ISAN o que se ha especificado en el apartado anterior y que no puede ser expresada en el diagrama E/R generado, por las limitaciones impuestas por el Modelo E/R utilizado.

4.6.1. Dominios

- Un participante puede realizar un papel en un trabajo audiovisual, papel que se encuentra dentro del siguiente rango de valores:

  Dominio Type (jerarquía de Participant)={Actor, Director, Others}

  Y si elegimos el tipo Others, el atributo tendrá que estar entre los valores del siguiente dominio:

  Dominio Others = {Screenwriter, Scriptwriter, Distributor, Producer, Character, Composer, Anchor}
- El idioma en el que se produce el trabajo audiovisual, será almacenado mediante un código de idioma, el cual seguirá la norma ISO 639-2.
- El color en el que se produce el trabajo audiovisual, debe encontrarse entre los valores que pertenecen al siguiente rango:

  Dominio Colour={black and white, full colour, combination}
- La co-producción únicamente podrá admitir valores comprendidos en el siguiente dominio:

  Dominio Co-Production ={Yes, No}
- El atributo Silent del trabajo audiovisual solamente podrá admitir valores comprendidos en el siguiente dominio:

  Dominio Silent={Yes, No}
- El tipo del trabajo audiovisual, que viene definido por Cod_type, sólo puede admitir valores del dominio:

  Dominio Type={AD,FF,TF,DO,ED,SE,SH,MM,NE,LV,PF,SP,TE, VC}
- La clase (Kind) del trabajo audiovisual, tendrá un valor comprendido entre los siguientes:

  Dominio Kind = {Live Action, Animate, Mix}

4.6.2. Otros

- El valor del atributo Episode_Number de la entidad [*EPISODIO*] será un número identificativo de dicho capítulo en la serie, será un número correlativo para tener organizados todos los capítulos de la serie según se hayan ido creando.
- El atributo Title de la interrelación *translated* deberá estar expresado en sus respectivas mayúsculas y minúsculas, nunca todo el nombre en mayúsculas. Si existe artículo en el título, deberá ir delante, de ninguna manera se aceptarán títulos de la forma "XXXX, The".
- Los atributos de la entidad [*PARTICIPANT*] Last_Name y First_Name vendrán completos (nombre y dos apellidos), éstos sin abreviaciones y en sus respectivas mayúsculas y minúsculas; en caso de que se trate un nombre compuesto, todo el nombre irá en el campo nombre; todo apellido vendrá expresado por más de un carácter y nunca admitirá el valor "None".
- Los nombres de las compañías productoras que sean distribuidoras, de derechos de autor y grupos no serán almacenados; y en caso de que sean varias empresas de adquisiciones o de quiebras solamente se plasmará la compañía principal.
- Los campos Reference_Year y Production_Year de la entidad principal [*AUDIOVISUAL WORK*] serán del tipo YYYY, por ejemplo, 2001, etc.

4.7. Análisis Crítico

ISAN es un sistema cuyo fin es incluir beneficios para todos los usuarios que decidan emplear dicho estándar para registrar sus fondos audiovisuales, ya que de esta manera podrán tener registrados a nivel internacional sus trabajos y así facilitar la búsqueda de éstos.

Existen algunos conceptos que el estándar no especifica con tanto detalle como debería; además se podrían realizar algunas mejoras en el sistema para que dicho estándar fuera más eficaz y eficiente.

ISAN describe como clave principal de la entidad [*PARTICIPANT*] los apellidos de éste. Es un grave error identificar a un participante por ese campo ya que no es una clave única y podrían darse casos de duplicidad y por lo tanto error en la

información. Lo más correcto sería identificar a un participante por su DNI, ya que es una clave única de identificación. Además se podría añadir como clave secundaria de identificación el atributo compuesto Name (First_Name y Last_Name), ya que identificando únicamente por los apellidos, se nos puede presentar algún caso de que dos personas se apelliden igual. También sería una buena opción almacenar el nombre artístico si dicho participante tiene. En la Figura 13 se reflejan los cambios que habría que realizar para llevar a cabo la mejora en la entidad de participante.

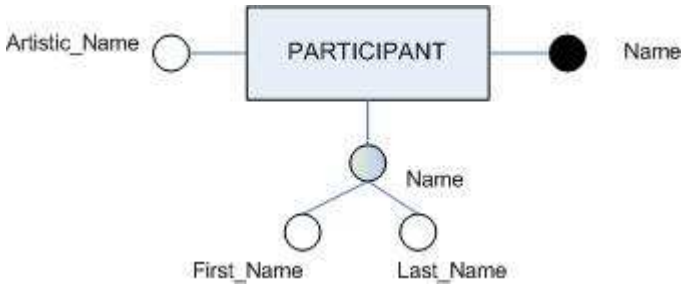

Figura 13 Propuesta de identificación en entidad PARTICIPANT

Para la entidad de [*COUNTRY*] añadiría una clave que fuera un identificador del país; ésta sería la clave principal de dicha entidad y el Name como atributo secundario. Según el estándar, únicamente identifica al país a través del nombre de éste, y es más eficiente si tenemos un código para cada país, al igual que se hace para los idiomas. Dicho código de país deberemos relacionarlo con el estándar ISO de código de países, para que sea más universal y genérica esa asociación. En la Figura 14 aparece la entidad país con los cambios que se acaban de proponer para mejorar la identificación de los países.

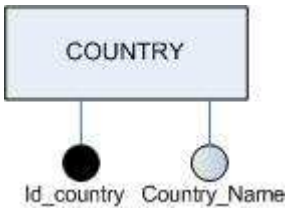

Figura 14 Propuesta de identificación en entidad COUNTRY

Según especifica ISAN en su estándar, no se almacena en ningún momento el país al que pertenece cada compañía productora del T.A.; por lo que si dicha información fuera almacenada en la base de datos, tendríamos una información más completa del trabajo audiovisual.

ISAN no relaciona el término País con el término Idioma, y a primera vista nos podemos dar cuenta que es una relación primordial; sería bueno que dicha relación se tuviera en cuenta en el estándar, ya que todo país tiene como mínimo un idioma, aunque la mayoría de países tienen más de uno. Aunque el estándar ISAN, trata los idiomas mediante códigos definidos por ISO 639-2, lo más eficiente sería tener almacenado por cada idioma el código (según ISO 639-2) y el nombre de dicho idioma y relacionarlo con los países en que se habla dicho idioma. Ver Figura 15.

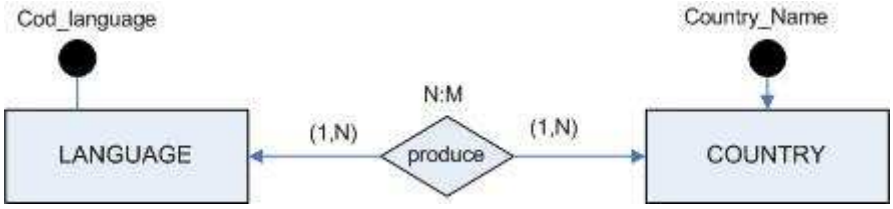

Figura 15 Propuesta de relación de países e idiomas

El estándar ISAN almacena por cada trabajo audiovisual tanto un año de referencia, como el año de producción (primera publicación). Por no almacenar información innecesaria, y por lo tanto disminuir espacio ocupado, se piensa que el año de referencia no tiene mayor sentido mantenerlo. El año que realmente necesitamos tener referencia, es el año en el que dicho trabajo se muestra por primera vez al público, es la única fecha importante de un trabajo audiovisual que sería necesario tener almacenada, por lo que no debe ser atributo opcional, sino atributo obligatorio del T.A.

Si se necesita guardar algún año que haga referencia al T.A. debería ir acompañado de un campo en el cual se indique esa fecha a que pertenece, por ejemplo, 1998-año en el que se empezó a crear el T.A.; ya que sino es muy relevante decir que 1998 es una fecha de referencia del trabajo.

En este caso, el subesquema E/R definido (Figura 16) no podría recoger la información de que en Year_of_reference no debería estar incluido el año de producción, ya que éste se recoge en el atributo Production_Year.

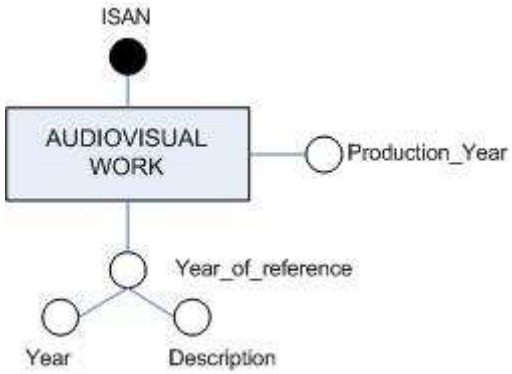

Figura 16 Propuesta de almacenamiento de fechas en entidad AW

En la especificación de ISAN viene definido que los actores como mínimo los tres que deben existir deben ser actores principales, y en ningún momento se está indicando si es o no principal, por lo que sería eficiente guardar por cada actor dicha información. En la Figura 17 se observa la mejora propuesta para dicha interrelación.

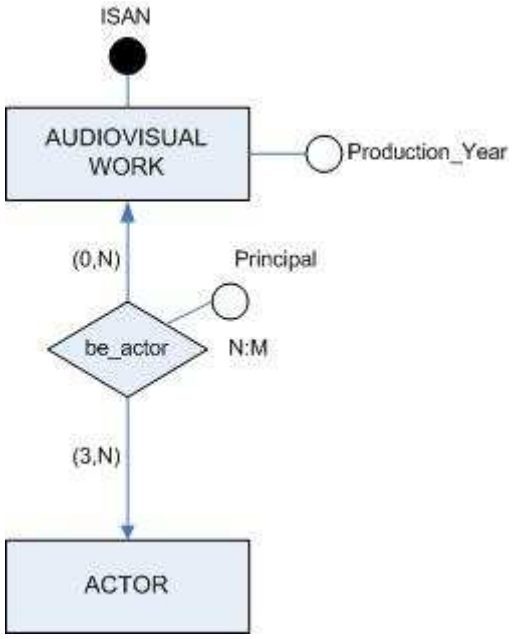

Figura 17 Propuesta de almacenamiento de actores principales

Otra deficiencia que tiene el sistema es el rango de valores que puede admitir un participante que sea de tipo Others, ya que los restringe a 7 tipos diferentes y en cualquier trabajo audiovisual pueden aparecer participantes que tengan un rol totalmente diferente a los 7 que ha definido el estándar ISAN, por ejemplo, podemos encontrarnos ante un participante cuyo rol en una película sea el maquillador, en este

caso, según está especificado ISAN no se podría almacenar información sobre este participante.

ISAN también trata capítulos de series, los cuales llevan asignados un número de episodio si éste es conocido, en otro caso dicho episodio se identifica por medio del título de la serie y del título de dicho episodio; pero sería más correcto que cada serie tuviera numerados todos los episodios que la compongan.

A parte de almacenar el número del episodio de dicho capítulo, deberíamos almacenar la temporada a la que pertenece dicho episodio, por motivos de posible duplicidad de información (dos capítulos con mismo nombre pero de diferente temporada).

La especificación de la información a almacenar de un determinado episodio, no es muy correcta, ya que según viene especificado el título de episodio sería la clave unívoca de dicha entidad, pero también aparece reflejado que dicho dato aparecerá siempre y cuando sea conocido, por lo que se está contradiciendo dicha propiedad en la documentación y no queda nada claro cómo se debe identificar un capítulo. Episode_Number lo indica como dato opcional, pero lo más interesante sería tener como campos obligatorios el número de episodio y la temporada a la que pertenece éste, a parte del título de dicho episodio, así se evitaría duplicidad y confusión en datos almacenados. Ver subesquema E/R en Figura 18.

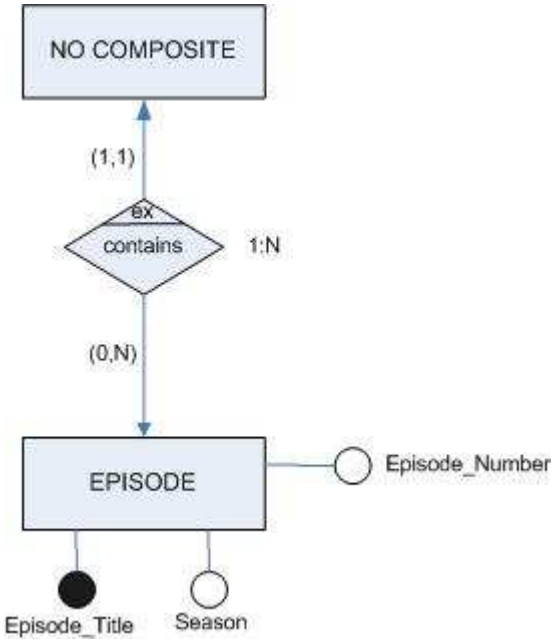

Figura 18 Propuesta de almacenamiento de episodios de una serie

Para la entidad [*COMPANY*] sería muy eficiente realizar su identificación a través de su CIF, ya que es una clave unívoca, al igual que ocurre con las personas y el DNI. Por lo que dicha clave sería la clave primaria y su Nombre sería un simple atributo, además no sería de más almacenar información sobre que tipo de actividad o servicio realiza dicha empresa. Ver Figura 19.

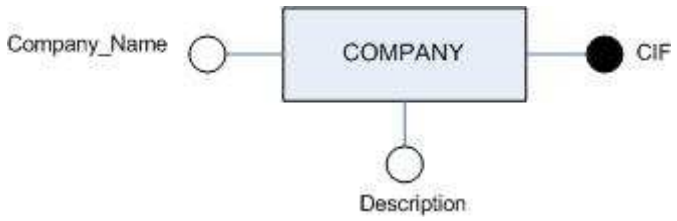

Figura 19 Propuesta de identificación en entidad COMPANY

El campo que guarda si ese trabajo audiovisual tiene o no co-producción no queda muy claro en el documento de ISAN, en concreto el tema de si es o no obligatorio almacenar dicho dato en la base de datos. En los pantallazos de la aplicación aparece como campo obligatorio, en cambio en las especificaciones de los campos aparece como campo opcional; por norma general, la errata estará reflejada en las especificaciones, ya que difícilmente estará el error en la propia aplicación.

Se debería de almacenar información sobre la planificación de retransmisión del trabajo audiovisual. Por ejemplo sería bastante eficiente registrar la duración del programa, la frecuencia con la que se retransmite (si es a diaria, anual, mensual), etc. Para lo que se debería añadir una entidad, denominada por ejemplo [*SCHEDULING*] con los atributos Duration, Frequency y Scheduling Transmisión; y por lo tanto el atributo Duration que tiene la entidad principal [*AUDIOVISUAL WORK*] debería suprimirse de ahí para no tener información duplicada en el sistema.

La documentación de la extensión del ISAN 2002, es decir, V-ISAN 2004 contiene información muy superficial sobre este tema; pero existe un documento denominado Aribsan-ISAN Version 2007 en el cual aparece de manera más detallada toda la información de las versiones de un ISAN.

Además en el documento Aribsan-ISAN 2007 viene definida más información para almacenar en cada versión que en la especificación de V-ISAN; dicha información sería de gran utilidad a la hora de almacenar cada T.A., ya que aportaría un gran número de datos que en principio, no se están definiendo por el estándar

ISAN, entre ellos encontramos el atributo Content_Alert, el cual define si un T.A. tiene un contenido que no es apto para todos los públicos.

Una gran ventaja sería añadir al documento ISAN 2002 todos los conceptos que en los documentos V-ISAN y Aribsan-ISAN se tratan, ya que la mayoría de dichos conceptos e información no se encuentran descritos en el documento principal; por lo que el documento quedaría totalmente actualizado y así se facilitaría el uso del sistema de dicho estándar.

En conclusión, este estándar tiene definiciones que son muy eficientes, el tratamiento de versiones, la diferencia de participantes, el tratamiento de diferentes tipos de trabajo (compuesto, no compuesto, serie, etc.) el único error, es que da la sensación tras leer toda la documentación, visitar su página web, etc. de que están un poco en el aire todas esas definiciones y especificaciones, no se encuentran maduras, se encuentran "cogidas por pinzas"; pero si se madurasen todas las ideas que se ven reflejadas sería un sistema de almacenamiento de fondos audiovisuales muy eficaz, sencillo y exacto para los tres tipos de usuario que trabajan con dicho sistema.

5. TV-ANYTIME

5.1. Introducción

TV-AnyTime es un estándar, creado por TV-AnyTime Forum, que describe la estructura de categorías de programas televisivos y de radio.

TV-AnyTime Forum es una asociación de organizaciones cuyo objetivo es desarrollar especificaciones que permitan el almacenamiento digital de servicios audiovisuales y de otro tipo, en plataformas de usuario, basándose en el mercado de masas. Esta asociación se formó en Newport Beach, California, USA, del 27 al 29 de Septiembre de 1999. Los fines más importantes marcados por este foro son:

- La definición de especificaciones que permitan a las aplicaciones explotar el almacenamiento local en plataformas electrónicas de usuario.
- El desarrollo de especificaciones para sistemas compatibles e integrados, desde los creadores/proveedores de los contenidos hasta los propios consumidores, pasando también por los proveedores de servicios.
- No intervenir en la distribución de contenidos a los equipos electrónicos de consumidor.

- La especificación de las estructuras de seguridad necesarias para proteger los intereses de todas las partes involucradas.

Hay que destacar que existen cuatro grupos de trabajo, que actúan sobre el sistema TV-AnyTime:

- Modelos de negocio, este grupo determina qué requerimientos se deben habilitar en cada fase de las especificaciones.
- Sistema, interfaces de transporte y referenciación de contenido, es el grupo responsable de toda la arquitectura y consistencia del Sistema
- Metadatos, este grupo se encarga de desarrollar la especificación de metadatos.
- Gestión de derechos y protección, se encarga de desarrollar estándares que permitan la expresión segura y la garantía de los derechos de la información distribuida hasta los PVR (Grabador de Video Personal-Dispositivo interactivo de grabación de televisión en formato digital).

Se especificó en dos partes, la primera en agosto del año 2003 (TS 102 822-1/2/4/5/7/8/9) y la segunda fue publicada por ETSI dos años después, en el 2005. Durante el 26 y 27 de Julio de ese mismo año, se celebró en Génova una reunión para ver el éxito que había causado la publicación de las especificaciones de dicho estándar.

El estándar de TV-AnyTime, a día de hoy, es empleado por más de 60 miembros, pertenecientes a un gran número de industrias, entre las cuales destacan emisoras de televisión, emisores por Internet, proveedores de servicios, fabricantes de equipos profesionales, de consumibles electrónicos y de componentes, vendedores de software, propietarios de contenidos. Este estándar básicamente se utiliza en Europa, Asia y EE.UU. además se emplea en grandes empresas como Nokia, Philips, Motorola, Microsoft, Walt Disney Televisión, etc.

La especificación de TV-AnyTime sobre la referenciación de contenidos, aporta una referencia independiente que identifica unívocamente el contenido deseado ubicándolo tanto en tiempo como en localización. Dicha referencia es el llamado C.R.IDentificador.

Las especificaciones del estándar TV-Anytime se basan en el gran área de Metadata. Emplea dicho término para relacionar información sobre contenido, tanto como el título del programa como la sinopsis de éste. A lo largo de las

especificaciones de este estándar trata el término metadata "attractors" a dicha información porque realmente atraen al consumidor el contenido.

El modelo definido por el estándar TV-AnyTime representa programas de los siguientes tipos:

- Un programa simple.
- Un programa con diferente número de versiones, por ejemplo una película con comentarios de los propios actores.
- Un programa que ha sido dividido en partes para ser publicado. Por ejemplo una película de larga duración, que se va a retransmitir en dos partes.
- Un programa formado por partes de otros programas, por ejemplo, un programa de tomas falsas, en el que aparecen partes de diferentes programas.
- Una colección de series y programas individuales que tienen en común el mismo concepto de programa, por ejemplo un *show*.
- Una publicación de un programa cuyo fin era otro, por ejemplo, una película de Lola Flores, su fin realmente es el de entretenimiento, pero tras su fallecimiento el retransmitir una película de ella tiene el fin de recordarla, hacerle un tributo a esta persona.
- Una serie de programas que pueden ser ligados.

5.2. Supuestos semánticos complementarios a la especificación de TV-AnyTime

Este apartado explica los conceptos que no han quedado claros, que son ambiguos o contradictorios en la especificación de TV-Anytime que estamos tratando.

- En caso de que el programa se elimine de la base de datos, toda la información relativa de sus "program location" serán también eliminadas de la base de datos.
- Un "program location" puede ser únicamente "On demand (bajo demanda)" o "Schedule (programado)".
- Un programa puede estar o nominado a uno o varios premios.
- Un premio siempre estará asociado a uno o varios programas diferentes.
- Un programa puede tener diferentes títulos.
- Un programa puede tener diferentes idiomas.
- Un programa puede o no tener diferentes episodios.
- Un programa puede o no tener almacenada información sobre:
    - Derivation Reason (listado por defecto, de criterios para la derivación de la versión de un programa desde un concepto de programa).
    - Derivated From (indica que el programa ha sido derivado de un programa o de un concepto de programa).
    - Base Derivation Reason (define el criterio de la derivación del programa).
- Un título pertenece a un único programa.
- Un episodio pertenece a un único programa.
- Un servicio puede estar formado por diferentes servicios a la vez.
- Un servicio puede o no tener almacenada información sobre la URL.
- Un servicio pertenece a un único Schedule.
- Un Schedule puede estar formado por diferentes servicios.
- Un programa puede estar formado por varios programas a la vez.
- Un programa puede tener diferentes críticas (review), pero una crítica pertenecerá a un único programa.

- Un credit (ficha técnica, persona) puede colaborar en diferentes análisis, pero una crítica estará realizada por un único credit.
- En un programa mínimo trabajará un credit.
- El personaje (character) y el papel (función que desempeña ese personaje en el programa, role) de un credit que trabaja en un programa, quedará almacenado en un fichero histórico en la base de datos, siempre y cuando ambos datos sean conocidos; ya que no es obligatorio almacenar el personaje y el papel que adopta un credit en un determinado programa.
- Un programa puede ser producido en diferentes países. Siempre se almacena en un fichero histórico la fecha de la producción de ese programa en ese país.
- Un programa es creado en uno o varios países. Si se dispone de la fecha de creación del programa en dicho país, ésta se almacenará en un fichero histórico.
- Un programa es realizado en un o varios países. La fecha de realización, siempre se almacenará en un fichero histórico.
- Un programa es representado en uno o varios países. La fecha de representación se almacenará en un fichero histórico, siempre y cuando, este dato esté disponible.

5.3. Esquema Entidad/Interrelación

En el diagrama de la Figura 20 se presenta un esquema conceptual definiendo los conceptos y relaciones más importantes a tener en cuenta en la base de datos que da lugar al sistema del estándar TV-Anytime. En dicho diagrama se ha plasmado exactamente la información según la define dicho estándar, las críticas y posibles mejoras a TV-Anytime se especifican en la sección 5.6 (Análisis Crítico).

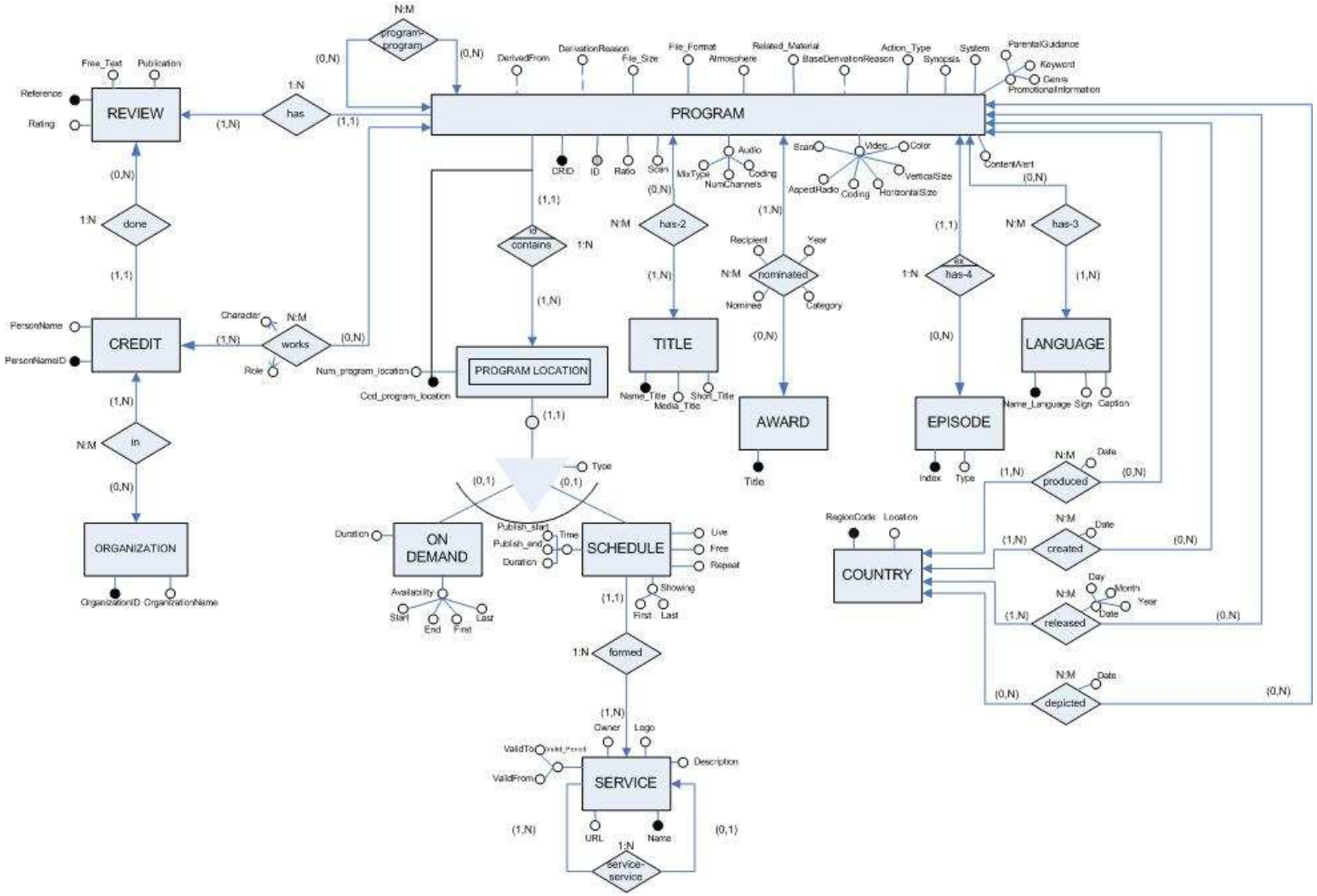

Figura 20 Diagrama E/R estándar TV-AnyTime

5.4. Explicación del Esquema E/R

En este diagrama existe una entidad principal, denominada [*PROGRAM*], en la cual se almacena la información de éste. Esta entidad se encuentra identificada por su clave primaria CRID, la cual identifica a cada programa con un código unívoco; y como clave secundaria se encuentra el código ID.

Dicha entidad tiene los siguientes atributos:

- Related Material: hace referencia a otros recursos de tipo media que se relacionan con el contenido del programa.
- System: describe el medio a través del cual se va a retransmitir el programa.
- Synopsis: es un texto que describe el contenido del programa.
- DerivedFrom: indica si el programa es derivado de otro programa.
- DerivationReason: se trata de un listado de valores que indican si se deriva o no a una nueva versión.
- Audio: siendo un atributo compuesto formado por los atributos Codin (formato en el que está codificado el sonido), NumChannels, MixType
- Video: es un atributo compuesto formado por los atributos AspectRadio, Coding (formato en el que está codificada la imagen), Scan (es el tipo de escaneo del video del programa), Color, Horizontal Size (la anchura de la imagen), Vertical Size (la altura de la imagen del programa).
- Scan: es el tipo de escaneo del color del video del programa.
- BaseDerivationReason: define el criterio de derivación de programas.
- File_Format: describe el formato de la instancia del programa.
- File_Size: indica el tamaño del fichero en el que se encuentra cargado la instancia del programa.
- Ratio: indica una dimensión, un punto en concreto, por ejemplo 3:4, donde 3 representa la dimensión horizontal y 4 representa la dimensión vertical.
- Promotional Information: (texto que contiene la información promocional del programa) es un atributo compuesto que viene formado por los atributos, Genre (género del programa), Keyword (es un listado con todos las palabras clave de un programa), Parental Guidance (es un código de clasificación del programa).

- ContentAlert: es un atributo que indica las alertas de contenido
- Intention: nos indica la intención con la que se ha generado dicho programa.
- Action_Type: define las acciones que pueden ser monitorizadas para analizar las imágenes.
- Atmosphere: es la valoración subjetiva de un programa.

Esta entidad principal [*PROGRAM*] se relaciona con otras entidades necesarias para el sistema, a través de interrelaciones y distintos atributos.

Entre las entidades [*AWARD*] y [*PROGRAM*] se encuentra definida una interrelación del tipo N:M, cuyo objetivo es el de poder tener registrado a qué premios se encuentra denominado dicho programa. En dicha interrelación se almacena el Recipient (especifica la persona o la organización que ha ganado el premio en una categoría dada), Year (indica el año en el cual ganó ese programa el premio o fue nominado a un premio), Nominee (únicamente especifica la persona que ha ganado el premio en una categoría dada) y Category (indica en qué categoría se encuentra nominado el programa) de cada nominación que un programa tenga a un determinado premio [*AWARD*].

Entre las entidades [*TITLE*] y [*PROGRAM*] existe una interrelación de tipo 1:N, en la cual se almacena cada título que tiene un determinado programa; destacando que de cada título se almacenan los siguientes atributos: Name_Title, Media_Title, Short_Title.

Entre las entidades [*LANGUAGE*] y [*PROGRAM*] también existe una interrelación de tipo 1:N pero nos encontramos ante una interrelación en existencia, con el objetivo de almacenar todos los idiomas en los que se ha realizado el programa, guardando las características de Name_language, Sign (especifica el idioma que se ha descrito como el original) y Caption (describe uno de los idiomas que se han empleado en el programa) de cada idioma.

Entre las entidades [*EPISODE*] y [*PROGRAM*] existe una interrelación en existencia de tipo 1:N, de la entidad Episode se almacenan los siguientes campos : Index y Type

Entre las entidades [*CREDIT*] y [*PROGRAM*] hay una interrelación del tipo N:M denominada *works*, la cual relaciona almacenando por cada relación el Character y Role. Además la entidad [*CREDIT*] tiene relación con otras dos entidades, las

cuales son, [*REVIEW*] y [*ORGANIZATION*]. La interrelación *done* relaciona [*CREDIT*] con [*REVIEW*], es de tipo 1:N, y su fin es tener informado en todo momento las personas que han realizado cada crítica. También almacena información sobre las organizaciones a las que pertenece un participante, de ahí la interrelación, *in*, de tipo N:M entre las entidades [*CREDIT*] y [*ORGANIZATION*].

Entre las entidades [*REVIEW*] y [*PROGRAM*] existe una interrelación, denominada *has*, de tipo 1:N cuyo objetivo es almacenar todas las críticas que pertenezcan a un determinado programa.

Entre las entidades [*COUNTRY*] y [*PROGRAM*] existen cuatro interrelaciones diferentes (*produced, created, released, depicted*), siendo todas del tipo N:M y en cada relación almacenando el atributo Date de ésta exceptuando en la interrelación *released* ya que dicha fecha es un atributo compuesto por Day, Month y Year de manera separada.

Existe una interrelación de identificación entre las entidades [*PROGRAM*] y [*PROGRAM_LOCATION*] de tipo 1:N, denominada *contains*, con el fin de tener almacenados todas las retransmisiones de un programa. Hay que destacar que [*PROGRAM_LOCATION*] es una entidad débil, es decir, en el momento que un determinado programa sea eliminado de la base de datos todas sus retransmisiones serán dadas de bajas. En la entidad [*PROGRAM_LOCATION*] se almacena como identificador de dicha entidad Cod_program_location y el Num_program_location.

La entidad [*PROGRAM_LOCATION*] a la vez de ser una entidad débil es el supertipo de una jerarquía parcial sin solapamiento, cuyos subtipos son [*ON_DEMAND*] y [*SCHEDULE*]. Con esta generalización dividimos las retransmisiones de los programas en retransmisiones bajo demanda o planificadas.

Para el subtipo [*ON_DEMAND*] se almacenan los siguientes campos Duration, y el atributo compuesto Availability formado por Start, End, First y Last. Para el otro subtipo, denominado [*SCHEDULE*] se almacenan los siguientes atributos Live, Free, Repeat y los atributos compuestos Time (Public_Start, Publish_End, Duration) y Showing (First y Last). Además el subtipo [*SCHEDULE*] se relaciona con la entidad [*SERVICE*] a través de una interrelación *formed* de tipo 1:N. La entidad [*SERVICE*] se caracteriza por almacenar los siguientes atributos: como identificador de la entidad se opta por Name, luego se informa también Logo (es el logo de la red, por ejemplo una imagen o un icono), Description (es la descripción del servicio), Owner (es la

marca propietaria del servicio) y el atributo compuesto ValidPeriod (indica el periodo de tiempo durante el cual un servicio es válido)formado por ValidFrom (es la fecha y hora a partir de la cual un servicio es válido) y ValidTo (es la fecha y hora hasta la cual un servicio es válido) y como información alternativa se puede informar el campo URL (es un elemento que indica la localización de dicho programa).

Existe una interrelación denominada *service-service* de la [*SERVICE*] con ella misma, ya que un servicio puede estar formado por varios servicios. Es de tipo 1:N.

TV-AnyTime realiza las siguientes definiciones para dichas entidades:

- Program Location contiene información sobre una instancia de un programa.
- Schedule Event es aquel tipo de Program Location que es apropiado para describir las retransmisiones de un programa. Para ello siempre vendrá informado del servicio, la hora y la duración del programa dado.
- Service representa formas, diferentes a las habituales, de retransmitir un programa.

5.5. Supuestos semánticos adicionales al diagrama E/R

En este apartado definimos las propiedades que no pueden ser expresadas en el diagrama E/R generado.

5.5.1. Dominios

En este apartado definimos los valores que pueden admitir los atributos de las entidades definidas. Dichos valores son definidos en el último nivel de detalle, ya que realmente es del que necesitamos la información. La información completa se puede encontrar en [TV-AnyTime, 2000].

- El atributo Action_Type admite valores del siguiente rango

  Dominio Action_Type ={1,2,3,4}

  Este dominio en realidad es más amplio ya que existe herencia de un segundo nivel.

  El primer nivel viene dado por dicho dominio

  1- Audio_Video

  2- Video

  3- Data

4- Commerce

Algunos ejemplos de segundo nivel son los siguientes:

1.5 Pause

2.1 Zoom

3.2 ScrollUp

4.3 AddTolart

- El atributo Related_Material tiene el siguiente dominio de valores:

  Dominio Related_Material ={1...14}

  Algunos ejemplos de este dominio son lo siguientes:

  1 - Trailer

  2 - Group Trailer

- El atributo Character admitirá valores comprendidos en el siguiente rango:

  Dominio Character ={V708,V709, V80, V485,etc.}

  Valores que se corresponden con los siguientes valores :

  V708 - Dubber

  V709 - Key character

  V80 - Choreographer

  V485 - Production Manager

- El atributo Role admitirá valores del siguiente rango:

  Dominio Role ={Anchor, Anchorman, Anchorwoman, Anchorperson, Reporter, Newsman, Newswoman, Newsperson, Interviewer,etc.}

- El atributo Intention únicamente podrá admitir valores comprendidos en el siguiente dominio

  Dominio Intention ={1.1.1, 1.1.2, ..., 1.8.1,1.8.2}

  Por ejemplo,

  1.1.1 - Pure Entertainment

  1.1.2 - Informative Entertainment

  1.8.1 - General Enrichment

  1.8.2 - Inspirational Enrichment

- El atributo Language tiene como dominio los valores definidos por ISO/IEC 639-1.

  Algunos ejemplos son:

Om - Oromo

Ab - Abkhazian

Aa - Afar

Af - Afrikaans

Ie - Interlingue

Ga - Irish

- El atributo Content Alert admite valores comprendidos en el siguiente rango:

  Dominio Content Alert ={6.0.1, 6.1.1, 6.1.5, 6.2,... 6.10.1, 6.10.2, etc.}

  Dichos valores corresponden a las siguientes alertas:

  6.0.1 - No content that requiers alerting in any of the categories below

  6.1.5 - Sexual Violence

  6.10.1 - No strobing

  6.10.2 - Stroging that Could impact on those suffering from photosensitive epilepsy

- El atributo Media admite valores comprendidos en el siguiente rango de valores:

  Dominio Media ={7.1.1, 7.1.4.1, 7.1.4.2, 7.1.4.3,..., 7.2, 7.3}

  Algunos ejemplos de valores de estos tipos de Media son los definidos a continuación:

  7.1.1 - Audio Only

  7.1.4.1 - Text

  7.1.4.2 - Graphics

  7.1.4.3 - Application

- El atributo Atmosphere se informará con valores del siguiente dominio de datos.

  Dominio Atmosphere ={8.1,..., 8.5.3}

  Donde por ejemplo:

  8.1 - Alternative

  8.5.3 – Wacky

- El atributo Scan de la entidad principal [*PROGRAM*] admitirá valores comprendidos entre los siguientes:

Dominio Scan={Interlaced, Progressive}

- El atributo FileFormat de la entidad principal [*PROGRAM*] admitirá valores definidos en la cláusula B.2.11 de ISO/IEC 15938-5 o MPEG-7 IPTCMimeTypeCS.
- El atributo System de la entidad principal [*PROGRAM*] admite valores del listado que define B.2.30 de ISO/IEC 15938-5.
- El atributo DerivationReason de la entidad [PROGRAM] admitirá valores del siguiente rango:

  Dominio DerivationReason = {Violence, Language, Sex, Duration, Other}

- Una retransmisión puede ser bajo demanda o planificada, para lo cual el campo Type de dicha jerarquía admitirá valores del siguiente rango:

Dominio Type (jerarquía de Program_Location) = {On_Demand, Schedule}

5.5.2. Otros

- El atributo Name de la entidad [*ORGANIZATION*] vendrá definido según ISO/IEC 15938-5.
- El atributo Parental Guidance que forma parte del atributo derivado Promotional_Information de la entidad [*PROGRAM*] vendrá definido como en la cláusula 9.2.3 de ISO/IEC 15938-5.

5.6. Análisis Crítico

Este estándar tiene una especificación y una definición de términos muy correcta, abarcando todos los puntos desde características del programa hasta información de toda la gente que trabaja en dicho programa, como de las organizaciones que contribuyen a la realización de éste, así como las características correspondientes a la retransmisión y planificación de dicho programa.

Un punto que diferencia este estándar de Escort e ISAN es el tratamiento de nominaciones de premios (Award), y almacenar las críticas que se realizan a cada programa (Review). Es una característica que hace más consistente al modelo TV-AnyTime, ya que es información importante sobre cada programa. La mejora que se debería realizar en correspondencia a este punto sería que las nominaciones de premios también estuvieran reflejadas a nivel de persona componente del programa,

lo que TV-AnyTime denomina como Credit, ya que un participante puede estar nominado, o haberlo estado, a un determinado premio.

En cambio una inconsistencia de este estándar es que al tratar un programa que sea de tipo serie, no se almacena en ningún momento datos de la propia serie, únicamente se hace referencia si dicho programa es un capítulo pero en realidad no se informa ni del título del episodio, ni la temporada a la que pertenece, ni el número de episodio dentro de éste. Este punto deja bastante desequilibrado al propio esquema TV-AnyTime.

Otra deficiencia del modelo es almacenar la alerta de contenido (Content_Alert) como un único atributo de la entidad [*PROGRAM*]. Un programa puede tener diferentes alertas de contenido y además éstas deberían estar asociadas a los países, ya que cada país tiene definidas unas alertas diferentes por cada tipo de contenido.

En ningún momento se almacena el nombre artístico del participante, únicamente se almacena el nombre y además como identificador de dicha entidad [*CREDIT*]. Por lo que es inconsistente ya que en ningún momento en las especificaciones del estándar TV-AnyTime se indica si el nombre es completo con ambos apellidos, o si sólo se encuentra informado primer apellido y nombre en cuyo caso se generarían confusiones en datos y duplicidad de datos, que realmente no son verdaderos. Nos referimos al caso de tener personas de la misma familia trabajando para dicho programa.

Las cuatro interrelaciones entre [*PROGRAM*] y [*COUNTRY*] se podrían agrupar en una misma, almacenando en dicha interrelación una fecha por cada tipo de actividad (date_produced, date_created, date_released, date_depicted), todas ellas con el formato DD-MM-YYYY. De esta manera el sistema sería más eficiente ya que no tendría tantas relaciones que generar.

Para cada episodio apenas se guarda información, lo correcto sería almacenar para cada uno de éstos características similares a las que un programa posee, es decir, Audio, Content_Alert, Video, etc. Además sería de gran importancia almacenar la temporada a la que pertenece dicho episodio dentro de su serie.

TV-AnyTime define que un programa puede tener varios títulos, por ejemplo por estar traducido ese programa a diferentes idiomas, pero en realidad no genera ningún tipo de relación de título con lenguaje. Por lo que se debería crear una relación

ternaria de tipo 1:N:M entre las entidades de Título, Programa e Idioma para solventar el problema existente.

A pesar de todas estas deficiencias, se puede denominar al modelo de TV-AnyTime uno de los modelos más completos de los estudiados en este proyecto fin de carrera, en cuanto a nivel de información importante a almacenar por cada programa tratado.

6. ESCORT

6.1. Introducción

Escort es un sistema de clasificación de programas tanto de radio como de televisión.

Escort 2.4 se creó para cubrir la necesidad que los productores tenían para acceder a diferentes informaciones de fondos audiovisuales, para poder analizar y realizar comparaciones entre diferentes trabajos. Dicho diseño se generó en el año 2004, año en el cual la mayoría de los productores comenzaron a digitalizar todas las instalaciones y actividades. EBU (*European Broadcasting Union*), ETS (*European Telecommunications Standards Institute*) y SMPTE (*Society of Motion Picture and Televisión Engineers*) se encargan del desarrollo de diferentes especificaciones para las actividades que los productores llevan a cabo, entre las cuales aparece el estándar Escort. Pero ya en el año 2007 se creó una nueva versión de dicho estándar, cuyo objetivo es el intercambio de información para la comparación de ésta en sus estadísticas y su público. Esta última versión también se emplea para temas de finanzas, contabilidad y marketing.

Escort 2007 se puede usar tanto para programas de televisión como para programas de radio, aunque en este proyecto fin de carrera únicamente se estudiará lo relativo a programas de televisión.

Este estándar, realmente se puede definir como un esquema de clasificación multidimensional, cuyas dimensiones permiten la separación de diferentes géneros en diferentes categorías. Identifica un conjunto en común de géneros y conceptos de programas, los cuales forman parte de una descripción de un servicio o de un programa.

Las aplicaciones del tipo EPGs (*Electronic Programme Guides*) emplean términos de géneros, como los definidos en Escort 2007, por lo tanto las dimensiones propuestas en este estándar sirven para generar un género EPG.

6.2. Estructura Escort 2007

Escort se basa en que un programa o servicio es completo siempre y cuando se conozca la siguiente información para cada uno de ellos.

- Descripción conceptual, idioma y participación
- Datos de adquisición de programa
- Datos de planificación
- Datos de retransmisión
- Datos de visualización y de audición
- Datos financieros

Este estándar engloba todos estos datos en un término llamado Broadcast Event, con el cual juega a lo largo de todas sus especificaciones.

Escort define 7 dimensiones diferentes para la parte de descripción conceptual, de idioma y participación. Identifica cada dimensión con un código unívoco.

Dichas dimensiones son las siguientes:

- Intention
- Format
- Content
- Participation
- Intended Audience / Target Group
- Origination
- Content Alert

Como podemos ver en las Figura 21 este estándar emplea una clasificación con herencia para la identificación de las dimensiones. Es decir, el identificador que tiene cada dimensión parte primeramente de un nivel global, nivel en el cual se va profundizando hasta llegar al nivel más bajo, cuyo fin es tener el mínimo detalle de ese atributo.

Required level	Recommend Level	Individual Level
Required by EBU	Recommended by EBU	For individual keywords and information

Figura 21 Formato de clasificación con herencia (géneros y subgéneros)

En la Figura 22 se muestra un ejemplo de como se realiza esa navegación entre niveles.

Required level	Recommend Level	Individual Level
NON FICTION / INFORMATION	News and current affaire	Daily News

Figura 22 Ejemplo de clasificación con herencia de la dimensión CONTENT

El identificador de cada dimensión estará formado por un número de dígitos, donde cada dígito identificará el tipo, subtipo y subtipo del subtipo, así que dependiendo del nivel de información al que se baje, su identificador vendrá formado por un mayor o menor número de dígitos.

Existen tres códigos, fijados por el estándar, para identificar las entidades, siempre y cuando nos encontremos bajo alguno de los siguientes casos excepcionales:

- En caso de que la especificación no se ajuste a los datos de la dimensión Program, se le aplicará el código 97 (*Other*).
- En caso de que no haya conceptos muy convincentes sobre el programa, se le aplicará el código 98 (*Mixed*).
- Si no se conocen unos determinados datos, se le aplicará el código 99 (*Unknow*).

6.3. Supuestos semánticos complementarios a la especificación de Escort

Este apartado explica los conceptos que no han quedado claros, que son ambiguos o contradictorios en la especificación de Escort que estamos tratando.

- Un programa tiene un único origen. Un origen puede pertenecer a diferentes programas.
- Un programa tiene una única intención, pero una intención puede pertenecer a varios programas.
- Un programa tiene un único formato y un formato puede pertenecer a varios programas.
- Un contenido puede pertenecer a varios programas pero un programa únicamente puede tener un contenido.
- Un programa tiene un único grupo de audiencia pero éste puede pertenecer a diferentes programas.
- Un idioma puede aparecer en diferentes programas. Y un programa solamente puede tener un idioma definido
- Se debe almacenar la descripción del origen, intención, formato, contenido, participación, idioma y grupo de audiencia del programa.
- Un idioma vendrá identificado por los tres códigos de idioma como original, doblado y subtitulado.
- De cada país participante en el programa se debe almacenar los dos códigos que el estándar ISO Internacional Codes For Names PF Countries and Curriences define por país. Además también se almacenará la moneda del país.

6.4. Esquema Entidad/Interrelación

En el diagrama de la Figura 23 se presenta un esquema conceptual definiendo los conceptos y relaciones más importantes a tener en cuenta en la base de datos que da lugar al sistema del estándar Escort. En dicho diagrama se ha plasmado exactamente la información según la define dicho estándar, las críticas y posibles mejoras a Escort se especifican en la sección 6.7 (Análisis Crítico).

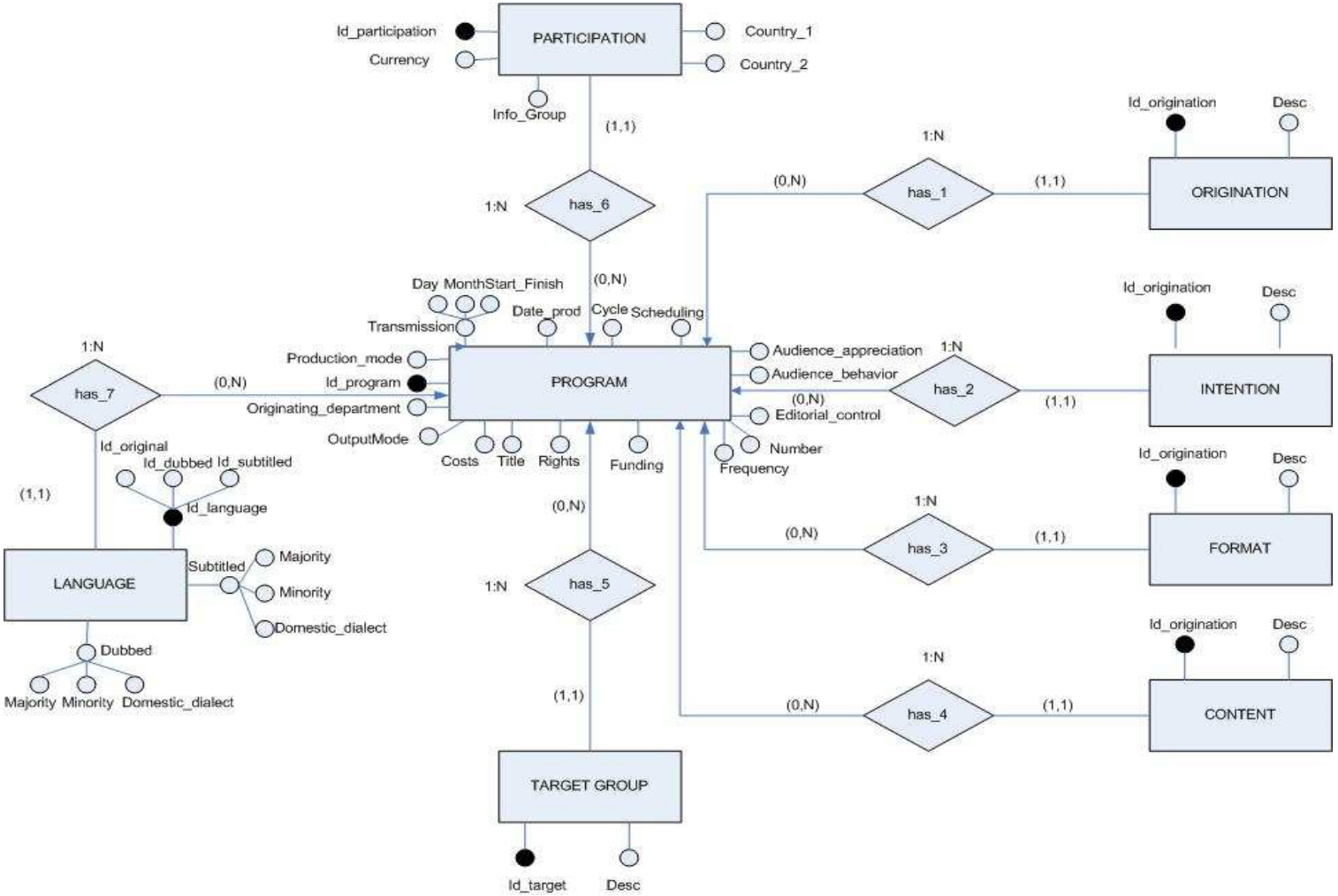

Figura 23 Diagrama E/R estándar Escort

6.5. Explicación del Esquema E/R

En este diagrama existe una entidad principal, denominada [*PROGRAM*], en la cual se almacenan datos sobre dicho programa. Esta entidad se relaciona con otras entidades necesarias para el sistema, a través de relaciones y diversos atributos.

Los atributos de la entidad [*PROGRAM*] son los siguientes:

- Rights, campo que almacena toda la información relativa sobre los derechos del programa, desde costes hasta derechos de propietarios, etc.
- Originating_department, es el departamento interno responsable de la producción o adquisición del programa (basado en la estructura interna de la organización de retransmisión)
- Title es el nombre que identifica al propio programa
- Number indica el número de programa .
- Scheduling almacena si el programa fue retransmitido por primera vez, o si se trata de una repetición (por primera vez, por segunda vez, etc.)
- Frequency indica la periodicidad con la que es habitual la retransmisión de este programa.
- Production_mode define todas aquellas instalaciones principales que intervienen en la producción del programa.
- Funding son los orígenes de finanzas a través de los cuales el retransmisor produce o adquiere el programa.
- Costs, son todos aquellos costes involucrados en la producción de un programa, incluyendo las inversiones directas del programa.
- OutputMode nos indica si el programa es en vivo o en partes sin un periodo de tiempo específico o desde un material de stock.
- Duration, es la duración neta del programa en minutos
- Cycle, indica si el programa forma parte de una serie o no.
- Editorial Control, es el área para la cual el retransmisor tiene el control del contenido del programa en la etapa de producción.
- Audience_behaviour, trata la información relativa a los niveles de de escucha y de vista.

- Audience_appreciation se basa en el sistema de medida interno que tenga la organización, con el objetivo de medir la audiencia que dicho programa tiene.
- Transmisión, es un atributo compuesto por los campos Day, Month y Start-Finish, cuyo fin es almacenar la fecha en la cual se retransmite dicho programa.

Además tiene como clave principal Id_program, la cual identifica de manera unívoca a un programa.

Se ha definido una interrelación de tipo 1:N entre la entidad [*ORIGINATION*] y la entidad principal [*PROGRAM*] denominada *has-1*, con el objetivo de almacenar el origen del programa, destacando que un único programa admitirá únicamente un origen, con la posibilidad de que dicho origen pertenezca a diferentes programas. El estándar define a la entidad [*ORIGINATION*] como aquella que almacena toda la información relativa al sector que crea el principal contenido del programa la primera vez. De esta entidad se almacena tanto el identificador como la descripción de ese origen del programa.

La entidad [*INTENTION*] se relaciona con la entidad principal [*PROGRAM*] a través de la interrelación *has-2* de tipo 1:N, con el fin de tener registrado el propósito que tiene dicho programa puntualizando que un programa tendrá asignada una única intención. El estándar define la entidad [INTENTION] como aquella entidad que almacena el principal objetivo que el retransmisor tiene a la hora de transmitir dicho programa; también se almacena el identificador y su descripción.

La entidad [*FORMAT*] se relaciona con la entidad principal [*PROGRAM*] con la interrelación *has-3* de tipo 1:N para almacenar el formato que tiene un programa (éste solamente admitirá exclusivamente un formato). Por cada formato se almacena su identificador y su descripción. El fin de esta entidad es clasificar los programas por su estructura formal

La entidad [*CONTENT*] se relaciona con la entidad principal [*PROGRAM*] a través de una interrelación de tipo 1:N llamada *has-4*. El objetivo de dicha interrelación es almacenar la información relativa al contenido del programa. Hay que destacar que un programa únicamente admite un contenido y que de dicho contenido

se almacena un identificador y una descripción. El fin de [*CONTENT*] es clasificar los programas por su contenido o su objetivo.

Se ha definido, una nueva interrelación de tipo 1:N, entre la entidad principal [*PROGRAM*] y la entidad [*TARGET GROUP*], dicha interrelación se denomina *has-5*. El objetivo es similar a las anteriores interrelaciones mencionadas pero en está se almacena información relativa al grupo de audiencia, el cuál será único para el programa. Esta entidad [*TARGET GROUP*] sirve para indicar la audiencia para la cual se ha creado dicho programa, clasificándola por edad, cultura, profesión, etc. Para ello se almacena el identificador y la descripción de cada grupo de audiencia que esté relacionado con el programa.

Existe una entidad denominada [*PARTICIPATION*], cuyos atributos son su identificador, los dos códigos que se almacenan del país que realiza dicha participación y la moneda de dicho país, a parte de su descripción correspondiente. Su finalidad es clasificar los programas por la información de los representantes de ciertos grupos relevantes como pueden ser partidos políticos, regiones, etc. Entre las entidades [*PARTICIPATION*] y [*PROGRAM*] se encuentra definida una interrelación del tipo 1:N llamada *has-6*, cuyo fin es poder tener registrada la participación que ha habido en el desarrollo de dicho programa audiovisual. Únicamente se acepta una participación por programa.

Por último, se ha definido una interrelación entre la entidad [*PROGRAM*] y la entidad [*LANGUAGE*] denominada *has-7* y de tipo 1:N, siendo su objetivo el mantener registrado el idioma en el que está definido dicho programa. Un programa puede estar doblado o subtitulado, por lo que dicha información se almacena en caso de que exista la posibilidad de que dicho programa sea de algún tipo de éstos; de ahí que la entidad [*LANGUAGE*] se encuentre definida con tres atributos compuestos, uno de ellos es para la parte del doblaje, otro para el subtitulado y por último, la clave principal que se encuentra formada por tres diferentes identificadores a la vez (Id_original, Id_dubbed e Id_subtitled), es decir, parte de la clave la compone el identificador del idioma original, otra parte el identificador del idioma en el que se ha doblado y finalmente otro identificador por el idioma en el que se ha subtitulado. Los atributos que forman los atributos compuestos Dubbed y Subtitled son los siguientes:

- Majority indica si el idioma es o no mayoritario entre los nativos.

- Minority indica si el idioma es o no minoritario entre los nativos.
- Domestic-dialect indica si ese idioma es o no el dialecto principal.

6.6. Supuestos semánticos adicionales al diagrama E/R

En este apartado definimos las propiedades que no pueden ser expresadas en el diagrama E/R generado.

6.6.1. Dominios

En este apartado definimos los valores que pueden admitir los atributos de las entidades definidas. Dichos valores son definidos en el último nivel de detalle, ya que realmente es del que necesitamos la información.

- El atributo Frequency admite valores del siguiente rango:

  Dominio Frequency={One off, Hourly, Daily, Every 2-6 days, Weekly, Monthly, Seasonnaly, Annually, Regularly}

  Por ejemplo, el programa de fin de año (campanadas) tendría como frequency Annually, ya que se retransmite una vez al año.
- El atributo Cycle tiene el siguiente dominio:

  Dominio Cycle={Single production, Anthology, Series, Mini-serial, Regular serial, Long-running serial}

  Escort ha definido los valores de ese dominio como:
  - Single production: es aquel programa que no se retransmite en más de dos partes.
  - Anthology: es aquel programa que comparte solamente el título principal y el concepto básico con otros programas.
  - Series: es aquel programa que comparte formato y título con otros programas.
  - Mini-Serial: es aquel programa que es una serie compuesta por no más de 6 episodios; también se les denomina como “mini series”.
  - Regular Serial: es aquel programa que tiene un número de episodios comprendido entre 7 y 26.
  - Long-running Serial: es aquel programa que tiene más de 26 episodios.

Por ejemplo Médico de Familia, Los Serrano,etc. podríamos tratarlos como programas con atributo Cycle Long-running, ya que son un conjunto de capítulos que relatan una historia continuada y formado por más de 26 capítulos en una misma temporada.

- El atributo Scheduling admite valores del siguiente dominio:

  Dominio Scheduling={On other channel, On two other channels, First repeat, Second repeat}

  Por ejemplo, uno de los programas que se realizan para Nochebuena; la retransmisión de éste en la noche del día 24 de Diciembre tendría como Scheduling First Repeat, en cambio al día siguiente suelen emitir una repetición de dicho programa, por lo que ese programa tendría como Scheduling Second Repeat.

- El atributo Desc de la entidad [*ORIGINATION*] admitirá valores comprendidos en el siguiente rango:

  Dominio Desc(Origination)={TV Studio Originated, Concert Hall/Theatre, Cinema Industry Orig., Videogramme, Public Event}

  Este estándar define los valores de este dominio como:

  - TV Studio Originated: aquel programa que su fin es retransmitirlo por televisión.
  - Concert Hall/Theater: es aquel programa basado en un evento musical o mostrado en un teatro o concierto,
  - Cinema Industry Orig.: es un programa que se ha realizado con el fin de publicarse en el mundo cinematográfico.
  - Videogramme: es aquel programa que únicamente está formado por videoclips.
  - Public Event: es un programa que recubre un evento, el cual va a tener lugar también en medios distintos a la radio y televisión.

    Por ejemplo, un meeting de un partido político tendría como descripción de Origination Public Event.

- El atributo Desc de la entidad [*INTENTION*] admitirá valores comprendidos en:

Dominio Desc (Intention)={Pure entertainment, Informative Entertainment, Pure Information, Infotainment, Advice, General Enrichment, Education, Inspirational Enrichment, Involvement}.

Escort define los valores de este dominio como:

- Pure Entertainment: es aquel programa que transmite relajación, sentimientos de belleza…
- Informative Entertainment: es aquel programa cuyo fin es el entretenimiento pero con connotaciones informativas.
- Pure Information: es aquel programa cuyo fin principal es el informar sobre hechos recientes, eventos, teorías además de consejos.
- Infotainment: es aquel programa que informa pero con elementos de entretenimiento.
- Advice: es aquel programa que da consejos sobre temas personales, económicos, de salud, salud, tráfico, etc.
- General Enrichment: es aquel programa cuyo fin es incrementar el conocimiento sobre hechos a través de una forma didáctica o inspiración espiritual.
- Education: es aquel programa que tiene como fin el aumentar el conocimiento sobre hechos de una manera didáctica.
- Inspirational Enrichment: es aquel programa que se basa en los diferentes tipos de creencias religiosas.
- Involvement: es aquel programa cuyo fin es el integrar a la gente en acciones sociales como por ejemplo en temas de "buena causa", etc.

Por ejemplo, un curso de inglés retransmitido por televisión tendría como Intention Education, ya que su objetivo es la enseñanza, enseñanza de un idioma en este caso.

- El atributo Desc de la entidad [*FORMAT*] admitirá valores comprendidos en el siguiente dominio:

  Dominio Desc (Format)={Bulletin, Magazine, Commented event, Documentary, Discussion/Interview/Debate, Lectura/Speech/Presentation, Textual, Performed drama, Drama with puppets, Animated drama, Hosted drama, Panel-show, Non-hosted show, Standup commediam(s), Solo

performance, Small ensemble performance, Orchestral, Choral music, Theatrical music, Ballet/Dance, Mixed music}

Algunas de las definiciones que el estándar Escort realiza a los posibles valores de esta entidad son:

- Discussion/Interview/Debate es un programa principalmente verbal en el cual participa más de una persona.
- Lectura/Speech/Presentation es un programa principalmente verbal en el cual participa una única persona.
- Drama with puppets es un programa de carácter dramático realizado con marionetas.
- Animated drama es un programa de carácter dramático realizado a través de imágenes, exactamente dibujos.
- Orchestral es un programa musical en el cual intervienen más de 7 personas.
- Ballet/Dance es un programa realizado por bailarines.

Por ejemplo, el discurso del Rey la noche del 24 de Diciembre sería de tipo Lectura/Speech/Presentation ya que es un discurso realizado por una única persona, los conciertos de música clásica tendrían como descripción en la entidad [*FORMAT*] Orchestral,etc.

- El atributo Desc de la entidad [*CONTENT*] admitirá valores del siguiente rango:

Dominio Desc (Content)={Daily news, Current affaire, Weather Forescast, Buddishm, Hinduism, Christianity, Islam, Judaism, Communism, Humanism, Political, Social, Economical, Legal, Reality, Show Business, Notable personalities, Personal problems, Do-it-yourself, Cookery, Gardening, Tourist/Travel, Motoring, Keep-fit, Fashion, Computing, Music, Plastic Arts, Theatre/Cinema, Advertising, Press, Literature, Languages, History, Applied sciences, Natural sciences, Physical sciences, Medicine, General light drama, Comedy, "Broken" comedy, Sitcom, Soap, Medical melodrama, Adventure/disaster, Detective/thriller, Epic, Horror, Science fiction, War, Western, Fantasy/Fairy tale, Erotica, Classical drama, Contemporary drama,

Docudrama, Poems/Stories, Simple game show, Big game show, Quiz/Contest, Variety Show, Surprise show, Humour, Candid camera, Early, Classical and Romantic, Contemporary, Light classical, Jazz, Middle-of-the-road, Pop/Rock/Pop-Rock, Traditional}.

*Además se incluye un listado de deportes, el cual lo detallamos más adelante en este documento.

Algunas de las definiciones que el estándar Escort realiza a los posibles valores de esta entidad son:

- Medicine es aquel programa que trata temas médicos, salud, etc.
- Natural Sciences es aquel programa que trata temas sobre biología, botánica, geología y zoología.
- Early Music es aquel programa que trata música escrita antes de la mitad del siglo 17.
- Contemporary Music es un programa que trata música escrita durante el siglo 20.
- Classical Drama es aquel programa dramático que ha sido escrito antes del año 1918.
- Serious Drama es aquel programa dramático con valor cultural y literario.

  Por ejemplo, la película de Romeo y Julieta tendría contenido Classical and Romantic, la retransmisión de la Pasarela Cibeles tendría como Content Fashion, la telenovela Pasión de Gavilanes tendría como atributo Content Soap.

- El atributo Desc de la entidad [*Target Group*] solamente podrá admitir valores comprendidos en el siguiente dominio:

  Dominio Desc(Target Group)={Young Children, Children, Teenagers, Young Adults, Senior Citizens, Regional, Local, Ethnic Groups, Inmigrant Groups, Linguistic Groups, Mainshoppers, Business People, Farmers, Fishermen /Sailors, Schoolchildren/Students, Unemployed People, Motorists, Tourists, Women}.

  Por ejemplo, los capítulos de Barrio Sésamo tendrían como Target Group Young Children, la serie Salvados por la Campana admitiría el valor de Teenagers ya que es una serie destinada a adolescentes.

- Un programa de contenido deportivo contendrá como descripción cualquier valor de los que vienen definidos en el listado GA 1SF/A GFIS (*General Association of Internacional Sports Federations*)*.

  Algunos ejemplos de estos contenidos deportivos son: Aeronautics, Golf, Judo, Skiing, American Football, basketball, Surfing, etc.
- La clave principal de la entidad [*LANGUAGE*] se encuentra formada por tres campos, los cuales admitirán valores según el idioma que le corresponda. Algunas de las asignaciones de idiomas que se muestran en el documento del estándar, son las siguientes:
  - 1 English
  - 17 Serbian
  - 21 Kurdish
  - 28 Slovak
  - 30 Hebrew
  - 43 Riff
  - 44 Romany

  Por ejemplo, una película rodada en Inglés tendría como Id_language 1 y como descripción English.
- Los atributos que componen los atributos Dubbed y Subtitled, tomarán valores comprendidos en los siguientes dominios.

  Dominio Majority = {yes, no, n.a.}

  Dominio Minority = {yes, no, n.a.}

  Dominio Domestic_dialect = {yes, no, n.a.}

  Por ejemplo, un programa con idioma principal el Inglés y doblado en Portugués y sin subtítulos tendría el siguiente conjunto de atributos:
  - Id_original con valor 1.
  - Id_dubbed con valor 5.
  - Id_subtitled con valor 0.
  - Majority (Dubbed) con valor yes.
  - Minority (Dubbed) con valor no.
  - Domestic_Dialect (Dubbed) con valor no.
  - Majority (Subtitled) con valor n.a.

- Minority (Subtitled)) con valor n.a.
- Domestic_Dialect (Subtitled)) con valor n.a.

- Los atributos Country_1, Country_2 y Currency de la entidad [*PARTICIPATION*] tomarán valores definidos por la ISO Internacional Codes For Names PF Countries and Curriences. Dicha definición relaciona a cada país con sus dos códigos de país y su identificador de moneda.

Por ejemplo, la película Torrente se ha realizado en España por lo que los campos Country_1, Country_2 y Currience tendrían los siguientes valores: ES, ESP y ESP/ESB.

6.6.2. Otros

- El atributo Duration se almacenará en minutos únicamente.
- El atributo Date_Prod tendrá el formato YYMMDD.

6.7. Análisis Crítico

Escort es un sistema de clasificación que tiene como fin satisfacer la necesidad de identificación de programas.

Se podrían realizar algunas mejoras en este sistema ya que de la manera que se encuentra definido es bastante ineficaz e ineficiente.

Esa ineficacia e ineficiencia surge porque el estándar Escort únicamente se basa en las características del programa, es decir, trata con delicadeza los contenidos, la audiencia, los formatos, etc. pero por ejemplo información sobre la gente participante en el programa, sobre posibles versiones, etc. no se contemplan en este estándar, y hacen que el estándar sea un poco escaso en cuanto almacenamiento de información se refiere

.Algunas de las inconsistencias existentes son las siguientes:

- En la entidad [*PARTICIPATION*] se está almacenando información relativa a los representantes de los grupos más relevantes de dicho trabajo; en principio, el nombre de la entidad es contradictorio, porque a simple vista el usuario cree que es información sobre cualquier persona física que ha trabajado en el desarrollo del programa. Además, hay que destacar, que en ningún momento se está almacenando información sobre dicho personal, es decir, no se sabe quien es el director, el productor, los actores principales, el estilista, el co-

productor, etc.Sobre esta entidad también destacar que no tiene ningún sentido almacenar en ella la información relativa a los países que intervienen en el desarrollo del programa. Lo más correcto sería incluir todos los atributos que hagan referencia al país, en la entidad [*ORIGINATION*]. Si únicamente metemos esa información en dicha entidad, no tendría sentido almacenar ni el segundo identificador del país (con tener un único código de país sería suficiente), ni la moneda de éste. Pero también cabe la posibilidad de crear una nueva dimensión, llamada por ejemplo [*COUNTRY*] y ahí almacenar todos los códigos del país como hace en Internacional Codes for names PF Countries and curriencies (ISO) junto con su moneda, y además ya se podría almacenar también el nombre completo del país.

- Escort define Programme Rights como una entidad, pero es imposible que se trate como una entidad ya que no se define ni la información, ni los campos que son necesarios a la hora de almacenarlos. Realmente se está tratando Rights como un atributo de la entidad principal [*PROGRAM*]. Por lo tanto, se deberían definir los campos de dicha entidad para que el sistema definido en los requisitos concuerde con lo que realmente se ha desarrollado.
- Una gran inconsistencia de este sistema es a la hora de tratar un programa que sea una serie. No se almacena en ningún momento el capítulo que es, la temporada a la que pertenece. Es decir, el sistema debería estar definido para que en caso de que se trate una serie, se pueda llevar una contabilización de los capítulos, de las temporadas que hay y así tener mayor información almacenada de dicho programa.
- Otra desventaja es que por lo que viene definido en la documentación, si un programa tiene dos grupos diferentes de posible audiencia, se almacenará dos veces ese programa pero cada uno con un diferente Id_target; lo ideal sería poder almacenar versiones de los programas.

A parte de tener un gran número de desventajas y/o inconsistencias, el estándar Escort tiene algunas ventajas como las siguientes:

- La manera en la que clasifica los programas, a través de diferentes características, como la intención, el formato,…llegando al nivel más bajo es un punto a favor para este estándar ya que a través de un mínimo código por

entidad sabemos de manera rápida que tipo de formato tiene ese programa, su intención,…

- Además, Escort al fin y al cabo es un sistema de clasificación para todo tipo de programas, ya que en caso de que dicho programa no posea alguna de las características definidas por el sistema, Escort lo almacena también con un identificador fijo (dependerá de si es otro tipo (97), si el identificador tiene que depender de varias características (98) o si no es conocida dicha cualidad (99)).

La documentación de este estándar es bastante intuitiva ya que es clara y concisa, con una gran colección de anexos en la cual se detallan todos los términos y dominios que se aplican en el sistema.

7. PROPUESTA DE NUEVO META-MODELO

7.1. Introducción

El esquema de almacenamiento de datos creado en este proyecto fin de carrera, es un sistema de clasificación y almacenamiento de materiales audiovisuales que cumplen las siguientes características:.

- El sistema considera un programa (o material audiovisual) todo aquel trabajo audiovisual (siempre tratándose de programas de retransmisión televisiva, no definido para retransmisiones en radio) que sea del siguiente tipo:
- Programa básico, por ejemplo el telediario.
- Programa formado por diferentes programas, como por ejemplo un programa de tomas falsas en el que salen imágenes de diferentes programas.
- Una serie, que es aquel programa que se encuentra formado por un número de capítulos (programas básicos) que tienen un título en común y tratan una historia continuada.

Para la creación del esquema conceptual de la base de datos se han tomado como base tres estándares actuales de clasificación de programas audiovisuales estudiados durante la realización del proyecto fin de carrera: ISAN, TV-AnyTime y Escort. De cada uno de ellos, se han elegido los conceptos e ideas que se han creído

más relevantes en cuanto a materiales audiovisuales de difusión vía televisión, de forma que en conjunto se genere un sistema eficaz y eficiente de clasificación y numeración de programas audiovisuales.

De una manera muy resumida, un desglose de las ideas escogidas de cada uno de los estándares estudiados para el diseño del nuevo sistema es el siguiente:

- *ISAN*. De este estándar se ha escogido la clasificación de participantes según su tipo, agrupándolos en los tres grupos principales (director, actor y otros), la manera de almacenar una nueva versión de un programa identificando ésta por el código del propio programa junto con el código que identifique a dicha versión.
- *TV-AnyTime*. De este estándar se ha escogido el almacenamiento de los premios a los que optan los programas como nominados, las diferentes relaciones de los países con los programas diferenciándolos por tipo de actividad realizada en éste, la interrelación de un mismo programa con otro programa, etc.
- *Escort*. De este estándar se ha optado por reutilizar ideas como el idioma empleado si es para subtitulado del programa o para el doblaje, o para ambas, los medios de transmisión, los atributos sobre la audiencia, etc.

Aunque los conceptos e ideas principales de estos estándares se han tenido en cuenta para la creación del nuevo esquema, el diseño ha podido ser modificado para mejorar la eficacia y eficiencia del almacenamiento de los datos y, al mismo tiempo, se han añadido ideas personales, que en cierta medida, se intuye que hagan más fuerte al propio esquema, como es la creación de la entidad que trata Capítulos, la posibilidad de tener almacenado quien es el ganador de cada premio, etc.

Todas estas similitudes a los modelos ya vistos anteriormente en este proyecto fin de carrera y los aportes personales para desarrollar un nuevo modelo se encuentran especificados claramente en la sección XXXX de este documento.

7.2. Supuestos semánticos complementarios a la especificación de la Propuesta

Este apartado explica los conceptos que no han quedado claros, que son ambigüos o contradictorios en la especificación de la propuesta que se está realizando.

- Un programa tendrá como mínimo un director y al mismo tiempo un determinado director podrá serlo en diferentes programas.
- Un programa tendrá como mínimo un director. Pero un determinado director podrá serlo en diferentes programas.
- Un actor podrá pertenecer a varios programas pero en un programa mínimo habrá un actor.
- Se destacará si un actor participante en un programa es principal o no en dicho programa.
- Un participante de otro tipo, que no sea ni director ni actor, podrá trabajar en diferentes programas, aunque en un programa mínimo habrá un participante de este tipo.
- Se debe identificar en cada actor o en cada participante (exceptuando al director) el papel que desempeña éste en el programa dado.
- Un participante se identificará por una clave y como clave secundaria su nombre artístico.
- Un premio vendrá identificado por su nombre.
- Un programa puede estar nominado a varios premios. Al igual que un premio puede tener como nominados a diferentes programas.
- Un participante puede estar nominado a más de un premio; además a un premio pueden estar nominados varios participantes.
- Una compañía vendrá identificada por su CIF como clave unívoca.
- Una misma compañía puede trabajar para diferentes programas. Además en un mismo programa pueden trabajar varias compañías diferentes.
- Un programa puede tener diferentes contenidos, pero como mínimo se debe tener un contenido almacenado.
- Un contenido puede pertenecer a diferentes programas.

- Un programa puede tener más de una crítica realizada; admitiendo ninguna crítica como mínimo.
- Una crítica de un programa al menos será realizada por un autor, y en caso de que sea realizada por más autores se almacenará como máximo dos autores más.
- Una crítica no puede pertenecer a diferentes programas ya que es algo único por programa.
- Un país puede tener actividad con diferentes programas, pero un mismo programa puede tener relación con distintos países. Se almacenará por cada relación de país con programa, la actividad que se realiza (si es producción, creación, realización) además de almacenar la fecha en la que se hace dicha actividad.
- Un país vendrá identificado a través de una clave definida por la norma ISO International Codes For Names PF Countries and Curriences.
- Un idioma vendrá identificado a través de una clave definida por la norma ISO-IEC 639-2.
- Un programa puede tener asociados diferentes idiomas y un idioma puede aparecer en varios programas.
- Un programa puedes ser traducido, subtitulado o doblado en diferentes idiomas.
- Un programa puede tener más de una versión especificada, pero no es condición necesaria que tenga al menos una versión.
- Un programa puede ser compuesto o no compuesto de más programas.
- Un programa no compuesto puede ser una serie.
- Una serie estará formada por capítulos, aunque cabe la posibilidad de tener almacenados los datos principales de la serie sin tener ningún capítulo almacenado.
- Un programa compuesto estará formado por otros programas. Como mínimo deben existir dos programas para que se genere dicha relación.
- En un programa de tipo serie se pueden modificar los participantes y los idiomas que le caracterizan.

- Un país al menos tendrá una alerta definida y una alerta al menos pertenecerá a un país.
- Un programa puede tener varios títulos en diferentes idiomas, mínimo un programa tendrá un título. Un título en un idioma pertenecerá a un único programa.

7.3. Esquema Entidad/Interrelación

En el diagrama de la Figura 24 y Figura 25 se presenta un esquema conceptual definiendo los conceptos y relaciones más importantes a tener en cuenta en la base de datos que da lugar al sistema de la propuesta de estándar. En dicho diagrama se ha plasmado exactamente la información que se ha definido en esta propuesta.

7.4. Explicación del Esquema E/R

En este diagrama existe una entidad principal, denominada [*PROGRAM*] en la cual se están almacenando diferentes datos característicos de éste. Esta entidad se relaciona con otras entidades necesarias para el sistema, a través de relaciones y diversos atributos.

Los atributos de la entidad [*PROGRAM*] son los siguientes:

- Desc_target nos indica el tipo de audiencia para la cual se ha creado dicho programa.
- Color nos indica si el programa está creado en color o es en blanco y negro.
- Audio nos informa sobre si el programa es con sonido o es mudo.
- Promotional_Information nos refleja la información que se está repartiendo entre la sociedad para promocionar dicho programa.
- Duration es un campo que nos indicará la duración de tiempo de dicho programa.
- Rights, es un atributo compuesto por dos atributos, Summary y URL_Rights. El atributo Summary es un texto con un pequeño resumen de los derechos de autor, etc. que tiene dicho programa, indicando en todo momento los más importantes. En cambio, el campo URL_Rights nos informa de la URL en la que podemos encontrar el documento original de manera completa.

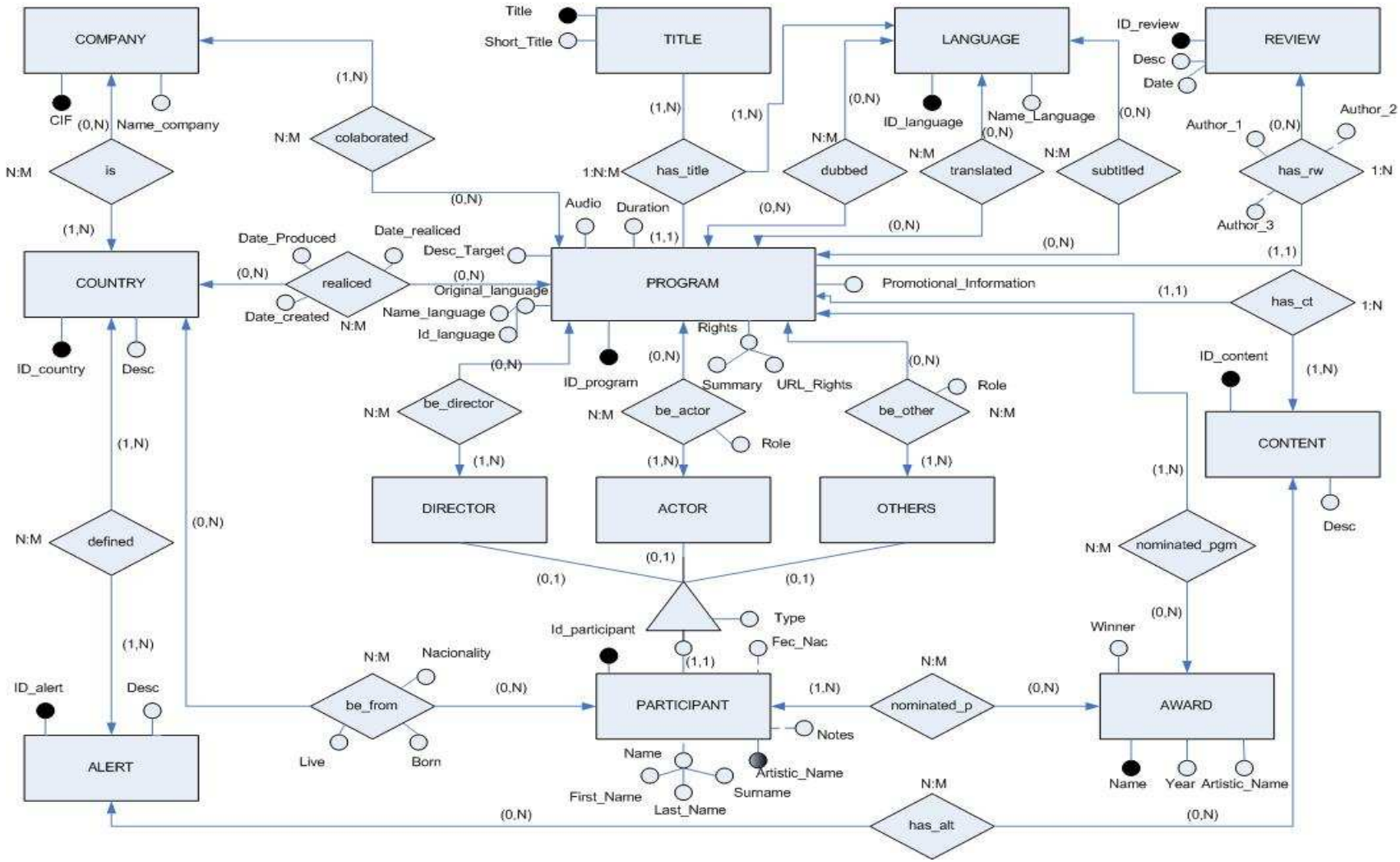

Figura 24 Diagrama E/R (1) propuesta de meta-modelo

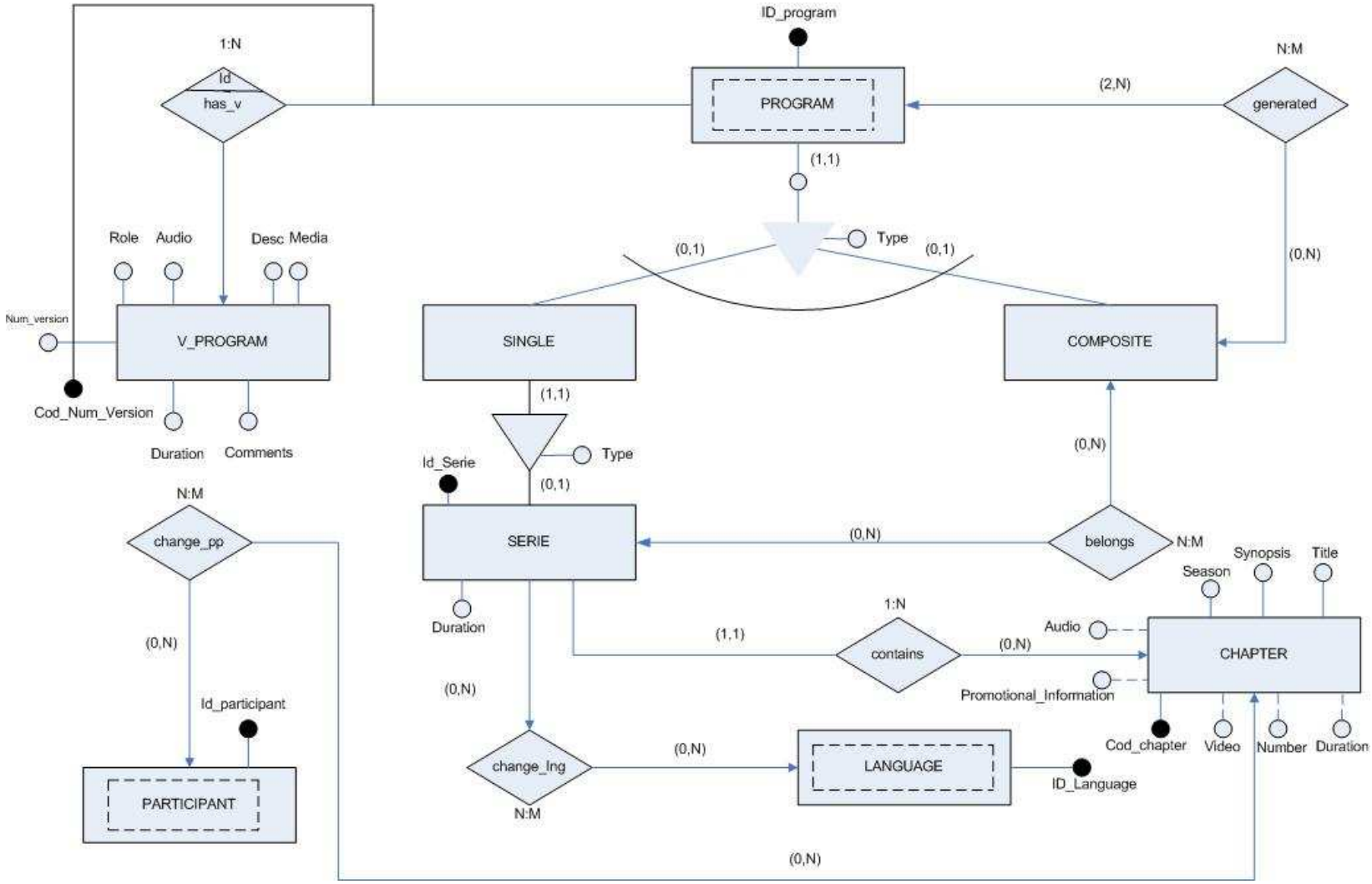

Figura 25 Diagrama E/R (2) propuesta de meta-modelo

- Id_program, es el identificador principal del programa, con esta clave podemos identificar de manera unívoca cada programa informado en este sistema.

Existe una interrelación denominada *has_rw* entre las entidades [*PROGRAM*] y [*REVIEW*]. Dicha interrelación es del tipo 1:N, donde una crítica es única y exclusiva de un programa. En dicha interrelación se almacena el Autor principal que ha desarrollado la crítica a través del atributo Author_1 y en caso de que existan más autores se almacenará como máximo dos autores más Author_2 y Author_3. El fin de esta relación es tener almacenadas todas las críticas que se realicen a un determinado programa. Estas críticas vendrán identificadas por Id_review, además almacenaremos el nombre del autor. También se guarda información sobre la fecha en la que se ha generado la crítica y una descripción, la cual refleja la crítica en sí.

Entre las entidades [*PROGRAM*] y [*LANGUAGE*] existen tres interrelaciones, denominadas *dubbed*, *translated* y *subtitled*, todas ellas de tipo N:M. La interrelación *dubbed* nos indica el idioma o idiomas en los que dicho programa se ha doblado. La interrelación *translated* nos indica los idiomas en los que se ha traducido, y por último, la interrelación subtitled nos informa sobre los idiomas en los que el programa se ha subtitulado. Por cada [*LANGUAGE*] se almacenará un identificador y el nombre del idioma sin abreviaturas ni siglas.

La entidad [*PROGRAM*] también se relaciona con una entidad denominada [*TITLE*] y con la entidad [LANGUAGE], a través de una interrelación de tipo 1:N:M llamada *has_title,* cuyo fin es tener almacenados todos los títulos que un programa adquiera en cada idioma. Por cada título se guardará su título principal (campo por el que identificaremos al programa) y su título de manera abreviada.

La entidad [*COMPANY*] tiene una interrelación de tipo N:M, denominada *colaborated,* con la entidad [*PROGRAM*], la cual tiene por objetivo almacenar los gastos que dicho programa ha generado a esa compañía a través de los servicios que ésta ofrezca. Una compañía vendrá identificada por su CIF, almacenando también como atributo el nombre de la compañía.

La interrelación *has_ct*, es una interrelación entre las entidades [*PROGRAM*] y [*CONTENT*]. Es de tipo 1:N ya que un contenido es exclusivo de un programa. Para cada contenido almacenaremos una descripción con el tipo de contenido que este programa tenga. Además la entidad [*CONTENT*] se encuentra relacionada con la

entidad [*ALERT*] a través de una interrelación, denominada *has_alt*. Dicha interrelación, de tipo N:M, nos informa de las alertas que tiene cada contenido del programa. De la entidad [*ALERT*] se almacena un identificador y una descripción con el tipo de alerta.

A la hora del tratamiento de los participantes de un programa, se ha generado una jerarquía parcial con solapamiento, cuyo supertipo es [*PARTICIPANT*] y como subtipos los tres tipos de participantes que pueden trabajar en un programa, es decir, [*DIRECTOR*], [*ACTOR*] y [*OTHERS*].

Cada participante, sea del tipo que sea, vendrán identificados por un código de participante. Se almacenará el nombre completo de cada participante (Nombre, Apellido1, Apellido2), su nombre artístico, su fecha de nacimiento, su nacionalidad y una pequeña descripción con características relevantes de cada participante (Notes).

Para relacionar estos participantes con el programa se han creado tres interrelaciones diferentes, las cuales son las siguientes:

- La interrelación *is_director*, de tipo N:M, entre las entidades [*PROGRAM*] y [*DIRECTOR*] , trata los directores que hay en los programas. En un programa como mínimo habrá un director.
- La interrelación *is_actor, de tipo N:M,* entre las entidades [*PROGRAM*] y [*ACTOR*], trata los actores que intervienen en dicho programa, almacenando para cada uno de ellos, en la interrelación is_actor, el role que toma en el programa. Dicho role nos informará si es actor principal, secundario u otros. Por ejemplo, en la película Jamón Jamón que se encontraría relacionada con el participante Penélope Cruz, el atributo de la interrelación role admitiría el valor de Principal.
- La interrelación *is_other,* de tipo N:M, entre las entidades [*PROGRAM*] y [*OTHERS*], trata todos aquellos participantes que no sean ni directores del programa, ni actores de éste, almacenando para cada uno de ellos el papel que desempeña en ese programa. Por ejemplo, el participante con identificador 12456 tiene como role en el programa Cifras Y Letras como estilista.

Entre las entidades [*COUNTRY*] y [*PARTICIPANT*] existe una interrelación llamada *be_from*, de tipo N:M. Esta interrelación realiza la función de tener ligados los países con los participantes para así tener almacenado si ese país es su

nacionalidad, si nació allí o si su domicilio habitual está registrado en ese país; de ahí que la interrelación tenga los atributos Nacionality, Born y Live.

Para el tratamiento de los diferentes tipos de programas se ha adoptado la idea de introducir una jerarquía parcial con solapamiento, cuyo supertipo es [*PROGRAMT*] y como subtipos los dos tipos de programas, es decir, [*SINGLE*] y [*COMPOSITE*], donde un programa de tipo Single será todo aquel programa que no se encuentre formado por ningún programa más, todo lo contrario que ocurre a un programa de tipo Composite, que es todo aquel programa que se compone de diferentes programas. Además, el propio subtipo [*SINGLE*] se convierte a la vez en el supertipo de otra jerarquía, en este caso total y sin solapamiento para poder tratar la información de los programas que son series. El subtipo de esta jerarquía es la entidad [*SERIE*] en la cual se almacena su identificador Id_serie formado por dos atributos, Identifier y Time, donde este último nos indica la temporada de la serie. Puede darse el caso de que nos encontremos ante una serie que sea un trabajo de tipo compuesto, por lo que existe una relación entre el subtipo [*SERIE*] y el subtipo [*COMPOSITE*], perteneciente a la primera jerarquía de tipos de programas, llamada *belongs* y de tipo N:M. A parte, este subtipo [*COMPOSITE*] se relaciona directamente con la entidad [*PROGRAM*] con la interrelación *generated* de tipo N:M, que nos indica los programas que forman dicho programa compuesto, para lo cual siempre un trabajo compuesto se encontrará formado por dos programas como mínimo.

El subtipo [*SERIE*] a parte de relacionarse con el subtipo [*COMPOSITE*] también tiene una interrelación [*LANGUAGE*]. Esta interrelación se ha creado con el objetivo de que una serie no tiene obligación de tener siempre los mismos idiomas, por lo que la interrelación *change_lng* es de tipo N:M y con ella se tiene almacenado en todo momento los diferentes idiomas que esa serie tiene. Al igual ocurre con la interrelación *change_pp* que relaciona la entidad [*PARTICIPANT*] con la entidad [*CHAPTER*] para tener almacenado en todo momento los diferentes participantes que colaboran en cada capítulo de la serie que se está tratando, dicha interrelación es de tipo N:M.

Además el subtipo [*SERIE*] se interrelaciona con la entidad [*CHAPTER*] a través de una interrelación de denominación *contains* de tipo 1:N, a través de esta interrelación se almacenan los capítulos que pertenezcan a una determinada serie de una determinada temporada. Por cada capítulo se almacena su código de capítulo, su

sinopsis, su título, su número de capítulo dentro de la serie si está disponible al igual que la temporada a la que pertenece y la duración de dicho capítulos, ambos son atributos opcionales, sin olvidar la posibilidad de almacenar el audio y video si éstos son conocidos.

Existe una interrelación denominada *nominated_pgm* entre las entidades [*PROGRAM*] *y [AWARD]* cuyo objetivo es tener identificadas las nominaciones que un programa tenga a diferentes premios. Por ejemplo, Titanic nominado en los Oscar con categoría de Mejor Película y con categoría de Mejor Guión.

Además existe una interrelación semejante a *nominated_p* pero denominada *nominated_p*, entre las entidades [*AWARD*] y [*PARTICIPANT*] . Dicha interrelación almacena el número de veces que ese participante ha sido nominado a dicho premio.
Se encuentra creada una interrelación con dependencia en identificación de la clave principal de [*PROGRAM*], denominada *has_v* con la entidad [*V_PROGRAM*]. La interrelación has_v es de tipo 1:N. Almacena todas las versiones que se creen de un programa, almacenando la descripción del cambio, campo obligatorio, y como campos opcionales se almacenan el color, el audio, la media, la duración (ya que podemos encontrarnos ante una versión extendida del programa, la cual alargará la duración de éste), los comentarios (en caso de que se añadan comentarios de los propios actores) de la nueva versión, ya que cualquier cambio en dichos atributos generarán una nueva versión. La clave principal será el código de versión del programa Cod_V_Program. También se almacena el número de versión de la que se esté tratando.

Existe una interrelación de tipo N:M, denominada *realiced,* entre las entidades [*PROGRAM*] y [*COUNTRY*] para relacionar todos aquellos países que se ven involucrados (en cualquier aspecto) en el programa. De cada país almacenamos un identificador y una descripción que tiene el nombre del país informado por completo (sin siglas alguna). En dicha interrelación se informan los atributos Date_Produced, Date_Created, Date_Realiced, los cuales almacenan las fechas de la creación, producción o generación en las que se hayan realizado esas actividades en ese país .

Por último, la entidad [*COUNTRY*] se relaciona con la entidad [*ALERT*] con el fin de tener registrado las alertas de contenido que cada país tiene definido como norma. Dicha interrelación es *defined* y de tipo N:M.

Básicamente este esquema se basa en identificar un programa a través de 8 dígitos y una letra al final (formato como el DNI). Los 8 dígitos identifican al programa original y la letra identifica a la versión. En caso de que el programa no tenga ninguna versión, se almacenará una X como último dígito del identificador. Se generará una versión nueva del programa por cada cambio que se realice en las características de color, audio y media. Los cambios en idiomas, aunque a primera vista, parezca que se debería generar una nueva versión, según se ha diseñado el sistema no es necesario, ya que por cada programa se almacenan todos los idiomas que se ven reflejados en la creación y transmisión de un programa. Por ejemplo, la película "Sex On The City" tendría como posible identificador 09788987X, pero la versión de esta película en la que aparecen comentarios de los propios actores, de diseñadores, etc. admitiría como identificador 097888987R.

7.5. Supuestos semánticos adicionales al diagrama E/R

En este apartado definimos las propiedades que no pueden ser expresadas en el diagrama E/R generado.

7.5.1. Dominios

En este apartado definimos los valores que pueden admitir los atributos de las entidades definidas. Dichos valores son definidos en el último nivel de detalle, ya que realmente es del que necesitamos la información.

- El idioma con el que se produce un programa será almacenado mediante un código de idioma, el cual seguirá la norma ISO 639-2 [5].
- Un programa puede ser de dos tipos compuesto por varios programas o ser único. Este tipo viene dado por el campo Type que tiene la jerarquía de programa. Los valores que admite son los siguientes:

  Dominio Type {jerarquía de Program} = {Composite, Single}
- Un participante puede realizar una actividad en un programa, actividad que se basará en el role que éste tenga y vendrá dado por algún valor de los del siguiente rango:

  Dominio Tye {jerarquía de Participant} = {Actor, Director, Others}

- El país que colabora en la producción del programa será almacenado mediante un código de país definido por la norma ISO International Codes For Names PF Countries and Curriences [16].
- Un programa de tipo Single a su vez podrá optar a tener como Type en dicha jerarquía Serial.

Dominio Single = {Serial, Otro}

- El atributo Colour de la entidad principal [*PROGRAM*] admitirá valores comprendidos en el siguiente rango de valores:

Dominio Colour = {Colour, Black & White}

El mismo dominio de valores será válido para el atributo Colour pero de la entidad [*V_PROGRAM*].

- El atributo Audio de la entidad principal [*PROGRAM*] admitirá valores del siguiente dominio:

Dominio Audio = {Silent, No Silent}

Este dominio también es el correspondiente para el atributo Audio de la entidad [*V_PROGRAM*].

- El dominio para el atributo Serial que tiene la entidad principal [*PROGRAM*] admite los siguientes valores:

Dominio Serial = {Yes, No}

- El atributo Ppal de la interrelación *is_actor* admitirá valores del siguiente dominio:

Dominio Ppal = {Yes, No}

- El atributo Media de la entidad [*V_PROGRAM*] admitirá valores correspondientes al siguiente rango de valores:

Dominio Media = {Television, Cinema, Other}

- El atributo Frequency de la entidad [*SCHEDULING*] solamente podrá admitir valores comprendidos en el siguiente dominio:

Dominio Frequency = {One off, Hourly, Daily, Every 2-6 days, Weekly, Monthly, Seasonnaly, Annually, Regularly}

- El atributo Desc_target se informará con valores del siguiente dominio:

Dominio Desc_target = {Young Children, Children, Teenagers, Young Adults, Senior Citizens, Regional, Local, Ethnic Groups, Inmigrant Groups, Linguistic

Groups, Mainshoppers, Business People, Farmers, Fishermen /Sailors, Schoolchildren/Students, Unemployed People, Motorists, Tourists, Women}

7.5.2. Otros

- Los campos Author_1, Author_2 y Author_3 vendrán informados con el siguiente formato:

  Apellido1, Nombre

  Apellido1 Apellido2, Nombre

  Por ejemplo,

  Alonso Piña, Raquel
- Todos los campos que tratan fechas, como Date de la entidad [*REVIEW*], Fec_Nac de la entidad [*PARTICIPANT*], y las de la interrelación realiced Date_Created, Date_Produced, Date_Realiced tendrán el siguiente formato de fecha DD-MM-YYY, por ejemplo, 21-01-1975.
- El campo Times en la interrelación *nominated_p* y *nominated_pgm* vendrá dado por un valor numérico entero y siempre positivo.
- El título del programa siempre tendrá la primera letra de cada palabra en mayúscula. En caso de que tenga como primera palabra el artículo The, se especificará de la siguiente manera, *World War, The.* Lo mismo ocurre para el campo Title de la entidad [*CHAPTER*].
- El campo Sypnosis de la entidad [*CHAPTER*] no admitirá una longitud mayor de 200 caracteres alfanuméricos. Al igual ocurre con el campo Notes de la entidad [*PARTICIPANT*].
- El campo Number de la entidad [*CHAPTER*] admitirá únicamente valores numéricos positivos y enteros. Nunca admitirá el valor 0.
- El campo Artistic_Name vendrá con las iniciales en mayúscula obligatoriamente.
- El campo Duration del atributo compuesto Transmission vendrá dado en minutos. No se debe indicar la medida de minutos en dicho campo.

7.6. Discusión

Como ya bien citamos anteriormente en este documento, para realizar un modelo eficaz, eficiente, a la vez que consistente, se optó por emplear ideas que se reflejaban en los modelos Tv-AnyTime, ISAN y Escort.

En la Figura 26 se puede observar un pequeño resumen de las ventajas y desventajas de cada uno de los tres estándares analizados durante este proyecto de fin de carrera.

Estándar	Ventajas	Desventajas
ISAN y V-ISAN	- Tratamiento de versiones. - Jerarquía de participantes. - Tratamiento de tipos de trabajos audiovisuales.	-Identificación de participantes por apellidos. -Identificación de países por nombres -Identificación de compañía por nombre. -Contradicción de identificación en los episodios. -No existe una numeración de episodios de una serie. -Restricción de almacenamiento de 7 actores de tipo Others. -No indica si los actores son principales o no. -Carencia de fecha de publicación. -Déficit al no relacionarse la compañía con el país perteneciente. -Déficit al no relacionarse el país con su idioma.
TV-AnyTime	-Buena especificación y definición de términos. -Tratamiento de premios por programa. -Tratamiento de críticas realizadas a dicho fondo audiovisual.	-Empleo ineficiente del campo Content_Alert. -No trata los premios con los participantes. -No se almacenan las series en sí.
Escort	-Correcta y amplia definición de dominios de atributos. -Identificación unívoca de un programa de sencilla, clara y rápida manera.	-Se basa solo en las propiedades "físicas" del programa. -Carencia de no almacenar versiones. -No se almacena información sobre las series ni sobre los participantes.

Figura 26 Ventajas y desventajas de ISAN, TV-AnyTime y Escort

A continuación desglosamos de manera más detallada las ideas que se han decidido plasmar en este nuevo modelo elaborado por cada estándar.

- **TV-AnyTime**

  Este modelo se caracteriza por ser un modelo bastante complejo a la hora de almacenar información sobre trabajos audiovisuales. Almacena demasiada información que no es necesaria para la clasificación de un trabajo de este tipo.

  Algunas de las ideas que se han tomado de este modelo para crear el modelo My_Model son las siguientes:

  - Las relaciones que un programa puede tener con un país definidas por Tv-AnyTime son de creación, realización y producción, pero al igual que en este modelo existe una relación por cada tipo de acción, almacenando para cada una de ellas una fecha de actividad, se ha decidido en el modelo que se ha creado, generar una única relación entre la entidad [*COUNTRY*] y la entidad principal [*PROGRAM*] y almacenar en esta la fecha de la actividad que se lleva a cabo en dicho país.
  - Este modelo TV-AnyTime tiene ideas reflejadas que ningún otro modelo de los estudiados en este proyecto fin de carrera han desenvuelto y que cobran importancia a la hora de almacenar y clasificar un programa. Nos referimos a información como las nominaciones a premios, tanto de programas como de participantes.
  - En el modelo que se ha creado, se ha relacionado a los premios con los propios programas y con los participantes, en cambio en el modelo TV-AnyTime únicamente se relaciona con los programas. Además, tampoco se almacena en ningún momento el número de veces que dicho programa ha sido nominado a ese mismo premio; en la propuesta si se encuentra reflejado tanto para las nominaciones a premios de programas como de participantes.

  Una idea que se ha plasmado en este metamodelo, es el tema de las críticas realizadas a un programa. Son opiniones que merecen estar almacenadas para tener una pequeña visión de lo que la audiencia opina sobre ese programa.

- **Escort**

  El modelo definido por Escort aunque es de fácil manejo, se observa que baja a nivel muy detallado en atributos que realmente no son de interés. Por ejemplo, en el origen de un programa, o llegar al detalle de si ese idioma lo habla un mayor o menor número de personas y si es un dialecto doméstico.
  Al igual que almacenar los costes que ha supuesto generar ese programa, es un dato totalmente inconsistente a la hora de realizar la clasificación de programas.

  Este estándar trata la planificación (*Scheduling*) de un programa como algo tan sencillo como es un atributo simple. Es totalmente imposible que se pueda recoger en dicho campo toda la información que conlleva la planificación de retransmisión de un programa.

  Por lo que de este modelo no se ha optado por adquirir muchas ideas y, básicamente, las que se han reflejado en la propuesta han sido con variaciones personales ya que sino creaban inconsistencias.

  Por ejemplo, la idea de almacenar los derechos que tenga un programa es muy eficiente, ya que es un tema legal que debe estar informado en cualquier momento. Pero se ha variado un poco con respecto al atributo que el modelo Escort define; en la propuesta se ha definido Rights como un atributo compuesto formado por dos campos Summary y URL_Rights, con el fin de poder tener almacenado tanto un resumen con los derechos más importantes como una URL como dirección para poder acceder al documento completo de dichos derechos.

  Se ha incluido una entidad denominada [*CONTENT*], que también el modelo Escort lo emplea, para almacenar los diferentes contenidos que tiene dicho programa. Pero se ha incluido una variación que involucra también al modelo TV-AnyTime, ya que dicho modelo inserta en sus especificaciones el tener almacenado una alerta del tipo de contenido del programa.

  Del modelo Escort no se han tomado más ideas de las que se acaban de indicar.

- **ISAN**

El modelo definido por ISAN es un modelo sencillo, de fácil manejo y entendimiento por el usuario. Además es bastante concreto a la hora de almacenar información, son datos de características principales, los más importantes. De ahí que en la propuesta generada se base concretamente en este modelo.

Los tres puntos más fuertes que se han tomado ideas para plasmarlas en el modelo de la propuesta las dos jerarquías, una para el tipo de programa y la otra para el tipo de participante, además de basarse en la idea de como tratar una versión de un programa.

La jerarquía empleada en ISAN para tener clasificados los diferentes tipos de programa ha sido modificada para el modelo de la propuesta. Se ha mantenido la primera clasificación entre trabajos compuestos y no compuestos, pero además se ha distinguido que un programa no compuesto puede ser tipo serie y por lo tanto estar formado por varios capítulos, de esta manera hemos hecho más eficiente el modelo al poder tener almacenado por un lado el nombre y temporada de la serie en sí, y luego ya se almacenan los capítulos que forman dicha serie. En nuestro modelo de propuesta de estándar además se ha optado por incluir una relación imposible en el modelo ISAN, que es si nos encontramos ante una serie que sea un trabajo compuesto. Aunque cabe decir que la idea de que un programa puede estar compuesto por varios tipos distintos de programas, es algo muy común en los programas actuales, por lo que se ha decidido plasmar esa relación de programa con programa tratando los programas compuestos.

La jerarquía empleada en el modelo definido por ISAN de diferenciar los participantes en tres tipos, Director, Actor y Otros también ha sido incorporada en el modelo de la propuesta desarrollada. Es una perfecta clasificación de tipos de participantes, pero se han modificado las cardinalidades mínimas y máximas que el estándar ISAN define ya que de esa manera se restringe de manera amplia los datos lo que lleva consigo una ineficiencia de dicho sistema. Por ejemplo, ISAN definía como mínimo el almacenamiento de tres actores y siempre y cuando estos fueran los actores principales. En el modelo que se ha definido en este proyecto, no se ha realizado dicha restricción y

además se indica en cada momento que se añada un nuevo actor si es principal, secundario u otro tipo de actor en dicho programa. Además se ha realizado una mejora en relación a la información que se almacena sobre los participantes. Como identificador secundario se optará por el nombre artístico del participante (este concepto en el modelo ISAN no se tiene en cuenta y hace que éste no sea demasiado consistente al ahora de poder identificar un participante), el nombre vendrá detallado con nombre, primer apellido y segundo apellido (en ISAN únicamente se almacenaba el nombre, opcionalmente, y primer apellido), por lo que podía haber problemas y duplicidad de datos a la hora de identificar un participante en ISAN, por ejemplo si había dos familiares trabajando como Penélope Cruz y Mónica Cruz.

Sobre la versión, el tratamiento de ésta, se ha modificado en relación a la definición de ISAN ya que se han añadido como posibles cambios para crear nuevas versiones: la duración (por si nos encontramos ante una versión extendida del programa) y un campo denominado comments (por si se añaden comentarios al propio programa), además de posibles cambios, como están definidos en ISAN, de color, audio.

Sin tener relación con ninguno de los modelos estudiados, se han añadido ideas propias.

Un programa puede tener diferentes idiomas relacionados, unos por traducciones que se realicen, otros porque se decida subtitular en algún determinado idioma, al igual que se decida doblar el programa en diferentes idiomas. Para ello, se ha decidido generar tres relaciones diferentes para así almacenar los idiomas según la actividad que se vaya a realizar con ellos.

De los participantes, es interesante almacenar la nacionalidad que tienen, por ello es bastante efectivo que exista una relación entre el país y el participante.

También una relación que no existe en ningún modelo, es el tener registradas las alertas de contenido por cada país, ya que cada país las tiene definidas de una manera u otra.

En conclusión, la propuesta que se ha creado consigue ser un modelo bastante eficiente y eficaz tanto en la clasificación como en el almacenamiento de información sobre programas audiovisuales; todo ello obtenido tras unificar las mejores ideas que

cada modelo aportaba y las modificaciones que se han creído oportunas en este proyecto de fin de carrera, con el fin de obtener un sistema consiste y sencillo en todos los términos.

8. PLANIFICACIÓN FINAL Y ANÁLISIS DE COSTES

Al finalizar el proyecto se realiza una planificación final y un análisis de costes comparando estos datos con los obtenidos tras la planificación inicial generada.

Con dicha comparación se puede ver cuáles son las tareas que se han realizado en el plazo definido previamente, en cuáles se han tardado menos tiempo del planificado, o en cuáles se ha tenido que alargar el plazo planificado además de identificar las causas que han generado dicho retraso en la entrega del proyecto. Todo ello se estudia con el fin de en un futuro poder ir mejorando la planificación y por lo tanto haya un menor margen de error en la fecha de entrega a estimar.

8.1. Planificación final

Para esta planificación final se he creado otro diagrama GANTT una vez ha finalizado el desarrollo del proyecto.

Este diagrama Gantt se puede ver en la Figura 28.

8.2. Costes

Al igual que se ha estudiado la variación del tiempo estimado con respecto a la fecha de entrega, también hay que estudiar los costes, para determinar si realmente el coste final corresponde con el estimado o ha variado con respecto al coste inicial.

Según la planificación final obtenida se ha tardado más tiempo en realizar dicho trabajo, lo que conlleva un aumento del presupuesto inicial. Se habían estimado 45120 euros como presupuesto inicial, en el que se veían reflejadas 188 jornadas laborales (1496 horas), pero en realidad se han empleado 209 jornadas laborales (1672 horas), lo que supone un presupuesto final de 50160 euros. Se observa un incremento de 5040 euros con respecto a lo calculado de manera previa a la realización del proyecto.

En la siguiente Figura 27 se observan las variaciones de coste por etapa definida:

Fase	Coste Inicial	Coste Final	Variación
Establecimiento de objetivos	240 euros	240 euros	0 euros
Gestión del Proyecto	2160 euros	2160 euros	0 euros
Documentación y Estado del Arte	2400 euros	2400 euros	0 euros
ISAN y V-ISAN	8640 euros	8640 euros	0 euros
TV-AnyTime	13440 euros	16320 euros	2880 euros
Escort	11280 euros	11280 euros	0 euros
Propuesta	6960 euros	9120 euros	2160 euros
Total Proyecto	45120 euros	50160 euros	5040 euros

Figura 27 Variación de costes Inicial-Final

Figura 28 Planificación Final - Diagrama Gantt

Id	Nombre de tarea	Duración	Comienzo	Fin
1	***Inicio Proyecto***	0 días	lun 01/10/07	lun 01/10/07
2	**Establecimiento de objetivos**	1 días	lun 01/10/07	lun 01/10/07
3	**Gestión del Proyecto**	543 días	lun 01/10/07	jue 29/10/09
4	Selección de Metodología	1 días	mar 02/10/07	mar 02/10/07
5	Selección de Ciclo de Vida	1 días	mié 03/10/07	mié 03/10/07
6	WBS	1 días	jue 04/10/07	jue 04/10/07
7	Planificación Inicial	2 días	vie 05/10/07	lun 08/10/07
8	GANTT	1 días	mar 09/10/07	mar 09/10/07
9	Planificación Final	2 días	mar 27/10/09	mié 28/10/09
10	Análisis de Costes	1 días	jue 29/10/09	jue 29/10/09
11	**Documentación y Estado del Arte**	538 días	mar 09/10/07	vie 30/10/09
12	Documentación del proyecto-1	1 día	mié 10/10/07	mié 10/10/07
13	Documentación del proyecto-2	1 día	mar 16/10/07	mar 16/10/07
14	Documentación del proyecto-3	1 día	mar 11/12/07	mar 11/12/07
15	Documentación del proyecto-4	1 día	jue 10/07/08	jue 10/07/08
16	Documentación del proyecto-5	1 día	mié 18/03/09	mié 18/03/09
17	Documentación del proyecto-6	1 día	mié 12/08/09	mié 12/08/09
18	Documentación del proyecto-6	1 día	vie 30/10/09	vie 30/10/09
19	Estudio del Estado del Arte	3 días	jue 11/10/07	lun 15/10/07
20	**ISAN y V-ISAN**	36 días	mar 16/10/07	mié 05/12/07
21	Análisis de ISAN y V-ISAN	11 días	mar 16/10/07	mié 31/10/07
22	Diseño Conceptual	10 días	jue 01/11/07	mié 14/11/07
23	Análisis crítico de ISAN y V-ISAN	15 días	jue 15/11/07	mié 05/12/07
24	**TV-AnyTime**	68 días	vie 04/04/08	mié 09/07/08
25	Análisis de TV-AnyTime	21 días	vie 04/04/08	lun 05/05/08
26	Diseño Conceptual	24 días	mar 06/05/08	vie 06/06/08
27	Análisis crítico de TV-AnyTime	23 días	vie 06/06/08	mié 09/07/08
28	**ESCORT**	47 días	vie 09/01/09	mar 17/03/08
29	Análisis de Escort	14 días	lun 12/01/09	jue 29/01/09
30	Diseño Conceptual	10 días	vie 30/01/09	jue 12/02/09
31	Análisis crítico de Escort	23 días	jue 12/02/09	mar 17/03/09
32	**PROPUESTA**	158 días	mié 18/03/09	lun 26/10/09
33	Estudio de Propuesta	11 días	jue 19/03/09	jue 02/04/09
34	Definición de Propuesta	10 días	jue 02/04/09	jue 16/04/09
35	Diseño Conceptual	4 días	vie 17/04/09	mié 22/04/09
36	Análisis Crítico 1	5 días	jue 23/04/09	mié 29/04/09
37	Análisis Crítico 2	8 días	jue 15/10/09	lun 26/10/09
38	***Fin de Proyecto***	0 días	vie 30/10/09	vie 30/10/09

Figura 29 Tareas y Tiempos de Planificación Final

En la Figura 29 se detallan los períodos de tiempo para cada tarea. Hay que destacar que el horario laboral es de 8:00 a 16:00 y que la jornada es de Lunes a Viernes.

9. CONCLUSIONES Y TRABAJOS FUTUROS

9.1. Conclusiones

En la actualidad no contamos con un estándar a nivel internacional que englobe todos los tipos de fondos audiovisuales y que clasifique y enumere a éstos de manera genérica. El estudio que se ha realizado en este proyecto de fin de carrera ayuda a contemplar de manera extensa cómo se encuentra el panorama del almacenamiento de dichos ítems.

Se ha observado que existe un gran número de estándares, dentro de los cuales podemos encontrar TV-AnyTime, Escort, DVB, SMEF, etc. Lo más llamativo tras el estudio de algunos estándares es que cada uno trata un tipo de fondo audiovisual y almacena diferente información sobre éstos, hay estándares que abren camino en la parte financiera y administrativa, otros únicamente se centran en los atributos físicos del ítem, como por ejemplo, idioma, formato, sonido, color, etc.

Como ya sabemos, la industria del cine y la televisión día a día va aumentando y por lo tanto mayor información para tener almacenada, por lo que se necesita en un futuro próximo, no muy lejano, un identificador universal para cualquier tipo de trabajo audiovisual, ya sea telediario, serie televisiva, comedia, anuncio de publicidad, concurso, reality-show, etc. que nos facilite la búsqueda de cualquier material de este tipo y así poder tenerlo almacenado en una base de datos.

En este proyecto de fin de carrera se ha generado una propuesta de identificador, el cual además de tener características de los estándares analizados .TV-AnyTime y Escort, tiene ideas propias que se han creído convenientes, para que dicho estándar sea lo más eficaz y eficiente posible, evitando la redundancia en la información, haciendo mayor hincapié en las características más importantes y dando menos importancia a las más relevantes y más ineficaces a la hora de numerar dicho trabajo. Se ha conceptualizado toda la propuesta de manera gráfica, a través de un diagrama E/R en el cual se puede observar claramente la información que se ha creído más oportuna para su almacenamiento y de la manera que se está tratando dichos atributos.

9.2. Trabajos futuros

Como bien se ha indicado en el apartado de las conclusiones, lo más importante es generar un estándar de manera universal para esa identificación de trabajos audiovisuales. Para ello se podría tomar como base la propuesta generada a la vez que se analiza cada estándar existente y así tomar como referencia características importantes e incluirlas en dicha propuesta; de esta manera se podrá obtener un estándar muy completo e identificar de manera unívoca a cada fondo audiovisual.

En este proyecto de fin de carrera se han realizado unos análisis críticos de los estándares analizados, con el fin de que cada Organización estudie dichas mejoras y si realmente éstas pueden ser necesarias para reforzar cada estándar. De esa forma podrán ampliar la definición de éste y así, establecer un modelo de datos más completo que el actual.

Además, se podría realizar el estudio de otro estándares y metamodelos relacionados con este proyecto fin de carrera.

En el metamodelo propuesto se podría incluir información sobre los identificadores de otros metamodelos, de forma que sea más fácil transferencias de información.

También se podría ampliar el estudio realizando tratando también el medio de difusión Radio, ya que únicamente se ha estudiado todo lo relativo a la televisión.

10. GLOSARIO DE TÉRMINOS

- AACS

Advanced Access Content System.

- AGICOA

Association of International Collective Management of Audiovisual Works.

- B2B

Bussiness to Bussiness.

- CISAC

International Confederation of Societies of Authors and Composer.

- CRID

Content Reference Identifier

- DVB

Digital Video Broadcasting

- E/R

Entidad/Interrelación.

- EBU

European Broadcasting Union.

- EPG

Electronic Program Guide.

- ESCORT

EBU System of Classification of Radio and Television.

- ETSI

European Telecommunicationes Standards Institute.

- FIAPF

International Federation of Film Producers Associations.

- I+D

Investigación y Desarrollo.

- ISAN

International Standard Audiovisual Number.

- ISO

International Standard Organization.

- IEC

International Electrotechnical Commission.

- KLV

Key Length Value.

- MPEG

Moving Picture Experts Group.

- MXF

Material Exchange Format.

- P/META

Conjunto de definiciones de EBU que proporciona metadatos para guardar información relacionada con materiales audiovisuales.

- PVR

Personal Video Recorder.

- SGBD

Sistema Gestor de Base de Datos.

- SMEF

Standard Media Exchange Framework.

- SMPTE

Society of Motion Picture and Television Engineers.

- TC/SC

Technical Comitee/Sub-Comitee.

- TVA

TV-Anytime. Conjunto de especificaciones de EBU que permiten el almacenamiento digital de servicios audiovisuales y de otro tipo, en plataformas de usuario, basándose en el mercado de masas.

- URL

Uniform Resource Locator.

- V-ISAN

Version Identifier of an ISAN.

11. BIBLIOGRAFÍA

11.1. Convencional

- [Aribsan-ISAN Version, 2007] Documento facilitado por la Agencia Iberoamericana del Registro ISAN.

- [ISAN, 2002] Information and documentation – International Standard Audiovisual Number (ISAN). ISO 15706: 2002

- [V-ISAN, 2004] Information and documentation – Guidelines on V-ISAN. ISO 15706-2: 2004

- [TV-AnyTime, 2000] J.-P. Evain: TV-Anytime metadata: preliminary specification on schedule! EBU Technical Review No. 284, (2000).

- [Escort, 1995] European Broadcasting Union: Escort: EBU System of Classification of RTV Programmes. Technical report (1995).

11.2. En Web

[1] Página Web de la Asociación Internacional sobre el estándar del número audiovisual (ISAN): http://www.isan.org

[2] Página Web del estándar TV-AnyTime: http://www.tv-anytime.org/

[3] Página Web de la Organización Internacional de estándares: http://www.iso.org

[4] Página Web de consulta del estándar MPEG-7: http://schemas.stylusstudio.com/mpeg7/frames.html

[5] Página Web de consulta de norma ISO 639-2: http://www.loc.gov/standards/iso639-2

[6] Página Web sobre DublinCore: http://dublincore.org

[7] Página Web de DVB: http://www.dvb.org

[8] Página Web de SMPTE: http://smpte.org

[9] Página Web de European Broadcasting Union sobre Escort (tech3322.php): http://www.ebu.ch/en/technical/metadata/specifications/index.php

[10] Página Web de consulta sobre la especificación XML1.0. http://www.w3.org/2001/XMLSchema

[11] Página Web de European Broadcasting Union sobre P/META (tech3295.php): http://www.ebu.ch/en/technical/metadata/specifications/index.php

[12] Página Web de la Conferencia Internacional sobre Dublin Core y Aplicaciones de Metadatos: http://dc2005.uc3m.es

[13] Página Web de consulta de norma NISO Z39.85-2007: http://www.niso.org/standards/z39-85-2007/

[14] Página Web de consulta de norma ISO 15836: http://www.iso.org/iso/catalogue_detail.htm?csnumber=37629

[15] Página Web de consulta de norma ISO 15938: http://www.iso.org/iso/catalogue_detail.htm?csnumber=39477

[16] Página Web de consulta de norma ISO International Codes OF Names PF Countries and Curriences: http://www.iso.org/iso/iso_cafe_country_and_currency_codes.htm