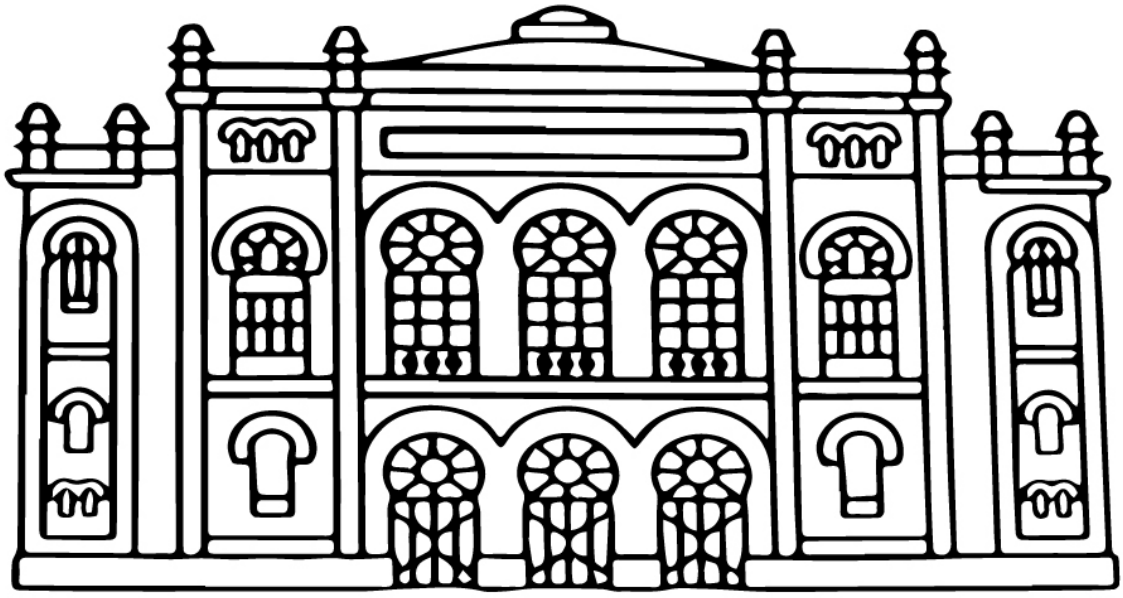

Base de Datos para el Centro de Documentación del Carnaval de Cádiz

Paloma Cortés Ocáriz
Grado de Ingeniería Informática, Tecnología de la Información
TFG Bases de Datos

Jordi Ferrer Duran
Santiago Codolà Vilahur

12 de Junio de 2017

Copyright © 2017 Paloma Cortés Ocáriz.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License".

FICHA DEL TRABAJO FINAL

Título del trabajo:	*Base de Datos para el Centro de Documentación del Carnaval de Cádiz*
Nombre del autor:	*Paloma Cortés Ocáriz*
Nombre del consultor/a:	*Jordi Ferrer Duran*
Nombre del PRA:	*Santiago Codolà Vilahur*
Fecha de entrega:	06/2017
Titulación::	*Grado de Ingeniería Informática*
Área del Trabajo Final:	*Bases de Datos*
Idioma del trabajo:	*Español*
Palabras clave	*Base de datos, Data Warehouse, Centro Documental, Carnaval*

Resumen del Trabajo

El presente proyecto tiene como objetivo la definición del modelo de datos que manejará el futuro Centro de Documentación del Carnaval de Cádiz y su definición mediante un esquema de base de datos.

Actualmente el proceso de documentación no está organizado ni sigue ningún tipo de protocolo archivístico; además, los datos se encuentran dispersos y desestructurados. Por tanto, la necesidad a satisfacer es tanto la estructura necesaria para el almacenamiento de datos como la necesaria para poner en marcha un proceso documental estandarizado y estricto, para lo que se introducen las recomendaciones y el uso de la ISAD(G).

El proceso de identificación de requisitos y definición del modelo conceptual, así como el de desarrollo y las pruebas, siguen una metodología iterativa e incremental. Una vez finalizado se diseña y desarrolla, en una segunda etapa, un conjunto de funcionalidades estadísticas dentro del enfoque de *data warehouse.*

El producto obtenido consiste en la definición del modelo a nivel físico sobre el sistema gestor de bases de datos PostgreSQL, el cual está completamente encapsulado mediante una API compuesta de los procedimientos almacenados suficientes para la actualización de la totalidad de los contenidos. Por otro lado, de forma totalmente independiente, se tiene el conjunto de funciones dedicadas a *data warehouse.*

Las conclusiones finales están relacionadas con la carga de trabajo que supone la construcción de un modelo de datos básico, principalmente atendiendo a la complejidad de implementar las peculiaridades de un dominio concreto del mundo real.

Abstract:

The main aim of this project is to define the data model needed by the future information system of the Cádiz's Carnival Center for Documentation through a database schema.

Nowadays the documentation process is not organized neither is under an archival description standard. Thus, the main needs are, on the one hand, the data storage schema and, on the other, a standardized documentation process; in order to supply the last need, the ISAD(G) recommendations and use are introduced.

The requirements gathering process, the conceptual model definition, the development and the testing phase are implemented through an iterative and incremental methodology. Once it is finished, a second phase based on the statistical features set is executed through a data warehouse approach.

The resulting product is the complete physical model. It is created in PostgreSQL as database management system and is completely encapsulated through an interface, as an API, which is composed of a minimal sort of stored procedures for the contents management. On the other hand, and in a self-contained way, the functions working on data warehouse are kept.

The final conclusions are related to the basis data model implementation workload, above all, taking into account the complexities of a specific domain in the real world.

Índice

1. Introducción 1
  1.1 Contexto y justificación del Trabajo 1
  1.2 Objetivos del Trabajo 2
    1.2.1 Generales 2
    1.2.2 Específicos 2
  1.3 Enfoque y método seguido 3
  1.4 Planificación del Trabajo 3
    1.4.1 Actividades 4
    1.4.2 Hitos 6
    1.4.2 Diagrama Gantt 7
  1.5 Breve descripción de los capítulos de la memoria 8

2. Gestión del Proyecto 9
  2.1 Alcance 9
    2.1.1 Incluye 9
    2.2.2 No incluye 9
    2.2.3 Líneas futuras 9
  2.2 Requisitos 11
    2.2.1 Requisitos funcionales 11
    2.2.2 Requisitos no funcionales 13
  2.3 Riesgos 15
    2.3.1 Análisis 15
    2.3.2 Plan de Contingencias 15
  2.4 Calidad 16
  2.5 Comunicación 17

3. Diseño conceptual 18
  3.1 Entidades 18
  3.2 Relaciones 26
  3.3 Restricciones de integridad 33
  3.4 Modelo UML 42
  3.5 Usuarios, perfiles y esquema externo 42
    3.5.1 Esquema Global 43
    3.5.2 Esquema *data warehouse* 43

4. Diseño lógico 44
  4.1 Esquema 44
  4.2 Normalización 52

5. Diseño físico 54
5.1 Nivel virtual 54
5.2 Tablas 54
5.3 Índices 58
5.4 Procedimientos almacenados 59
5.5 Construcción 65
6. *Data warehouse* 66
6.1 Diseño lógico 66
6.3 Implementación y esquemas externos 67
6.3.1 Agrupaciones ganadoras por concurso y año 67
6.3.2 Premios por entidad y función 68
6.3.3 Dedicación oficial al Carnaval por persona 69
6.3.4 Temas recurrentes en letras por año y modalidad 70
6.3.5 Soporte al proceso documental de las agrupaciones 71
7. Pruebas 72
7.1 Enfoque y metodología 72
7.2 Rendimiento 75
7.3 Baterías de pruebas y automatización 75
8. Conclusiones 76
9. Glosario 77
10. Bibliografía 80
11. Anexos 81

Lista de figuras

Figura 1. Tabla cronológica de actividades 6
Figura 2. Diagrama Gantt 7
Figura 3. Modelo conceptual 42

1. Introducción

El presente proyecto está destinado a la construcción de una base de datos que permita gestionar de forma centralizada los depósitos documentales y materiales del Aula de Cultura del Carnaval de Cádiz, en lo siguiente el Aula, que es institución sin ánimo de lucro vinculada oficialmente a la Universidad y al Ayuntamiento de Cádiz. En última instancia, el Aula requiere de la creación de un Sistema de Información para el Centro de Documentación del Carnaval de Cádiz, que de soporte integral a su labor de descripción y protección del patrimonio material e inmaterial relacionado a las fiestas de carnaval de la ciudad.

1.1 Contexto y justificación del Trabajo

El Carnaval de Cádiz es el evento por excelencia de la ciudad, declarado como Fiesta de Interés Turístico Internacional en 1980; es, de hecho, la representación artística de la idiosincrasia de sus ciudadanos. Tanto en la época de Carnaval como durante el resto del año involucra a miles de personas, muchas de forma profesional; y, además de la participación humana, se da un ingente de creación material e inmaterial, incluso se puede decir que recoge su historia a lo largo de los últimos siglos –tiene sus primeros registros documentales en el siglo XVI–.

La primera, y mayor, razón para tener un repositorio bien estructurado y regulado es el interés cultural que esta cuestión representa para sus ciudadanos. Por la misma, en 1995 se formaliza oficialmente el Aula de Cultura del Carnaval de Cádiz como entidad dedicada a la recopilación, documentación, protección y puesta en valor de todo el material relativo a esta fiesta.

El Aula surge de la necesidad que tiene la ciudad y los aficionados de compartir y conocer los orígenes e información de este evento anual. Hasta entonces, dicha información se había transmitido oralmente y los documentos físicos se encontraban distribuidos entre instituciones y colecciones privadas. Este primer intento de coordinación y unificación genera unos conjuntos de registros que datan de 1800.

Actualmente el Aula de Cultura del Carnaval de Cádiz, junto con otros entes vinculados a estas fiestas, se encuentra inmersa en diversas propuestas relacionadas con el reconocimiento y la protección de la misma. Las intenciones actuales van desde la creación del Museo del Carnaval hasta la declaración, por la Unesco, del Carnaval de Cádiz como Patrimonio Inmaterial de la Humanidad.

Este último proceso también requiere de la reestructuración y formalización de la información relevante, en este caso como objeto de defensa de la labor del Aula. Dicha información se encuentra dispersa entre los socios y colaboradores de la misma y tiene deficiencias en lo relativo a normativas de descripción documental.

Se consideran estos hechos como oportunidad para inicializar la creación de un Sistema de Información que permita coordinar todos los procesos involucrados en la protección, difusión y resto de actividades concernientes al patrimonio histórico, cultural y documental del Carnaval.

Para satisfacer la primera necesidad, la defensa de la institución ante el proceso de declaración del Carnaval de Cádiz como Patrimonio Inmaterial de la Humanidad, mediante el presente proyecto se desarrollará la base de datos que dará soporte al SI. No obstante, aunque el alcance del proyecto se limita al modelo mínimo necesario para la documentación estricta del inventario de bienes inmateriales registrados por el aula, el objetivo final es la puesta en producción de un SI dedicado al Centro de Documentación del Carnaval de Cádiz que permita gestionar todas las actividades del Aula, desde los fondos hasta la gran diversidad de actividades que promueve esta institución.

1.2 Objetivos del Trabajo

1.2.1 Generales

El objetivo principal es unificar, normalizar y centralizar todos los datos relativos a los depósitos propios y las colecciones privadas reconocidas por el Aula, incluidas las personas y entidades involucradas.

El proyecto general, la implantación de un sistema de información, de forma subyacente se presenta como oportunidad para la reorganización de los procesos documentales y la introducción de los procedimientos legislativos que afectan a los datos que se gestionan; y, en consecuencia, de la propia estructura del Aula en lo referente a responsabilidades.

Por otro lado, considerando los intereses de la institución en el contexto social, la falta de normalización es un obstáculo para el reconocimiento de la labor de documentación y difusión desde un punto de vista oficial o gubernamental.

Dentro de estos mismos intereses sociales, una vez consensuados y definidos los protocolos de actuación, surge otra oportunidad: el fomento de la participación de investigadores y aficionados mediante la apertura controlada a los datos. Cabe aclarar que, en el Carnaval de Cádiz, la generación anual de obras lingüísticas y musicales excede las capacidades de catalogación del Aula, a la vez que existen miles de aficionados con presencia web que podrían participar del proceso documental en lugar de generar contenidos paralelos.

1.2.2 Específicos

El presente proyecto, considerado como subproyecto de la implantación del SI para el Centro de Documentación, tiene como finalidad la construcción de una base de datos que permita la gestión del inventario inmaterial custodiado.

El modelo mínimo a definir debe incluir, para cada elemento, todas sus características específicas más las recomendadas por la normativa archivística. Por otro lado, el diseño físico de la base de datos debe ser robusto, eficiente y escalable en cuanto al modelo de datos y al volumen de los mismos.

Paralelamente, se pretende implantar protocolos y procesos de trabajo correctos, completos y estrictos, acorde a los realizados por centros de documentación y archivos públicos que, en consecuencia, permita el trabajo colectivo de forma organizada y eficiente. Por último, aprovechando la oportunidad, reorganizar de forma coherente a las normativas vigentes los datos digitalizados y el multimedia –esencial en el inventario–. Respecto a esto habrá que valorar los niveles de descripción y los esquemas de signatura en uso.

1.3 Enfoque y método seguido

Considerando el presente proyecto como punto de partida para la propuesta general de implantación del "SI del Centro de Documentación del Carnaval de Cádiz", se propone para la gestión general las prácticas del marco PMBOK y para el desarrollo una metodología ágil, como Scrum[1]. Aunque, al ser un sistema totalmente nuevo a construir, y el alcance de este proyecto solo el modelo de BBDD mínimo –a ejecutar en una primera iteración–, se podría gestionar con metodologías en cascada; sin embargo, la complejidad del dominio, la diversidad de perfiles interesados y las altas tasas de fracaso en el diseño de *software* a medida que aún afectan a este mercado, han llevado a descartar las metodologías en cascada.

Se opta por las metodologías ágiles por su adaptabilidad a los cambios, a la vez que por la importancia que dan a la comunicación entre los implicados; por añadidura, conllevan el enfoque iterativo e incremental. Por un lado, se aumenta el compromiso y la participación y, por otro, facilitan la gestión de los requisitos y los riesgos; en definitiva, atendiendo a las características del proyecto, con objetivos claros pero requisitos extensos y aún desconocidos, los principios de desarrollo *lean* facilitan la gestión de proyecto y desarrollo, a la vez que promueven la implementación de productos totalmente funcionales por cada nueva fase o iteración.

1.4 Planificación del Trabajo

La estructura de distribución del trabajo está orientada a actividades[2], puesto que dentro de cada iteración del proceso de desarrollo del producto siempre se dan las mismas: requisitos, diseño, desarrollo y revisión. De hecho, tras cada revisión del estado por parte del Director del Proyecto, y tras cada reunión con el equipo de expertos del cliente, se plantea la actualización de la documentación de gestión y del producto; es decir, se repiten las actividades realizadas en distinto grado y, en consecuencia, se siguen las pautas de las metodologías ágiles.

Por otra parte, el proyecto será ejecutado por una única persona; por tanto, aunque se paralelizan las tareas en la medida de lo posible, no podrán ejecutarse simultáneamente en ninguno de los casos. Por todas estas cuestiones la gestión del tiempo tiene un enfoque pesimista.

En cuanto a recursos, se utilizará un ordenador personal, y dos más para la fase de pruebas; todo el *software* –ofimático, editores multimedia, IDE, etc.– y el SGBD son *open source,* más la aplicación que se utilizará como interfaz también; por último, y considerado el más importante, el equipo de expertos del Aula, del que solo se puede disponer durante 2 horas a la semana.

Se procede a comentar brevemente las tareas que implica cada una de las actividades, en el siguiente apartado se detalla el tiempo que han requerido.

---

[1] "*Ingeniería del software*". Pradel, J.; Raya, J.
[2] "*Gestión de proyectos*". Rodríguez, J. R.; Mariné, P.

1.4.1 Actividades

Planificación

- **Documentación y consulta.** Proceso continuo a lo largo del proyecto, desde consulta de fuentes externas hasta producción de documentos auxiliares (detalle de los diversos procesos, manuales de instalación, etc.).
- **Definición del alcance**
  - **Requisitos.** Análisis de necesidades, normativas, datos existentes y formas de actuación del Aula. Incluye las reuniones con el cliente y la documentación de los mismos.
  - **Priorización.** Evaluación y propuesta del orden de implantación considerando las dependencias y las limitaciones temporales.
- **EDT.** Plan de distribución del trabajo basada en actividades.
- **Planificación Temporal.** Atendiendo a la EDT se asignan días a las diversas tareas con un enfoque pesimista, se introduce paralelización atendiendo a las dependencias y se incluye como actividad la revisión y actualización tras las entregas parciales.
- **Plan de Gestión de Riesgos.** Conocidos los requisitos del producto y las tareas genéricas de gestión del proyecto y del desarrollo, se procede con el análisis y documentación inicial de los riesgos; en función de los resultados se define el plan de contingencias.
- **Plan de Gestión de Calidad.** Con base en los requisitos a implantar se definen los mínimos de calidad; luego, sobre estos mínimos, se diseñarán las baterías de pruebas.
- **Entrega parcial EP1.** Toda la documentación generada hasta el momento. Las actividades anteriores se consideran finalizadas.
- **Revisión global.** Según las recomendaciones del Director del Proyecto, respecto a los contenidos, extensión y estructura de la EP1, se revisa la documentación anterior.

Ejecución

- **Diseño**
  - **Conceptual.** Detalle y modelado del dominio del discurso mediante UML, incluido el esquema necesario para los estadísticos e históricos.
  - **Revisión de requisitos.** Se verifica con la Junta Directiva del Aula el modelo conceptual y se corrige en caso necesario.
  - **Lógico.** Traducción directa desde el modelo conceptual. Detalle de las decisiones de diseño, la normalización y las restricciones asociadas a cada relación.
  - **Físico** Descripción las tablas a construir, según las descritas por el modelo anterior, y procedimientos almacenados para actualización.

- **Entrega Parcial EP2.** Documentación actualizada y los modelos, al Director del Proyecto. Paralelamente, también se revisa con el equipo de trabajo del Aula y se atienden dudas y peticiones.
- **Revisión global.** Según las recomendaciones del Director del Proyecto, respecto a la EP2, se revisa la totalidad de la documentación y los modelos.

- **Construcción**
  - **Preparación del entorno de desarrollo.** Instalación del SGBD y del entorno de desarrollo.
  - **BBDD**
    - **Tablas.** Traducción del modelo físico al lenguaje SQL específico del SGBD.
    - **Procedimientos Almacenados.** Definición e implementación de los procedimientos almacenados.
    - **Consultas.** Se paraleliza con la anterior, con la misma pauta referente a las dependencias de estos mecanismos, a medida que se cuenta con las tablas se plantean las consultas.
    - ***Data warehouse***. Diseño lógico y físico, incluida la definición e implementación de los disparadores y funciones asociadas.
  - **Pruebas.** Esta actividad también se realiza junto a las tres anteriores, procedimientos, consultas y *data warehouse*, con objeto de construir una batería de pruebas lo más completa posible.
    - **Diseño.** En función a las pruebas generadas en dichas actividades, se seleccionará un conjunto reducido que fuerce el producto final.
    - **Automatización.** Creación de *scripts* basados en el diseño anterior.
    - **Optimización.** Según los resultados de las pruebas, revisión integral de los modelos, revisión de los índices y reajuste de vistas.
  - **Entrega parcial EP3.** Modelos revisados y código completo.
- **Revisión global.** Según las recomendaciones del Director del Proyecto respecto a la EP3, se revisa la totalidad del proyecto.
- **Entrega final.** Modelos y código revisados.

Cierre

- **Memoria.** Redacción de la memoria de trabajo, los anexos y el código, por: corrección lingüística, cohesión y consistencia en los contenidos.
- **Presentación audiovisual.** Exposición del proyecto y del producto mediante un documento multimedia. Planificación del contenido, grabación y edición.
- **Exposición.** Ante el Tribunal de Evaluación de la UOC.

1.4.2 Hitos

Tareas

WBS	Nombre	Inicio	Fin	Trabajo	Terminado
1	**Planificación**	**Mar 1**	**May 29**	**14d 5h**	
1.1	Documentación y consulta	Mar 1	May 29	11d 2h	100%
1.2	**Definición Alcance**	**Mar 1**	**Mar 8**	**1d 3h**	
1.2.1	Requisitos	Mar 1	Mar 8	1d	100%
1.2.2	Priorización	Mar 4	Mar 6	3h	100%
1.3	EDT	Mar 4	Mar 6	3h	100%
1.4	Planificación Temporal	Mar 5	Mar 7	3h	100%
1.5	Plan de Gestión de Riesgos	Mar 6	Mar 7	2h	100%
1.6	Plan de Gestión de Calidad	Mar 6	Mar 7	2h	100%
1.7	Entrega parcial EP1	Mar 8	Mar 8		
1.8	Revisión global	Mar 6	Mar 11	6h	100%
2	**Ejecución**	**Mar 9**	**May 22**	**20d 7h**	
2.1	**Diseño**	**Mar 9**	**Apr 14**	**7d**	
2.1.1	Conceptual	Mar 9	Mar 28	2d 4h	100%
2.1.2	Revisión Requisitos	Mar 20	Mar 29	1d 2h	100%
2.1.3	Lógico	Mar 22	Mar 31	1d 2h	100%
2.1.4	Físico	Mar 31	Apr 10	1d 3h	100%
2.1.5	Data Warehouse	Apr 10	Apr 14	5h	100%
2.2	Entrega parcial EP2	Apr 10	Apr 10		
2.3	Revisión global	Apr 19	Apr 22	4h	100%
2.4	**Construcción**	**Apr 2**	**May 22**	**13d 3h**	
2.4.1	Preparación del entorno de desarrollo	Apr 2	Apr 3	2h	100%
2.4.2	**BBDD**	**Apr 3**	**May 10**	**6d 2h**	
2.4.2.1	Tablas	Apr 3	Apr 12	1d 2h	100%
2.4.2.2	Procedimientos Almacenados	Apr 12	Apr 26	1d 7h	100%
2.4.2.3	Data Warehouse	Apr 24	May 8	1d 7h	100%
2.4.2.4	Consultas	May 1	May 10	1d 2h	100%
2.4.3	**Pruebas**	**Apr 17**	**May 22**	**5d 5h**	
2.4.3.1	Diseño	Apr 17	May 16	3d 6h	100%
2.4.3.2	Automatización	May 8	May 22	1d 7h	100%
2.4.4	Optimización	May 1	May 10	1d 2h	100%
2.5	Entrega final EP3	May 8	May 8		
3	**Cierre**	**May 8**	**Jun 9**	**4d 6h**	
3.1	Memoria	May 8	Jun 1	3d 1h	100%
3.2	Presentación audiovisual	Jun 2	Jun 9	1d	100%
3.3	Revisión global	May 29	Jun 2	5h	100%

Figura 1. Tabla cronológica de actividades

1.4.2 Diagrama Gantt

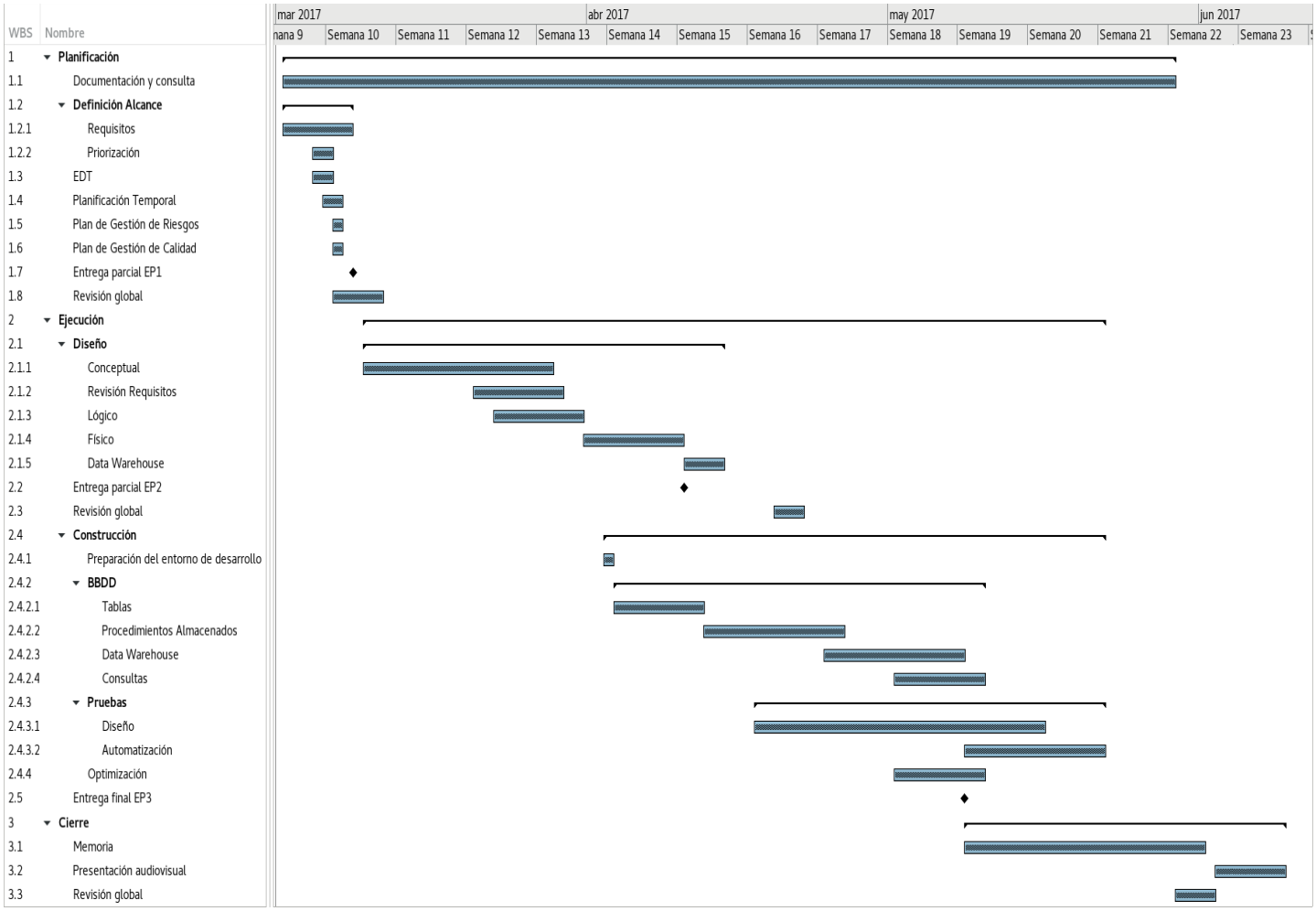

Figura 2. Diagrama Gantt

1.5 Breve sumario de productos obtenidos

Se detallan a continuación los entregables parciales, o intermedios, identificados como EP y, seguidamente, se describe el que se considera producto final.

- **EP1: Planificación del proyecto**. Plan de gestión del proyecto y desarrollo del producto, en contenido y tiempo, una vez los requisitos han sido consensuados con la Junta Directiva del Proyecto, donde participan el personal del Aula de Cultura del Carnaval de Cádiz y los expertos en archivística.
- **EP2: Diseños de la BBDD**. Documentación actualizada y modelos definidos en los niveles: conceptual, lógico y físico. De nuevo, una vez que han sido expuestos y verificados en detalle con la Junta Directiva del Proyecto.
- **EP3: BBDD**. Documentación relativa a la construcción y las pruebas de eficiencia y rendimiento, incluye detalle de las tareas de optimización y resultados de las pruebas.

Se considera producto final al conjunto completo, que incluye:

- Los *scripts* necesarios para construir la estructura de la base de datos y las funcionalidades *data warehouse*; las funciones que gestionan las reglas de integridad de tipo usuario; todos los procedimientos almacenados necesarios para la gestión de los datos (alta, baja y modificación); y, por último, las vistas necesarias para acceder a la totalidad de los contenidos de forma clara.
- La documentación de soporte utilizada a lo largo del proceso, la memoria, normativas y anexos, donde se detalla todo el proceso de integración de los conocimientos y las decisiones de más bajo nivel.

1.5 Breve descripción de los capítulos de la memoria

El siguiente capítulo se dedica a la gestión del proyecto, en el se detallan en cierta medida las actividades realizadas a este respecto: alcance, requisitos y riesgos, además de los planes de comunicación y calidad, por la importancia que se da estos dos factores. Dicho capítulo, y este primero, por el carácter iterativo de la gestión y del desarrollo, son los más revisados y actualizados a lo largo del proyecto; mientras que, los siguientes, describen en cierto modo lo que sería cada fase de desarrollo que construye la BBDD.

El tercer capítulo, tal como se acaba de exponer, detalla el primer proceso de modelado; concretamente, incluye la descripción de las entidades y relaciones, más la representación diagramática en UML.

Seguidamente, en el cuarto y quinto capítulo se desarrollan los modelos lógico y físico. En estos se describen todas las estructuras, así como las implicaciones y decisiones de diseño de sendos niveles de abstracción; por ejemplo, el grado de generalización de las entidades, el SGBD elegido, los componentes de control utilizados, etc. Por otro lado, se incluye en estos el detalle de las tareas de optimización, que realmente se realizan luego.

El sexto capítulo se centra en los componentes *data warehouse*, en como estas funcionalidades se alimentan del inventario, de nuevo respecto al modelo y las decisiones. Por el contrario que en los capítulos anteriores, se describe un producto que se desarrolla paralelamente a la BBDD; es decir, aun no detallando estas funcionalidades en los dos capítulos anteriores, se desarrollan dentro de las mismas actividades de diseño de forma conjunta, solo que se documentan de forma independiente.

El séptimo capítulo está dedicado a la fase de pruebas, desde el diseño y los resultados hasta la implementación de los *scripts* que las automatizan. De nuevo, en este caso, se documenta un proceso que se lleva a cabo desde el diseño físico.

Finalmente, los capítulos 8, 9, 10 y 11 se dedican a información general: conclusiones, glosario, bibliografía y anexos.

2. Gestión del Proyecto

2.1 Alcance

El objeto del presente proyecto es la construcción de la base de datos que necesitará el futuro SI del Centro de Documentación del Carnaval de Cádiz. Dentro del alcance de la BBDD está únicamente la funcionalidad de inventario inmaterial, o documental; concretamente, de los conceptos esenciales asociados a las agrupaciones, su tipo, las letras, los involucrados y los premios. Por tanto, se consideran suficientes los siguientes conceptos generales: obras, personas, documentos, eventos, agrupaciones y empresas.

2.1.1 Incluye

- Análisis del estado actual de los datos que posee el Aula y sus socios, con enfoque tanto legal como estructural. Con objetivo de introducir las las normativas actuales de documentación archivística, concretamente se introducirá el uso de las recomendaciones de la ISAD(G), *International Standard of Archival Description (General)*[4].
- Diseño del modelo de datos conforme al dominio del discurso y a las normativas documentales; posteriormente la definición de los modelos lógico y físico, incluidas todas las funciones para gestión de la integridad y el encapsulamiento de la BBDD respecto de aplicaciones externas.
- Construcción de la base de datos, definición de consultas y pruebas.
- Informes generados acerca del modelo de datos y el proceso.

2.2.2 No incluye

- El diseño y la construcción de la interfaz de la BBDD.
- La formación requerida para su uso sin interfaz.
- La migración de los datos existentes.
- Ni todas las funcionalidades que se describen en el siguiente apartado.

2.2.3 Líneas futuras

Se enumeran, priorizadas, las funcionalidades a introducir en las próximas fases del proyecto general del SI para el Centro de Documentación del Carnaval de Cádiz.

Migración de datos

Incorporación de los registros existentes a la estructura de datos de forma automatizada, dentro de lo posible. Esta actividad deberá de producirse más veces, ya que se cuenta con datos que no se incluyen en este primer modelo.

Interfaz web

Una vez rescatados los datos existentes, resulta imprescindible para poner en producción la base de datos. Se propone que la funcionalidad de control de integridad entre las referencias contenidas en la BBDD hacia la estructura de

[4] "*General International Standard Archival Description*". Consejo Internacional de Archivos

directorios sea implementada junto con la interfaz –los documentos y el material multimedia están ya organizados según instrumentos de descripción archivística adecuados, pero es necesario controlar la consistencia entre espacio de datos y referencias a los mismos–.

Legislación

Análisis de los marcos legales que afectan a los datos contenidos en el sistema e introducción de la gestión de los mismos, en principio la Ley Orgánica de Protección de Datos y la Ley de Propiedad Intelectual –la cual se aplica en este Carnaval de forma, aveces, implícita debido a la singularidad del mismo–. Esta línea de trabajo también debe darse por cada nueva funcionalidad a introducir.

Usuarios

Como objetivo general del proyecto de implantación de un SI de gestión documental se hace necesaria la gestión de usuarios y perfiles, que permita la organización global de las diversas actividades. En este aspecto habrá que incluir la planificación de las vías y pautas que permitan la participación de personas ajenas al Aula; así como utilidades de recuperación y auditoría.

Ampliación del modelo

Dentro de los conceptos ya descritos por el modelo resultante de este proyecto restan cuestiones a incluir, omitidas en el primer diseño, cabría finalizarlo completamente antes de proceder con la implementación de otros módulos independientes, como el siguiente.

Biblioteca y Hemeroteca

Son inventarios que ya gestiona el Aula, pero su prioridad se considera inferior a la puesta en producción y la aplicación de las pautas legales que afectan a la BBDD; no obstante, su desarrollo se puede paralelizar, sobretodo porque estas últimas son mayormente administrativas.

Difusión

Entre los objetivos principales del Aula, y del resto de instituciones participantes, está la difusión del patrimonio cultural del Carnaval. Entre las actividades de difusión cabe resaltar la gestión de exposiciones, el préstamo de fondos y la promoción de obras.

Conservación

De manera natural, tras tener el inventario regularizado por las leyes necesarias y documentado de forma estandarizada por los perfiles técnicos adecuados, se deben incluir al sistema las facilidades necesarias para gestionar las tareas de conservación que requieren los fondos materiales.

Estas dos últimas líneas de expansión del SI tienen la misma prioridad, habrá que valorar el orden llegado el momento.

Comunicaciones

El Aula, como se ha ido exponiendo, forma parte de un conjunto de entes participantes en estas fiestas; incluso, en muchos casos, se realizan procesos administrativos y documentales conjuntos. Por otro lado, mantiene relaciones con otras entidades relacionadas con carnavales del resto del mundo.

Respecto a las actividades donde existen convenios de cooperación, sería de gran interés estudiar las posibilidades de automatización; por ejemplo, la regulación de los intercambios o la organización de las tareas colaborativas.

2.2 Requisitos

La definición de requisitos se realiza mediante reuniones con los expertos en Carnaval de Cádiz y expertos en Ciencias de la Información; y, por otro lado, también de gran importancia, mediante la observación[1] de los datos y los fondos con que cuenta el Aula. A continuación se exponen las necesidades a satisfacer ya priorizadas, agrupadas en requisitos funcionales y técnicos[1], o no funcionales, dentro de los contemplados en el ámbito de las BBDD.

2.2.1 Requisitos funcionales

RF01 Gestión del inventario inmaterial del Centro de Documentación, centrado en las agrupaciones y las obras que interpretan, considerando todas las información que obtiene el Aula de forma oficial o directa.

RF01.1 **Obras**. Cualquier tipo de producción inmaterial, como las lingüísticas, musicales y artes plásticas. Serían las letras, partituras, bocetos y diseños de maquillaje. Estos tres últimos no se encuentran en formato texto, solo de imagen.

RF01.2 **Agrupaciones**. Se considera una por año porque cada año cambian varios factores. Se clasifican en modalidades (lo que afecta a la obra que interpretan de forma directa).

RF01.3 **Personas**. Serán descritas por los datos comunes de descripción de personas. La importancia radica en los roles que cumplen dentro de las agrupaciones o empresas, ya que pueden ser cantantes, músicos, artesanos, etc.

RF01.4 **Documentos**. Físicos o digitales; en principio, como mínimo, se solicita implementar la gestión de documentación genérica, el material gráfico, la fonoteca y la videoteca.

RF01.5 **Eventos**. Inicialmente solo se contemplarán los concursos oficiales con las cuantías por premio, teniendo en cuenta que cada año cambia la configuración de los premios. No obstante, cabría incluir la descripción de otros como cabalgatas, homenajes, premios, pregones, etc.

RF01.6 **Entidades.** Desde instituciones organizadoras hasta empresas (artesanos, notarías, productoras, etc.).

Inicialmente se introduce este concepto para describir a empresas, concretamente por la necesidad de vincular los contratos de artesanía y maquillaje a las agrupaciones.

[1] "*Ingeniería del software*". Pradel, J.; Raya, J.

RF01.7 Gestión de la veracidad de los contenidos.

Como parte del proceso documental, donde las fuentes son diversas y su credibilidad arbitraria, se gestiona información a cerca de la información registrada; es decir, ciertos controles donde se detallan las fuentes y observaciones respecto del proceso documental en sí.

RF02 Implementación de todas las relaciones posibles entre los conceptos anteriores, más todos los condicionantes que implican.

RF02.1 **Obras – Personas**, describe la relación de autoría, en ocasiones pueden ser anónimos o desconocidos.

RF02.2 **Obras – Agrupaciones**, relación de interpretación.

RF02.3 **Obras – Documentos**, letras, partituras y otros diseños digitalizados; además, todo el material multimedia asociado.

RF02.4 **Agrupaciones – Personas**, con diferentes roles e intereses, tanto musicales como legales y económicos. Una persona puede cumplir más de un rol por agrupación o agrupaciones.

RF02.5 **Agrupaciones – Documentos**, libretos digitalizados y todo el material multimedia asociado a las agrupaciones.

RF02.6 **Agrupaciones – Eventos**, se entiende la participación por la existencia del vínculo y se debe registrar la obtención de premios si gana alguno. Dentro del alcance del presente proyecto se considerarán exclusivamente las peculiaridades de los concursos.

RF02.7 **Agrupaciones – Entidades,** roles de artesanía y maquillaje.

RF02.8 **Personas – Documentos**, en el caso de las autorías.

Relaciones fuera del alcance del modelo:

- Personas a eventos y a otras entidades, con roles de participación, afectación o beneficio.
- Documentos a entidades, considerando el enfoque administrativo. Un buen ejemplo es el fallo del jurado, donde la entidad sería el Patronato del COAC, incluso el jurado por ser anual, y como evento el COAC.
- Entidades a eventos, para roles de organización, relaciones administrativas o por ser afectados.

RF03 Introducción de recomendaciones de descripción indicadas por las normativas de documentación de obras, documentos y materiales.

RF03.1 ISAD(G), de reconocimiento internacional, para la descripción de documentos y obras inmateriales, así como de fondos y colecciones.

RF04 Consulta e interpretación. Acceso a la información de forma directa según los conceptos de interés definidos en el requisito RF01.

- **RF04.1** Listado de agrupaciones ordenadas por año y modalidad, incluidos todos los datos asociados y los premios, o el orden.
- **RF04.2** Listado de documentos multimedia asociados a agrupaciones, ordenadas por año y modalidad, incluida la autoría y la dirección del grupo.
- **RF04.3** Listado de personas relacionadas a agrupaciones, ordenadas por año, modalidad y agrupación, centrada en los roles.
- **RF04.4** Listado de eventos y características de los mismos, en el caso de los concursos también el detalle de los premiados.
- **RF04.5** Listado de obras relacionadas a autores y agrupaciones, ordenadas por año y modalidad, premios y referencias a documentos (físicos y digitales).
- **RF04.6** Listado de documentos, textuales, gráficos y multimedia ordenados por año y tipo.
- **RF04.7** Tipos de documentos que no se han registrado e información que no ha sido verificada.

RF05 Gestión de históricos y consultas precalculadas.

- **RF05.1** Agrupaciones ganadoras por concurso y año.
- **RF05.2** Premios por agrupación, persona, empresa.
- **RF05.3** Dedicación al Carnaval de forma oficial por persona.
- **RF05.4** Temas recurrentes en letras por año y modalidad.

2.2.2 Requisitos no funcionales

RNF01 Rendimiento

- **RNF01.1** El factor de mayor criticidad es el volumen de datos, concretamente por las frecuentes búsquedas sobre campos de texto de longitud indeterminada –la transcripción de las letras–. Debería soportar búsquedas por texto libre en dichos campos para más de 20000 registros.
- **RNF01.2** Respecto a la concurrencia y el tiempo de respuesta se prevé un mayor nivel de accesos en la época de carnaval; no obstante, para esta instancia inicial de la BBDD no se plantea el uso público, o remoto, de la misma.

RNF02 Restricciones de diseño.

- **RNF02.1** Se limita el desarrollo al uso productos de *open source*, tanto para la BBDD como para el entorno de producción.

RNF02.2 El modelo resultante debe alcanzar el mayor nivel de normalización posible, con objeto de permitir cualquier tipo de expansión, eliminar redundancias y, con esto, anomalías durante la gestión de datos.

RNF03 Robustez y operatividad, mantenibilidad y portabilidad.

RNF03.1 Se proveerá de robustez y consistencia mediante la encapsulación del modelo tras los procedimientos almacenados necesarios para las acciones de actualización de la totalidad de los contenidos. Solo se permitirá el uso de SQL desde aplicaciones externas para consultas y creación de vistas.

RNF03.2 Toda operación de actualización debe proveer de un resultado, ya sea un conjunto de datos, un mensaje de éxito o de error lo suficientemente explícito.

RNF03.3 El SGBD elegido, a pesar de la restricción del RNF02.1, debe poder ejecutarse en cualquier sistema operativo. El conflicto se plantea por el juego de caracteres únicamente, por tanto se acuerda usar UTF-8.

RNF04 Seguridad

RNF04.1 El control de acceso, en primera instancia, solo requiere de dos perfiles: el administrador y el editor de contenidos. Dentro de este proyecto no se consideran los controles ni perfiles que impondrá la futura consideración de la LOPD.

RNF04.2 Las restricciones de integridad serán controladas por los mecanismos del SGBD; excepto en los casos que son de tipo semántico, o de usuario, donde se controlará mediante disparadores y los procedimientos almacenados de alta, baja y modificación.

RNF04.3 La privacidad solo es requerida para los datos amparados por la LOPD; sin embargo, por el carácter de lanzamiento de un proyecto mayor que tiene el presente, se omite la gestión de las prerrogativas de esta ley orgánica. Aun así, se recomienda que el producto una vez poblado sea utilizado por el menor número de personas posible.

RNF04.4 La corrección de los datos se controlará mediante la imposición de utilizar unicamente los procedimientos de almacenado para la actualización, según el RNF03.1.

La disponibilidad no entra en el alcance del presente proyecto. Se considera la planificación de copias de seguridad como algo sumamente sencillo, a la vez que insuficiente, si no se monitorizan los servidores. También queda fuera del alcance la gestión burocrática relativa a la LOPD.

2.3 Riesgos

Por las razones ya expuestas se utilizarán al uso de metodologías ágiles; dentro del ámbito de la gestión de riesgos, permiten una temprana detección, así como a la reducción del impacto de los riesgos en gran medida. Por su componente iterativo, por la importancia que dan al seguimiento del proyecto y por los frecuentes procesos comunicativos que recomiendan, facilitan la adaptación de los planes ante cambios o desviaciones inesperadas.

En primer lugar, se considerá como factor principal del éxito la participación del mayor número de especialistas posibles, tanto específicos del Carnaval como de perfil archivístico; para, de esta forma, obtener una descripción completa y detallada de las entidades mínimas a incluir en el modelo de datos, además de todas las posibles relaciones y sus condiciones.

En segundo lugar, los factores del éxito relativos a la gestión de proyectos:

- Control exhaustivo del alcance semántico del modelo, considerado mínimo, con objeto de finalizar el producto con la mayor calidad posible.
- Seguimiento del compromiso por parte del cliente, que suele resultar en un producto no deseado y, en consecuencia, no utilizado.
- Control del consenso en la especificación de requisitos, tanto por deficiencias como por inconsistencias en los procesos comunicativos.
- Y, principalmente, manejar la resistencia al cambio que pueda generar la carga de trabajo que introducirá el nuevo modelo de datos.

Finalmente, a nivel técnico, la aplicación estricta de reglas de diseño asociadas al modelado de datos y de BBDD.

2.3.1 Análisis

R01 Excesos en el alcance semántico del modelo básico de inventario; así como defectos del modelo a este respecto.

R02 Falta de compromiso y retrasos provocados por los procesos de coordinación, negociación y consenso.

R03 Imprevistos personales, por motivos laborales o familiares.

R04 Fallos en la infraestructura.

R05 Defectos en el diseño de alto nivel.

2.3.2 Plan de Contingencias

CR01 Reuniones periódicas con la Junta Directiva del Proyecto para intercambio de información y validación por parte del Director del Proyecto.

CR02.1 La falta de implicación por parte del cliente queda controlada a partir de los procesos comunicativos; lo que, implícitamente, aumenta el componente participativo y decisional.

CR02.2 Los retrasos no pueden evitarse; pero, la paralelización de actividades permitirá que siempre haya trabajo por realizar.

CR03 La jornada laboral es de 4 horas, pudiendo duplicarse los fines de semana. La planificación del diagrama Gantt implica una hora de trabajo por actividad y día; es decir, por cada actividad paralelizada se tiene una hora, lo que resulta en una media de 3 horas diarias.

En la fase de construcción se duplica el número de horas por día, pero también se duplica el número de horas necesarias por cada actividad. En definitiva, se planifica de forma totalmente pesimista; no obstante, se consideran jornadas cortas según el tiempo disponible.

CR04.1 Respecto al entorno de desarrollo, se utilizará la virtualización para el SGBD y el IDE; de esta forma, restablecerlo no llevará más de 10 minutos.

CR04.2 Se programa una de copia de seguridad en una unidad de disco externa, mediante Cron y Rsync. Cada hora se sincroniza el directorio de trabajo con una réplica en dicha unidad y se almacenan comprimidos del directorio de datos del SGBD.

La recuperación de la normalidad, en el peor de los casos, implica ubicar dichos datos en el lugar correspondiente, y relanzar nuevas instancias de los objetos virtualizados.

CR05.1 Se introducen horas de más en los procesos de diseño para verificar los modelos al detalle. Además, en cada nueva fase de diseño se plantean horas de revisión de los diseños de más alto nivel, ya que cada nivel afecta a los previos.

CR05.2 Se introducen horas dedicadas a documentación que deben utilizarse para hacer un buen uso de las capacidades del SGBD.

2.4 Calidad

El control de calidad, por un lado, tiene como base de evaluación concreta cada uno los requisitos funcionales y técnicos, detallados en el apartado “2.2 Requisitos” de este capítulo; y, por otro, en el control y seguimiento de los riesgos identificados en el apartado anterior, como cuestiones que prevenir y a las que anticiparse. Es decir, se plantea la lista de requisitos a modo de *checklist* sobre el producto y los riesgos como circunstancias a prevenir.

A nivel técnico, desde el inicio se toman las decisiones de forma que satisfagan los requisitos no funcionales. A nivel funcional, se ejecuta el proyecto atendiendo a la priorización de todos los requisitos.

De la misma manera que la mayoría de procesos de gestión del proyecto, el control de la calidad se hace de forma iterativa e incremental; es decir, a medida que evoluciona la documentación, el diseño y el desarrollo del producto se realizan las verificaciones, y correcciones, pertinentes. Dentro de los dos enfoques, el técnico y el funcional, resulta de interés el seguimiento constante, ya que permite una adaptación gradual a las necesidades que no hayan sido detectadas; por tanto, facilita la gestión de requisitos y mantiene la consistencia entre el conjunto de procesos involucrados en la ejecución del proyecto.

Por otro lado, a este proceso de gestión se adscribe la tarea de mantener la documentación actualizada y coherente al proyecto y al producto.

2.5 Comunicación

Los procesos comunicativos son parte esencial del éxito del presente proyecto. Resulta vital para fortalecer el compromiso del equipo, fomentar la cooperación y la participación; y, principalmente, para que ganen consciencia de la necesidad y el beneficio que aportará el producto a la labor del Aula de Cultura del Carnaval de Cádiz.

Por otro lado, aunque se parte de un modelo de datos muy sencillo, hay una alta complejidad en cuanto a las propiedades de las entidades y en las relaciones entre las mismas; por tanto, es necesaria la participación de los expertos en la materia, principalmente durante el modelado del dominio de interés –más aun considerando las ideas preconcebidas que se puedan tener–.

Con todos estos objetivos como principios de actuación, se realizan encuentros semanalmente con los interesados durante el diseño conceptual, luego las reuniones se reducen a las estrictamente necesarias para la resolución de dudas y el fomento de la participación. Por otro lado, se cuenta con el contacto directo del identificado como patrocinador del proyecto, este socio solventa dudas puntuales y concretas a diario.

Para las reuniones se prepara el material generado y resúmenes de las tareas a realizar durante el periodo que transcurrirá hasta la próxima reunión. Las pautas generales son:

- Detallar claramente el estado del proyecto, mediante la exposición de los modelos, y de forma acorde al registro de los diversos interlocutores[1].
- Atender todas las posibles dudas hasta su total resolución.
- Consensuar las decisiones de diseño.
- Promover la participación colectiva como método de incorporación de conocimiento constante; por ejemplo, mediante mecanismos como el *brainstorming*, el análisis colaborativo de los datos y la valoración de las diversas opiniones personales, etc.
- Más, como parte del proceso de gestión del alcance, y en defensa de la finalidad del proyecto, se debe controlar continuamente la percepción del cliente al respecto, tanto de los beneficios como de las limitaciones que este considera.

[1] “*Ingeniería del software*”. Pradel, J.; Raya, J.

3. Diseño conceptual

Para modelizar el contexto de acción, o mundo real de interés, se ha realizado un proceso de descripción descendente[3], desde los objetos principales del dominio: personas, agrupaciones, obras, concursos y documentos. A partir de ellos se han realizado diversos refinamientos atendiendo a: las disertaciones compartidas con el cliente, las posibilidades de generalización y la complejidad de las interacciones entre los objetos a conceptualizar. A lo largo de todo este proceso se ha ido delimitando el alcance semántico del modelo, excluyendo toda particularidad considerada secundaria.

De nuevo, atendiendo a las premisas generales de gestión del proyecto, por cada iteración se han verificado las peculiaridades bajo diversos enfoques, con objeto de integrar los diferentes marcos de referencia, o concepciones del mismo mundo, que tienen los usuarios finales del SI.

El diseño que se describe, y muestra, a continuación es el obtenido una vez aplicadas las correcciones relativas a trampas comunes de diseño (abanico, corte, pérdida de afiliación, aridad y de semántica)[3]. Todos los nombres siguen las convenciones de nomenclatura Pascal y Camel[1], según proceda, y no se hace uso de caracteres fuera del código ASCII; además, en caso de utilizar nombres de entidades o atributos a lo largo de la descripción los tres modelos, se marcan mediante subrayado. Por otro lado, no se describen en detalle los atributos ni multiplicidades, esta información puede extraerse del diagrama de clases y, posteriormente, a nivel de atributo, del modelo lógico.

Finalmente, cabe puntualizar ciertos factores específicos del dominio que afectan al diseño global, haciendo necesaria la introducción de un número elevado de atributos nulos. Por ejemplo: la flexibilidad que requiere la tipología de las agrupaciones, de las callejeras solo se conoce el nombre y la obra literaria –si ha sido adquirida–; la incertidumbre o la falta de veracidad, de la información histórica; y, por el mero hecho de intentar modelar el arte.

3.1 Entidades

En este apartado se satisfacen todos los requisitos funcionales enumerados en RF01, centrados en los objetos del dominio. Antes de proceder con la descripción se deben aclarar ciertas decisiones, puesto que podrían llevar a confusión en cuanto a la terminología utilizada en dicho requisito; concretamente, se ha marcado la diferencia entre ciertos conceptos y se han propuesto ciertos vínculos que, tras las fases del diseño, han sido alterados. En algunos casos por restricciones de diseño y en otros por redundancia.

- En los requisitos se trata a las personas y entidades como diferentes, se considera a estas últimas como empresas, instituciones u otros entes –aunque el presente proyecto solo contemple las empresas–. En el modelo resultante, por la necesidad de asociar a empresas, personas y agrupaciones de forma indistinta a otros conceptos (como las obras, por interpretación, autoría o creación) se abstraen todas ellas a nivel de entidad; por tanto, desde este punto del proyecto, se consideran que tanto las empresas como las personas y agrupaciones son entidades.

---

[1] "*Ingeniería del software*". Pradel, J.; Raya, J.

[3] "*Diseño de bases de datos*". Burgués, X.; Cabré i Segarra, B.; Casas, J.; Costal, D.; Juanola, P.; Ortego, S.; Pérez, J.; Rius, À.; Segret, R.

- Según lo recién detallado y atendiendo al caracter anual que se confería a las agrupaciones en los requisitos, durante el modelado conceptual surgía la necesidad de considerar a las agrupaciones como entes permanentes a la vez que anuales; de nuevo, por existir la posibilidad de autoría por parte de una agrupación, se hace necesaria la distinción. Por añadidura, el hecho de diferenciar entre las agrupaciones y sus propuestas anuales permite satisfacer otro requisito: listar el historial de cada agrupación en el Carnaval de Cádiz por separado.

El resto de conflictos de significado entre los requisitos y el modelo definido incumbe a las relaciones, por ello se expondrán en el siguiente apartado. Además, en el Anexo I se tiene la descripción detallada de todos los conceptos junto con los diversos argumentos asociados a las decisiones de este diseño.

A continuación se procede con las entidades, empezando por las cinco consideradas auxiliares o de soporte; para cada una se da una breve descripción y el detalle de las implicaciones que se han considerado complejas en el modelado.

Modalidad

El concepto de modalidad es fundamental en la clasificación de agrupaciones y, por extensión, de las obras; principalmente son: Chirigota, Comparsa, Cuarteto, Coro y Romancero. Únicamente contiene un atributo, que inicialmente se modelizó como atributo de AgrupacionAnio en forma de enumeración, por ser conjuntos de valores ciertamente estáticos y limitados; esta decisión se toma respecto a los tres siguientes, así mismo, para todos estos la clave candidata es dicho atributo.

Por otra parte, en este caso y en Instrumento, la extracción del atributo se hizo necesaria por su uso desde otra entidad; aun así, la mayor ventaja de ambos es que permitirán utilizar los valores contenidos a través de identificadores –los cuales serán introducidos en el modelo físico de la BBDD–. De esta manera, se evitan errores tipográficos, que puedan generar duplicidades, y se reduce enormemente el espacio utilizado en los atributos modalidad e instrumento de las dos entidades que los utilizan.

Instrumento

Respecto al modelado de la lista de valores que representa esta entidad como enumeración, además, suponía un obstáculo a la diversidad que se está introduciendo en los últimos años en cuanto al uso de intrumentos musicales.

Actividad

De nuevo con un solo atributo y la misma evolución y razones descritas para los anteriores, en cuanto a reducción de espacio y de errores tipográficos. Por otro lado, esta solo será utilizada por las empresas, las cuales pueden desarrollar más de una actividad; por tanto, respecto a la entidad Empresa esta característica es multivaluada, aun síendo el conjunto de valores más invariable que el anterior.

Descriptor

A diferencia del tipo de documento, que es de nivel muy general, y del tipo de soporte, que se refiere al contenedor físico de un documento en concreto, se

tienen los descriptores. Este concepto describe el contenido de los documentos de forma más específica y centrada en la estructura lógica de los contenidos; por ejemplo, serían valores como: libreto, entrevista, ensayo, antología, reglamento, etc. De nuevo, por la flexibilidad del dominio de este concepto, se propone como lista abierta mediante esta entidad a modo de tabla de *lookup*. Además de todas las ventajas anteriores, en este caso tener un identificador numérico aportará eficiencia en las búsquedas y ordenaciones en Documento.

Tema

Como concepto será utilizado igual que los descriptores para letras y tipos –descritos mediante la obra gráfica utilizada por las agrupaciones–, a modo de etiquetas. En este caso, aunque la funcionalidad es similar a las recién detalladas, se necesitan dos atributos: el tema en si y su descripción, necesaria para los muy específicos, obsoletos o históricos; siendo el primero la clave candidata. Por otro lado, la lista de valores en esta ocasión es totalmente abierta, ya que las temáticas surgen a la par que los tiempos.

GestionInformacion

Mediante esta entidad y la siguiente se satisface el requisito dedicado al soporte del proceso documental, el identificado como RF04.7. Contiene atributos booleanos que permiten indicar el estado de verificación, más la vía de obtención, de la información relativa a una agrupación determinada; por ejemplo, si se ha registrado a partir de datos en prensa o de documentos oficiales. Si la información está marcada con estas señales tendrá, además, indicadas las fuentes concretas y alguna observación si el archivero lo estima necesario. Por su semántica se modela como entidad débil de las agrupaciones de nivel anual, con dependencia de existencia y cardinalidad 1:1; por tanto, su clave candidata es la que conforma su vínculo con AgrupacionAnio.

GestionDocumento

Similar a la entidad anterior en objetivo y mecanismos de diseño, pero centrada en los documentos. En este caso las marcas refieren, más que a la veracidad, al estado de descripción, existencia, conservación y protección de los mismos.

Entidad

Se modelan los conceptos de persona, agrupación y empresa mediante el uso de una generalización disyunta y completa, con varios objetivos. Por un lado, cualquiera de estos tres conceptos puede participar de una relación de autoría –aún considerando este modelo de datos como básico–; y, por otro, se plantea como mecanismo para gestionar que haya un identificador único en el dominio por cada entidad, del tipo que sea.

En cuanto a la costumbre de abstraer los atributos comunes de las subclases, que además se modela como abstracta (no instanciable), se desestima. La abstracción de alguno de los atributos podría generar utilidad funcional, pero no resulta muy coherente separar el nombre de los apellidos u otros conjuntos que incluso forman parte de la clave candidata de las subclases.

Persona

Primer caso de concepto que especializa a Entidad. En este caso los atributos son los descriptores comunes, más la historia; este último se deduce del uso

que hace el cliente del campo observaciones en su origen. La complejidad en este caso deviene del como establecer las claves candidatas, se tienen las siguientes condiciones: el DNI admite valores nulos, porque antes de 1984 no era un requisito, y solo se utilizaba un apellido en las inscripciones; en consecuencia, no se puede establecer como clave candidata ni el DNI ni la tripleta compuesta de nombre y dos apellidos; por otro lado, ante estas ausencias, se requiere la introducción del domicilio como forma de identificación unívoca.

Una información de especial interés en relación a las personas es el nivel de dedicación o la trayectoria, sobre todo en el caso de los autores; sin embargo, por tener un carácter evolutivo, se tratará dentro de la funcionalidad de *data warehouse.*

Empresa

Segunda subclase de Entidad, también con datos comunes; en este caso los mínimos necesarios para cubrir el alcance, centrado en el rol de creadores de obras gráficas (disfraces, atrezos y forillos). Se modelan como obligatorios los estrictamente necesarios para identificarlas de forma única, se decide en este caso el conjunto nombre, localidad y provincia, con objeto de evitar coincidencias en el primero. Esta decisión se toma de nuevo por el impacto de la incertidumbre, la cual aquí es aún mayor que en el resto porque que estas entidades no aportan sus datos de forma oficial; de hecho, por esta razón el CIF no es obligatorio, aunque pueda ser clave candidata.

Una decisión de diseño, que afecta también a la entidad anterior, es la multiplicidad que aceptará el valor del atributo telefono; se decide no generar una relación de tipo 1:M para modelarlo como multivaluado, por considerarlo un dato totalmente accesorio, afectaría al rendimiento general sin ser de interés en si mismo. Se da una circunstancia similar en caso del contacto, modelable mediante una relación con instancias de tipo persona; también por su carácter secundario y la falta de rigurosidad de la característica se desestima.

Agrupacion

Última subclase de Entidad, en este caso se describen las agrupaciones carnavalescas con enfoque temporal indefinido, teniendo en cuenta a los grupos de personas que cada año interpretan determinada obra. En este nivel, sin embargo, no se alude a los componentes sino a la entidad, propiamente dicha, y a la procedencia geográfica de las mismas, que no varian en el tiempo.

En cuanto a las características, solo es obligatoria la procedencia. Luego, se tiene alias en lugar de nombre, porque realmente no tienen un nombre como agrupación; de hecho, no es obligatorio por que este identificativo lo da el tiempo y el público. Por último el contacto, en este nivel no es posible asignarlo, ni reasignarlo, en función del rol de representante legal; aveces no existe el rol y, otras, puede ser el autor o cualquier integrante de larga duración.

Con vistas a la futura interfaz y sus posibilidades, para todos los casos como el recién comentado, se plantea que existan facilidades que permitan acceder a información vinculante acerca de los atributos que podrían ser deribados. Más adelante se dan varios conceptos susceptibles a la derivación, tanto precalculados como por las relaciones entre entidades.

AgrupacionAnio

Mediante una relación de composición esta entidad describe la prupuesta anual de una agrupación concreta. A este concepto se le asocian personas, mediante roles y las obras que interpretan; luego, tienen características como la categoría (determinada por la edad de los integrantes), la modalidad, la composición, el lugar de ensayo o el presupuesto; todos, naturalmente, dentro del enfoque anual. Por otro lado, debido a la incertidumbre que afecta al dominio, algunos atributos, como el total de componentes, no pueden ser derivados –en este, concretamente, se podría controlar como valor mínimo el total de personas distintas vinculadas mediante instancias de rol en que se alude a una agrupación–.

Por último, en cuanto a la clave candidata, se da un escenario que obliga a introducir el nombre de la agrupación como parte de la clave: hay ocasiones en que un grupo de personas completo interpretan dos obras con distinto nombre –por regla general, el nombre de la obra coincide con el de la propuesta de la agrupación en determinado año–, de no ser por esto el identificador de la agrupación que la presenta más el año serían clave candidata de esta entidad.

Obra

Este es un nuevo caso de generalización, al igual que Entidad, es completa, disyunta y abstracta, donde la entidad raíz no contiene más que el identificador y un atributo discriminante. También es esta ocasión, por la necesidad de relacionarlas en un nivel abstracto, en este caso con Documento, y por la utilidad de identificarlas unívocamente sean del tipo que sean. En este caso, de tener vínculados los documentos a los distintos subtipos de Obra, se tendrían muchas más relaciones de tipo N:M intermedias, lo que complicaría enormemente la gestión de las relaciones y las consultas; cuando, además, esta es la ventaja de tener el identificador único a nivel de Obra.

Otra decisión, a nivel operativo, es tener una documento mínimo por cada obra; esto se basa en la necesidad de describir las obras con enfoque de documento según las recomendaciones de la norma ISAD(G)[4]. Atendiendo al modelo de datos y al uso práctico, se considera que al crear una obra quizás no se cuente con documentación tangible dentro del nivel de obra –según la ISAG(G) sería con nivel de ‘’unidad documental compuesta’’–. Hasta este punto se tendría una instancia de obra con su descripción mediante una instancia de documento sin documento representativo; desde este momento, el primer documento físico que se añada se registrará dentro de esta primera instancia existente, para el resto se irán creando nuevas instancias de documentos asociados a la obra.

Cabe señalar que en ciertos tipos de obras no hay documentos tangibles de nivel de obra, pero estas serán descritas, y en otros tantos hay varios; dentro de esta casuística hay que mencionar el material multimedia que suele aludir a dos o tres tipos de obras.

Por el contrario, en lo referente a las interpretaciones de las mismas, por parte de las agrupaciones anuales, la relación se hace a nivel de los subtipos de esta entidad; para el concepto de interpretación hay que tener en cuenta el uso de tres (o menos) obras de diferente tipo, lo cual sería más dificil de gestionar dentro del ámbito de la integridad referencial.

---

[4] “*General International Standard Archival Description*”. Consejo Internacional de Archivos

Finalmente, en cuanto a su estructura, para este concepto se utiliza como identificador signatura, por ser numérica y de denominación más acorde a la terminología archivistica; luego se comentará la necesidad de introducir este valor en el campo identificador de los documentos, el cual es más complejo sintácticamente. En cuanto a la posibilidad de introducir aquí características como el año, el título y demás, se desestima por estar necesariamente en la entidad Documento –con todos los datos recomendados por la ISAD(G)–; de hecho, el año es también derivable a partir de la agrupación que la utiliza.

Literaria, **Musical** y **Grafica**

Mediante estos subtipos de Obra se modelan las producciones textuales, musicales y los diseños, que abarcan desde el tipo (disfraz, caracterización y actitud) hasta las diversas clases de atrezo. La función principal de estas entidades es actuar de compuestos en una relación de composición; no obstante, en los dos primeros además se tiene un atributo para el repertorio donde se resume la composición de la obra. Este último es otro caso de característica que podría ser generada; sin embargo, y por ejemplo, considerando que solo se cuente con un documento de audio a nivel de obra –que sería una grabación de la actuación completa de una agrupación sin dividir en piezas– se describe el número de piezas por cada género.

Tal como se ha expuesto en la descripción de Obra se creará una instancia de Documento para describir cada uno de los tipos; de hecho, la creación de uno de estos subtipos implicará la creación de una instancia de Obra y otra de Documento, puesto que la Obra se modela como abstracta. De la misma forma, para cada uno de los tres tipos de componentes, que se describen seguidamente, se dará la misma operación; por cada nuevo componente a asociar a una obra se creará un documento que permita describirlo dentro de las pautas de la ISAD(G), solo que en ese nivel tendrán carácter de ‘’unidad documental simple‘’.

Por último, atendiendo a las signaturas de los componentes y de estos tres subtipos de Obra con rol de compuestos, dentro del atributo correspondiente a signatura de la ISAD(G) estarán contenidas las dos, según la estructuración jerárquica dada por las pautas de la norma CDU.

Letra

Esta entidad representa el primero de los componentes de las relaciones de composición recién detalladas, el de las obras literarias, y contendrá el detalle de las letras con enfoque individual. Este concepto es, si cabe, uno de los focos de mayor interés del dominio del Carnaval de Cádiz. Como características tiene el género, la transcripción y la signatura, el resto de información asociada estará contenida en la instancia mínima de Documento que tendrá asociada.

Musica

En este segundo componente solo se contiene la signatura y el género. A este nivel pueden tenerse tanto documentos de audio como partituras, o vídeo siempre que contenga una pieza musical indivisible.

Cabe decir que en esta entidad y la anterior el género se modela mediante enumeraciones, además difieren entre las dos entidades en gran parte. Por lo demás tiene las mismas características de diseño comentadas para las letras.

Disenio

Último componente de Obra, concretamente de las gráficas, en lugar de género tiene un atributo de tipo descriptor, también modelado como enumeración. En primera instancia, dentro del alcance del presente proyecto, no se plantea la especialización; sin embargo, de aplicarla, los valores de este atributo serían los subtipos candidatos.

En lo referente a la semántica de la entidad, se pueden crear instancias para bocetos, imágenes, carteles, etc.; pero, dichos valores quedarán descritos en el Documento asociado, mediante los descriptores de esta última –los modelados mediante la entidad Descriptor–. El atributo dedicado al tipo refiere al tipo del diseño, no al del documento, ni su estructura o contenido.

Evento

Mediante esta entidad se describen los diversos tipos de eventos que se dan en el carnaval, desde el pregón a las cabalgatas; sin embargo, el alcance del proyecto está centrado en los concursos. Para este último caso se da la necesidad de especializar, de nuevo por tener un vínculo con las agrupaciones, dentro de la semántica de concurso, además de estar compuestos de premios.

Como características a nivel de evento se tiene el año, el nombre, la localidad y la provincia, todos estos componen la clave candidata porque los nombres suelen ser poco específicos, y los mismos eventos se dan en diversos lugares. Así mismo, aunque opcionales, se consideran de interés la ubicación o emplazamiento y la descripción.

Concurso

Esta entidad, al igual que los subtipos de Obra, es utilizada para especializar a Evento por las asociaciones concretas de este tipo de eventos dentro de su nivel semántico. En este nivel las características de interés son la categoría y los premios, los cuales se modelizan mediante el siguiente concepto, por ser varios y tener características propias. Por otra parte, tal como se ha mencionado anteriormente, se vincula con las agrupaciones de carácter anual por una relación de participación.

Premio

Describe las características propias de cada premio dado en un concurso: nivel del premio alcanzado, modalidad y cuantía, más obserbaciones, donde se podrán aclarar algunas particularidades; por ejemplo si son de tipo accésit, etc. En este caso la clave candidata requiere la modalidad y el premio, además de la referencia a la entidad de la que es componente, Concurso. Hay que tener en cuenta que un mismo premio se da tantas veces como modalidades participan, entendiendo premio como clasificación o puesto.

Documento

Esta entidad se contienen toda la información que recomienda la ISAD(G) para la descripción archivística, lo que satisface el requisito RF03. En consecuencia, no por semántica del dominio, sino por el objetivo normalizador del mismo, esta entidad es de gran interés y la más extensa, a la vez que la menos óptima en lo que a diseño de BBDD refiere. Los campos de la ISAD(G) son en su mayoría narrativos, aunque semánticamente puedan considerarse atómicos; de hecho,

generan redundancias en cuanto a datos existentes en otras entidades, incluso en esta misma –por ejemplo, la ubicación física del documento tiene atributos dedicados para facilitar la gestión de los mismos, más dicha localización está descrita en uno de los campos de ISAD(G)–.

Por tanto, además de los campos de ISAD(G) se incluyen los recién comentados, necesarios para la gestión de los mismos y otros tantos para uso y conocimiento interno del Aula, como el tipo de documento, el tipo de soporte y la URL si está digitalizado. Por último, aunque ciertamente compleja, la clave candidata es la que propone la ISAD(G), el campo 3.1.1 Código de Referencia; donde, como se ha indicado en la descripción de las entidades Obra, Letra, Musica y Disenio, en función del nivel de descripción de la unidad documental, simple o compuesta, la signatura tendrá más o menos niveles.

En lo referente al tipo de documento, se modela mediante una enumeración con los valores: Texto, Imagen, Audio, Video, Expediente, Serie, Fondo y Colección. Esta lista es cerrada, solo podría admitir valores con similar nivel de abstracción, donde únicamente se distingue entre los tipos de documentos existentes en el dominio; solo diferencia entre objetos compuestos e indivisibles y, dentro de estos últimos, el tipo de contenido sin entrar al detalle del soporte físico, formato o tipo de la información. Un modelo alternativo supone tener una entidad con este conjunto de valores, el cual se consideraba de utilidad para implementar la funcionalidad descrita por el requisito RF04.7; sin embargo, implica pérdidas de eficiencia derivadas del uso un índice con alto porcentaje de valores repetidos, más la necesidad de utilizar la combinación para obtener los valores útiles. Tener un índice sobre el atributo alfanumérico directamente no introduce más complejidad que la alternativa comentada y resulta en un modelo más sencillo. En cuanto al requisito funcional RF04.7, con base en los valores se puede generar la consulta adecuada; en cualquier caso, por su carácter estadístico, se introduce dentro de la propuesta de *data warehouse.*

Retomando la importancia que gana esta entidad en el modelo desde el enfoque descriptivo, en ella se contienen todos los atributos necesarios para catalogar las obras y sus componentes, según las pautas de asociación dadas en el detalle de la entidad Obra. Se planteaban diferentes posibilidades de acceso y creación de las instancias de esta entidad; por un lado, se tenía la opción de crear documentos y asociarlos a obras existentes, esto llevaba a posibles inconsistencias si las obras ya tenían alguno vinculado –para la descripción de la misma–, pero sin elemento tangible asociado; por otro lado, restringir el acceso a la creación de documentos únicamente a partir de eventos, obras o componentes de estas últimas.

Este último esquema operativo ha sido la opción elegida, aunque limita en cierta medida la usabilidad del sistema aporta consistencia al conjunto, a la vez que elimina redundancias y facilita la organización jerárquica; la cual pasa a estar dominada por las obras y, en menor grado, por los eventos. Cabe aclarar que un documento puede estar asociado a obras y/o eventos; es decir, como mínimo a una obra (o alguno de sus componentes) o a un evento, aunque puede aludir a ambos conceptos. Toda esta casuística se desarrolla considerando en todo momento las posibilidades que ofrecerá la futura interfaz.

Finalmente, para contribuir a la coherencia del conjunto de valores contenidos en la entidad, respecto al hecho de tener asociado o no un documento tangible a una instancia de Documento, se introduce un tipo de soporte llamado “No

disponible”; el cual, además, estará condicionado por los datos gestionados por la entidad GestionDocumento. En el Anexo I se desarrollan todos estos argumentos y deciciones.

DepositoLegal

Este concepto está ligado a las instancias de Documento, de forma opcional por la premisa general de incertidumbre. Es una información que se genera cada vez que una obra sale de la editorial, la imprenta o la productora. El conjunto ha sido extraído de la entidad Documento por la interdependencia entre sus atributos, los cuales implican información en si misma y se identifican mediante uno de ellos. Además, con respecto de los documentos, todos ellos son multivaluados, se generan simultáneamente y deben estar ordenados, aun así los únicos requeridos son los participantes en la clave candidata. Cabe añadir que, aun siendo la clave candidata compuesta, por el tipo de relación de la que participa, el valor de Depósito Legal es clave candidata en si mismo, independientemente del documento. Por añadidura, se prevé un ingente de documentos sin este concepto asociado, prácticamente todo el multimedia, por tanto esta decisión extrae un alto porcentaje de valores nulos de Documento.

3.2 Relaciones

Las decisiones de diseño de las relaciones abarcan los requisitos funcionales acerca de las mismas definidos en el RF02. Por otra parte, como se mencionaba en el caso de las entidades, se detallan a continuación las divergencias semánticas entre los requisitos de proyecto y el modelo definido, en lo referente a las relaciones:

- Se mencionaba la relación de autoría entre personas y obras, tras varios encuentros con el cliente se aclara que en la creación de una obra pueden participar varios autores, incluso que la autoría puede ser de la agrupación; pero que, a nivel de componente de obra (letra, composición musical o diseño) solo puede haber uno, al menos oficialmente.
- Se contempla también la relación de agrupaciones con documentos, con dos enfoques distintos. Por un lado, están los documentos administrativos que se relacionan a entidades –entiéndase ya personas, agrupaciones y empresas– con eventos, lo cual no se modela por estar fuera del alcance del presente proyecto. Y, por otro lado, los documentos dedicados a la obra que interpreta la agrupación. En ambos casos, la relación se implementa a través de la entidad Documento; en el primer escenario, un documento puede pertenecer a un evento y aludiría a una entidad –cosa que no se ha modelado–; en el segundo escenario, un documento describe o pertenece a una obra y, esta última, es interpretada por la agrupación; por tanto, en este caso la relación es indirecta entre documentos y agrupaciones, ya que predomina la pertenencia de los mismos a las obras.
- Por último, a modo de inciso, la relación entre agrupaciones y entidades pasa a ser entre agrupaciones y empresas más, de nuevo, de forma indirecta. Una empresa diseña y/o ejecuta la puesta en escena de la obra gráfica que, por otro lado, es utilizada por la agrupación. Por tanto, la contratación de empresas por parte de las agrupaciones queda descrita, también en este caso, a través de la obra gráfica; donde, la

obra tiene relación directa con la agrupación y la empresa, se considera la derivación de la información contractual entre empresa y agrupación.

Cabe añadir que, dentro del dominio del Aula como institución y del Carnaval como evento, las agrupaciones no tienen porque proveer de su documentación administrativa interna y, por añadidura, existen premios a los diseñadores de las obras gráficas –también fuera del alcance del proyecto–; con base en estos dos hechos, se confirma la corrección del modelo conceptual propuesto.

rol (Persona ↔ AgrupacionAnio)

Una persona puede participar con más de un rol en una o más agrupaciones por año; por ejemplo, ser director y guitarrista. Es especializada mediante las entidades Cargo, Orquesta y Voz debido a las peculiaridades de cada uno.

La complejidad de esta relación es la forma de identificar sus ocurrencias de manera única. En primer lugar, una persona puede participar en más de una agrupación en un año; en segundo lugar, una persona respecto de una agrupación concreta, puede adoptar varios tipos de roles; consecuentemente, según el tipo de rol se podrán dar repeticiones de pares Persona y AgrupacionAnio. Por otro lado, en el caso de los roles de voz y orquesta la adopción de uno impedirá la del otro, siendo exclusivos entre ellos; mientras que cargos se pueden ostentar varios.

Otro rol de gran importancia es la autoría, de letra y música, se decide modelarlo a través de la obra interpretada por la agrupación en ese año, concretamente desde los componentes Letra y Música de los compuestos Literaria o Musical; de esta forma, se admite la existencia de más de un autor en el nivel de la obra. De nuevo surge la opción de tener un atributo derivado que obtenga, por ejemplo, el autor mayoritario pero se desestima –cabe señalar que actualmente no suele ocurrir que más de un autor escriba para una agrupación, es un condicionante impuesto por los datos históricos–.

Por otra parte, con esta decisión se evita la trampa de diseño de corte, en caso de eliminar una agrupación y sus roles se perdería esta información; además, no habría autorías de obras no interpretadas; por último, a nivel semántico, es más pertinente la autoría asociada a las obras que a su agrupación interprete. La problemática proviene de que, en el dominio de interés que se está modelizando, al hablar de autor se hace a nivel de agrupación; pero, a nivel legal la autoría está en el nivel de componente de obra.

Cargo y **Voz**

Especializan la entidad que modela la relación de asociación rol, entre Persona y AgrupacionAnio, recién descrita. En estos subtipos y en el siguiente se incluye un solo atributo, inicialmente propuesto como enumeración, por ser un conjunto semicerrado de valores. En la entidad Cargo se tratan conceptos políticos y organizativos, donde se hace necesario un espacio para observaciones, y en Voz los diversos registros participantes.

Para la identificación en estas subclases, ambas heredan el identificador compuesto mencionado antes. La restricción de unicidad debe permitir que una persona pueda adoptar varios cargos y solo una voz; así como que, si adopta un rol de voz no utilizará ningún instrumento. Por tanto, la clave candidata en Cargo introduce a la heredada el tipo de cargo; mientras que, para Voz, la clave

candidata es la heredada, lo cual no permite tener más de una ocurrencia en Voz por persona para una agrupación dada.

Orquesta

Tercer caso de especialización de la relación asociativa rol. La característica específica de la subclase es el instrumento que utiliza la persona descrita mediante una instancia de rol en concreto, solo de entre los reconocidos por los reglamentos de concurso. Como se indicó en la descripción de la entidad Instrumento, la lista de valores se extrae para su uso desde este subtipo de rol y desde la agrupación.

En cuanto a la identificación de las instancias de esta relación, al igual que en Voz, la única clave candidata es la heredada con objeto de no admitir repeticiones del par persona y agrupación, porque no es posible que un integrante toque más de un instrumento reglamentario por agrupación.

Orquesta → **Instrumento**

Tal cual se ha indicado, mediante esta relación se modela el valor de instrumento del rol que puede adoptar un integrante de tipo orquesta; así mismo, por el carácter accesorio de los valores de esta lista, se tiene navegabilidad en solo en una dirección.

Respecto a la multiplicidad, cabe indicar que esta entidad representa la vinculación de un integrante con una agrupación, y por año no utiliza más de un instrumento –atendiendo al reglamento–; aunque, si puede hacer uso de otros puntualmente, estos serán descritos mediante la siguiente relación; por ende, la clave candidata es el identificador del rol Orquesta unicamente.

usa (AgrupacionAnio → Instrumento)

El uso de la entidad Instrumento es similar al anterior; sin embargo, en este caso la multiplicidad es incierta a ambos lados; aunque se de la navegabilidad indicada por la orientación de la flecha, por el carácter adjetivo de esta relación, cada valor de instrumento puede ser utilizado por varias agrupaciones.

El objetivo es describir el total de instrumentos diferentes utilizados por una agrupación, se conozcan o no las personas que los utilizan; por ejemplo, en el caso de las callejeras no es conocida la participación ni las personas (al menos oficialmente). En definitiva, se trata de una relación de cardinalidad N:M donde la clave candidata está compuesta de los identificadores de las involucradas.

AgrupacionAnio →**Modalidad** y **Premio** → **Modalidad**

Este es otro caso en que varias entidades utilizan un mismo concepto, la modalidad. En ambos la extracción del atributo, que fuera enumeración, resulta de utilidad por ser foco de interés en búsquedas y ordenaciones. De nuevo por el carácter adjetivo de la entidad Modalidad respecto de las agrupaciones y los premios, la clave candidata de estas relaciones son los identificadores de AgrupacionAnio y Premio.

participa (AgrupacionAnio ↔ Concurso)

Mediante esta relación se describen las particularidades asociadas al hecho de participar en un concurso, como son la clasificación, la fase o los puntos obtenidos. Por la premisa general de incertidumbre que afecta a los datos los

atributos son todos opcionales, exceptuando los elementos de la clave candidata –ambos estructurales–; por tanto, de existir la relación, lo que describe es el hecho de la participación en el concurso. Por otro lado, de ganar algún premio, la cuantía se obtiene de la relación de composición entre esta relación junto con Concurso y Premio.

En lo referente a la multiplicidad, en un concurso hay un mínimo de una agrupación, de hecho más, pero al declararlo quizás no se le haya inscrito ninguna; por esto, la inserción de un concurso no debe forzar la creación de una instancia de la relación participa; es más, quizás no se hayan definido aún las agrupaciones del año del concurso recién insertado. Dicho esto, la clave candidata resulta del compuesto entre las referencias a las entidades involucradas.

autoriaLetra (Letra ↔ Entidad), **autoriaMusica** (Musica ↔ Entidad) y **autoriaGrafica** (Disenio ↔ Entidad)

La autoría puede ostentarla tanto una persona como una agrupación, o una empresa, si se trata del tipo o el atrezo; por otro lado, podría ser anónima o desconocida, para ello se crearán los valores “anónimo” y “desconocido”, con objeto de forzar la relación como obligatoria.

En cuanto a la identificación de las ocurrencias de estas relaciones, la multiplicidad es 1:M, desde el punto de vista de los autores, los cuales pueden haber compuesto innumerables obras; por tanto, la clave candidata sería el identificador de estos tipos de piezas, siendo posible un único autor por cada Letra, Musica o Disenio. Tal como se indicó se trata de una cuestión legal, aunque podría no ser del todo cierto.

crea (Disenio ↔ Entidad)

Se hace necesaria esta segunda relación entre estas entidades porque la autoría de obras gráficas no implica necesariamente la creación del diseño; aunque, puede darse que una empresa sea la autora, lo normal es que se adscriba el diseño a personas y su construcción a empresas. En cuanto a la multiplicidad, un diseño puede ser ejecutado por varias entidades y una entidad puede crear varios diseños al año; por ello, con cardinalidad N:M y como clave candidata el compuesto de los identificadores implicados.

Empresa → Actividad

Este es otro concepto que fue modelado como enumeración por su carácter adjetivo. Se propone como relación de uno a muchos desde Empresa para poder describir empresas de múltiples capacidades; aún así, un mismo tipo de actividad puede ser desarrollado por varias empresas, teniendo en cuenta que esta relación fuera inicialmente una enumeración, Actividad para a ser una estructura de tipo *lookup*; en definitiva, se modela como relación N:M con su clave candidata compuesta de las referencias a las entidades implicadas.

Documento

Para describir las relaciones en que participa esta entidad se tendrán en cuenta diferentes enfoques, según el concepto con que se le vincule; en primer lugar, respecto a las propiedades del documento en si, luego las que lo relacionan con los objetos que mediante instancias de Documento se describen.

→ **Descriptor**

Relación mediante la que se describe el documento desde un punto de vista estructural y de contenido. Se puede considerar una relación con cardinalidad 1:M, desde Documento, de nuevo con carácter adjetivo; pero, al igual que en el caso de empresas y actividades recién detallado, un descriptor podrá utilizarse por muchos documento; por lo tanto, como identificador de la relación el par de identificadores involucrados.

↔ **DepositoLegal**

Con multiplicidad indeterminada, da el número de veces que una Obra ha sido publicada y revisada; y, en caso de desconocimiento de esta información, evita valores nulos en Documento. En lo referente a identificación unívoca, cada instancia de DepositoLegal solo se vincula a un único documento y puede haber varias por documento; además, en si mismo el valor de Depósito Legal es único en el conjunto de documentos que pasan por este proceso de registro.

Respecto a productores o implicados, puede tener relación con:

↔ **Evento**

Una unidad documental se asocia mediante esta relación a uno o más eventos. Son ejemplos: los fallos del jurado, las inscripciones, normativas, boletines, programaciones, etc. Por otra parte, se considerará suficiente que un documento esté relacionado solo con Evento, y no con Obra, Letra, Musica y Disenio; es decir, un documento puede estar vinculado a obras y eventos, o solo a eventos, siendo válido. Con base en estas condiciones, la relación se modela con multiplicidad N:M pero la unicidad debe ser gestionada en otro nivel, por la interdependencia que tiene la existencia de instancias de esta relación respecto de las obras.

La relación con la Entidad, como generadora o afectada, se obtiene a través de la autoría de los componentes de las obras, más de los creadores en caso de las gráficas. Cabe señalar que se podría relacionar Documento con Entidad para describir los no vinculados a las obras, pero debido al alcance que se ha propuesto para el proyecto no se describen las relaciones entre documentos y entidades –habría que añadir la Entidad de tipo Institucion y restringir que no exista relación con obras si existe una directa con Entidad, para evitar trampas de diseño–.

Respecto a las unidades documentales compuestas:

↔ **Obra**

Por defecto, al crear una nueva Obra se crea una instancia de Documento y se vinculan mediante esta relación. La cardinalidad entre Documento y Obra implica que una Obra puede tener varios documentos vinculados; pero, desde el punto de vista de Documento tener varias obras vinculadas implica que el documento es de nivel serie, fondo o colección, lo cual no se implementará dentro del alcance del proyecto.

Por otra parte, el modelado de esta relación se hace al nivel de Obra, no de sus subtipos, y se deben considerar ciertas particularidades según el tipo de documento.

- En cuanto a la subclase Literaria, si el documento es de tipo texto e incluye la obra completa, podría tratarse de un libreto y solo existiría una instancia de esta relación.
- Para la obra Musical, el caso sencillo sería un conjunto de partituras u otro tipo de documentación asociada a la misma obra, a modo de expediente; sin embargo, la mayoría de las veces que se cuenta con documentos musicales con alcance de obra son de tipo multimedia. Por ejemplo, sería el audio o vídeo de una actuación completa, en cuyo caso habría que relacionar el documento a las obras gráficas y literarias correspondientes, todas del mismo nivel documental.
- Por último, en el caso de las obras gráficas; de nuevo, el simple sería un conjunto de imágenes que describen el tipo completo de determinada agrupación, incluido el atrezo, el forillo y la caracterización –o un cartel asociado a un evento, que estaría a nivel de obra en un solo documento–; y, documentación diversa.

Tanto en este nivel como en el siguiente, los documentos no podrán estar relacionados con una obra y un componente simultáneamente, o están relacionados a Obra o a Letra, Musica y Disenio; a diferencia de si un documento contiene referencias a más de una tipología de obras, el nivel documental es lógicamente diferenciable, ya que es divisible o indivisible. Mientras que, según lo expuesto, dentro del mismo nivel documental se pueden dar varias relaciones en función del contenido.

Los documentos de tipo video son un caso especial en el uso de estas relaciones. Si tienen alcance de obra completa se pueden vincular a los tres tipos de estas; por añadidura, suponiendo una excepción a la restricción de vincular obras con componentes, las obras gráficas estarán completamente descritas de forma visual en cualquier video con alcance a componente de obra; es decir, los que contienen piezas indivisibles, vinculables a componentes de tipo Letra y/o Musica –ambos están en el nivel de unidad documental simple, según la ISAD(G)–.

Por último, aunque se han tratado tres tipos de relaciones todas se modelan entre las entidades Obra y Documento, por lo que la relación tiene cardinalidad N:M con la clave candidata compuesta; en consecuencia, la unicidad de las instancias de esta relación de enfoque genérico de Obra se debe controlar en otro nivel.

Respecto a las unidades documentales simples:

↔ **Letra**

Se utiliza esta relación por la necesidad de describir las letras según las recomendaciones de la ISAD(G), por los argumentos dados en la descripción de esta entidad y las dos siguientes. En cuanto a los documentos físicos de este nivel, rara vez se conservan individualmente –siendo el pregón de nuevo una excepción–, pero suelen existir vínculos entre Documento y Letra si se cuenta con audios de alcance de pieza.

↔ **Musica**

Además de la instancia mínima de Documento y de esta misma para relacionarlos, se crean documentos para adjuntar partituras, audios u

otra documentación asociada a un único componente de la relación de composición Musical. Por tanto, al igual que la anterior modela una relación de cardinalidad 1:M, pudiendo haber más de un documento por cada componente de tipo Musica.

↔ **Disenio**

Permite asociar al diseño varios documentos, como bocetos y fotos, considerando solo los individuales y estáticos; pero, como en los anteriores, la gran mayoría suelen adscribirse al tipo multimedia, concretamente de video, en cuyo caso la asociación entre el diseño y el documento está a nivel de Obra.

Por la complejidad que introduce la descripción del material multimedia, habrá que tener en cuenta el tipo de documento para localizar las obras o componentes relacionados con el mismo. Como en otros casos, aunque se planteen las vistas necesarias para representar esta complejidad, será en el ámbito de la interfaz donde sea accesible uno u otro esquema para gestión de la información. Por ejemplo, de ser un documento de tipo audio, será necesario mostrar los registros contenidos en las entidades Literaria y Musical, o Letra y Musica; mientras que de ser tipo video, se podrá acceder a la obra gráfica y a Letra y Musica, o Musical y Literaria, según el nivel documental.

Como se ha expuesto durante la descripción de las entidades Obra y Documento, la inserción de instancias en estas cinco últimas relaciones debe provenir del lado Evento u Obra. En el único caso que se editarán los documentos respecto de sus relaciones –lo que implica crear instancias de estas cinco relaciones desde el lado Documento– será en caso que ya exista al menos una instancia en cualquiera de las otras cuatro. De esta forma se facilitará la vinculación de documentos con las demás entidades, en cualquier caso habrá que localizar los nuevos extremos de la relación previamente; de nuevo, cabría plantear las posibilidades que debe ofrecer la futura interfaz. En resumen, la jerarquización de los niveles documentales, relativos a las obras y sus componentes, queda modelada en estas últimas relaciones.

Por último, teniendo en cuenta que estas tres relaciones se modelan por separado, resta aclara la multiplicidad independiente que cada una tiene respecto de Documento; es decir, sin considerar la condiciones de vinculación entre Documento y estas tres en función del tipo del mismo –lo cual debe gestionarse a un nivel superior–. Todas ellas son de tipo 1:M, donde un Documento solo puede aludir a un componente o pieza de una Obra simultáneamente, y un componente de cualquiera de los subtipos de Obra tendrá un número indeterminado de documentos, con un mínimo de uno; en consecuencia, la clave candidata es el identificador de la instancia de Documento participante en estas relaciones.

ComposicionLiteraria, **ComposicionMusical** y **ComposicionGrafica**

Describen la estructura definida para las obras respecto de sus piezas, todas modeladas como composición, por ser los componentes (Letra, Musica y Disenio) obras en si mismas cuando son uno solo y necesitar de diferente nivel de descripción. Al hacer uso de este tipo de relación se admite la descripción de un compuesto (Obra) que aun no tiene componentes (piezas o diseños), pero no se podrá describir un componente del que no existe obra, puesto que el componente en sí sería una obra de ser único –de nuevo el pregón–. Esto

se controla, en cierto modo, mediante la creación de instancias de Documento para cada uno de estos niveles documentales.

ComposicionAgrupacion

Bajo las condiciones recién dadas, se entiende como dos niveles de descripción de las agrupaciones, la de la entidad indefinida en el tiempo y la del producto anual; aun considerando las agrupaciones de aparición eventual. Cabe comentar que en esta, y las tres relaciones de composición anteriores, la clave candidata se compondrá del identificador del compuesto más uno propio del componente, que es lo común en esta técnica de modelado conceptual.

GestionDocumento ↔ **Documento** y **GestionInformacion** ↔ **AgrupacionAnio**

Ambas son relaciones 1:1, tal como se ha detallado para estas entidades, que modelan la asociación entre documentos e información registrada acerca de las agrupaciones, respecto de la información necesaria para dar soporte al proceso documental de los anteriores. La cardinalidad de estas relaciones representa una dependencia de existencia, el identificador es el mismo en estas que en las dos entidades dominantes de cada una.

Letra ↔ **Tema**

Mediante esta relación de multiplicidad N:M se gestiona el etiquetado de las letras mediante temas. La navegabilidad es necesariamente bidireccional porque será habitual hacer búsquedas en Letra según los temas, y viceversa; luego, la clave candidata es la común para relaciones de tal multiplicidad.

Disenio ↔ **Tema**

Similar a la anterior en diseño, cualquier tipo de diseño puede ser catalogado con varios temas; sin embargo, suele ser el tipo, o la idea a representar, el que se describe mediante temas.

interpretaMusical (AgrupacionAnio ↔ Musical),
interpretaLiteraria (AgrupacionAnio ↔ Literaria) y
utilizaTipo (AgrupacionAnio ↔ Grafica)

En todos estos conceptos, se modela el hecho por el que las agrupaciones anuales existen: la interpretación. Se hace notar mediante la multiplicidad 1:1, opcional a ambos lados, que las agrupaciones deben interpretar/utilizar una obra por año pero las obras no tienen porque ser interpretadas por las agrupaciones. Por ejemplo, el cartel del carnaval no se utiliza por ninguna agrupación, o el pregón (que lleva música y letra) tampoco. Por otro lado, las obras que se interpretan/utilizan mediante estas relaciones tienen el nombre de la agrupación para ese año, ya que en realidad están generadas a tal efecto.

3.3 Restricciones de integridad

Se enumeran a continuación las implicaciones semánticas y referenciales de las entidades y relaciones según el orden utilizado en el apartado anterior, omitiendo las que no tienen condicionantes. Por otra parte, se consideran dadas las referentes a claves candidatas y datos obligatorios, los cuales están explícitamente marcados en el modelo lógico que se describe en el apartado

4.1 del presente documento; así mismo, se tienen como no pertinentes las restricciones a nivel de tipo de atributo –como los intervalos o mínimos en los numéricos–.

Modalidad, **Instrumento** y **Actividad**

En todas esta entidades, con funcionalidad *lookup*, la única restricción es la no repetición de los valores contenidos.

Descriptor y **Tema**

Estos dos conceptos son similares a los anteriores pero por la extensión que pueden alcanzar, sobre todo el segundo, mantener la unicidad de los valores contenidos será complejo. Para ambos casos la única vía se situa en el nivel de la interfaz, incluso de los usuarios; problemáticas como los paralelismos, la inversión de palabras que no modifican su significado o la sinonímia no pueden ser controlados mediante las funcionalidades de los SGBD o la programación.

GestionInformacion

Se debe controlar que cuando uno de los booleanos pase a verdadero se introduzca algún dato acerca de las fuentes de forma obligatoria.

GestionDocumento

Se debe controlar que cuando el booleano que indica que el documento está digitalizado pase a verdadero exista algún valor en URL; y, si el que indica que está en depósito pasa a verdadero, se aporte información de la ubicación concreta en la instancia de Documento asociada.

Por otro lado, de nuevo sobre la instancia de Documento modificada, el tipo de soporte no podrá ser “No disponible”. Con este valor se diferencia entre los documentos utilizados para describir obras y componentes a nivel abstracto y los que contienen algún documento representativo de las mismas; por tanto, al indicar que el documento está digitalizado o en depósito el soporte físico será conocido.

Entidad

Este concepto se ha modelado como abstracto, por tanto no podrán crearse instancias del mismo de forma directa; pero, por cada instancia de cada subtipo, existirá una en este nivel que debe contener el dato del tipo donde se encuentra, lo que suele modelarse a modo de atributo discriminante; por ende, esta entidad contendrá una lista con el total de identificadores de entidades posibles, junto con el subtipo donde se encuentra en uso cada uno.

Para este caso y el de Obra, que es el otro caso de entidad abstracta, se controlará la no inserción directa mediante la creación de procedimientos almacenados de alta solo para los subtipos, de ambas; de hecho, atendiendo al requisito no funcional RNF03.1, se encapsulará la BBDD completamente mediante procedimientos almacenados que restrinjan la actualización directa de cualquiera de las entidades.

Estos procedimientos de alta crearán la instancia en Entidad, luego tomarán el identificador generado por esta para crear la instancia del subtipo y su relación con respecto a esta superclase.

Persona

Como soporte a la identificación unívoca, de forma accesoria, se forzará que de no haber segundo apellido o un DNI exista un domicilio con localidad. Para ello se cuenta con que siempre se tiene la dirección cuando falta el segundo apellido y el DNI –con el año 1984 como punto de inflexión–; y, todo esto, con la presunción de que no se registrarán personas de forma extraoficial.

En cuanto a integridad referencial, para este y los otros dos subtipos de Entidad, se restringirán la eliminación de la instancia de Entidad a la que refieren; no obstante, tal y como se ha expuesto, no se permitirá la edición directa de las entidades modeladas como abstractas, sería el procedimiento almacenado para bajas el que eliminaria el subtipo y luego la instancia de la superclase. Por otra parte, la actualización se implementa en cascada en todos los conceptos donde se admite la inserción del identificador por parte del usuario.

Agrupacion

Unicamente las restricciones de integridad referencial recién comentadas para su análoga, Persona, y como mínimo la información relativa a la procedencia.

AgrupacionAnio

Debe existir una tupla de Agrupacion a la que se asocie, si no existe la indicada o no se indica ninguna hay que forzar la creación de una en blanco. Esto se gestionará mediante procedimientos almacenados, según el requisito de encapsulación, dicho procedimiento tendrá que recibir un identificador de Agrupacion o, en su defecto, como mínimo localidad y provincia (ambos obligatorios en la relación Agrupacion).

Por otra parte, el número de componentes debe ser coherente con el total de roles de tipo Voz y Orquesta asociados a la agrupación; por simetría, habría que actualizar este atributo si está establecido antes de crear nuevas tuplas en Voz y Orquesta. Se desestima esta operativa, podría registrarse el total de componentes y luego vincular personas, lo que crearía inconsistencias.

Por último, en cuanto a la integridad referencial, por ser una relación de composición, no podrá eliminarse ninguna Agrupacion si existe AgrupacionAnio asociada. Cabría valorar si se permitirá, a través del procedimiento de baja asociado a esta entidad, realizar el borrado en cascada de toda información asociada a la instancia; por ejemplo, en la futura interfaz se podría informar al usuario de las consecuencias de realizar la baja de determinada agrupación.

Obra

De la misma forma que en el caso de Entidad, no se podrán introducir instancias de esta entidad abstracta directamente, según los mismo argumentos expuestos en la anterior y mediante los mismos mecanismos. Por otro lado, la documentación ya registrada en las BBDD del cliente está signaturada de forma numérica, por lo que también se debe admitir la recepción de este dato.

Respecto a la integridad referencial, se modela y plantea la herencia de forma análoga a la generalización en Entidad; es decir, no se podrá editar el identificador de esta entidad de forma directa, sino mediante el procedimiento almacenado de altas y bajas de las subclases de la misma.

Literaria, **Musical** y **Grafica**

Al crear uno de estos subtipos de Obra se creará una instancia en Documento y, en consecuencia, otra en la relación DocumentoObra, según los argumentos dados en el detalle del actual modelo en cuanto a multiplicidades mínimas de las relaciones Documento, Obra y DocumentoObra. Al ser compartida la clave primaria entre los subtipos, la condición de mínimo una tupla por documento en DocumentoObra facilita esta restricción a nivel de Obra.

Con base en este esquema de vinculación, que se repetirá en el nivel de los componentes de estas entidades compuestas, se procede a describir la forma de signaturización de los documentos asociados a estos subtipos. Teniendo en cuenta que, en estas entidades la signatura es numérica, la complejidad es componerlas para identificar el Documento. El identificador de la entidad Documento se utilizará como segundo nivel para el compuesto de signaturas que se plantea como valor del campo 3_1_1_CodigoReferencia de la ISAD(G). En este nivel, dedicado a las unidades documentales compuestas, sería la compuesta por: la signatura de Obra y el identificador del Documento. Podría refinarse más la gestión, pero con este esquema se consideran completamente diferenciados los elementos totales a describir y se admite la reutilización de los identificadores ya asignados por el Aula (el cliente) en la catalogación de sus fondos.

Por último, debido a las relaciones de herencia, en estas subclases de Obra la integridad referencial impedirá el borrado de instancias de la superclase.

Letra, **Musica** y **Disenio**

Al igual que para las entidades anteriores, se crea una tupla en las relaciones Documento y Documento_Letra, Documento_Musica o Documento_Disenio, respectivamente, de forma conjunta; por tanto, el procedimiento almacenado que actualiza esta relación debe recibir mínimo los campos obligatorios de estas entidades y de Documento para las operaciones de inserción. De nuevo en este nivel debe controlarse la introducción de su valor de signatura, ya sea dado por el usuario o asignado sintéticamente por el SGBD.

Por tener estas entidades con rol de componente en relaciones de composición, se plantea que los identificadores sean secuenciales por cada compuesto; es decir, por cada obra se tendrán las signaturas de los componentes numeradas del 1 en adelante.

Análogamente a la composición de las signaturas en dos niveles que se ha detallado en la descripción de las restricciones de Obra, el valor para el atributo 3_1_1_CodigoReferencia de la ISAD(G) a nivel de unidad documental simple será la tripleta compuesta por: el valor de la signatura de Obra (que esta contenida en estas entidades por la composición), la signatura del componente en concreto y, por último, el identificador del Documento. Podría refinarse más la gestión, pero con este esquema se consideran completamente diferenciados los elementos totales a describir y aún se admite la reutilización de los identificadores del cliente.

Evento

Por el contrario que en las relaciones de generalización anteriores, la creación de eventos no forzará la creación de instancias del subtipo, puesto que no están todos los posibles especializados por herencia; sin embargo, de existir sí

se debe forzar la creación, como en el caso de Concurso. Por tanto, el procedimiento almacenado de altas de esta entidad dará error si se le pasa el valor "Concurso" como tipo de evento.

Concurso

Tal como se ha indicado habrá un procedimiento almacenado de alta para esta entidad, el cual recibirá los atributos de Evento como argumentos, y gestionará la integridad referencial entre las dos relaciones. En cuanto a esta última, de nuevo debido a la relación de herencia que protege la restricción, la eliminación del Evento del que depende un Concurso se evitará. Con esto, de tener que eliminar un Evento con un concurso asociado, habrá que eliminar el concurso; más, si dicho concurso tiene premios y relaciones de participación por parte de las agrupaciones, tampoco podrá ser eliminado sin retirar primero todas aquellas relaciones.

Premio

Aparte de que la cuantía de los premios sea mayor que cero, cosa que se pospone a la fase de diseño físico, se tienen las restricciones de integridad referencial. Considerando que se trata de una relación de composición, se bloqueará la eliminación de un Concurso en caso de tener información de premios asociada; de igual forma en caso de la Modalidad, en la cual se contienen valores de referencia, o soporte. Por otra parte, si hay alguna agrupación beneficiada del premio en el concurso concreto, aunque es de forma indirecta, tampoco se debe permitir la eliminación de estos.

Documento

En primer lugar, se tiene la restricción de dependencia con la entidad GestionDocumento; en segundo, las comentadas en cuanto a la composición del campo de ISAD(G) 3_1_1_CodigoReferencia que han sido desarrolladas respecto a los subtipos de Obra y sus componentes.

En cuanto a esto último, siguiendo la misma pauta de reutilización de identificadores existentes, se tiene el caso de los documentos vinculados a Evento que no tienen relación a las obras; en este caso se propone la combinación del identificador del evento y del documento a vincular. Con esta propuesta ya se tiene completamente gestinada la automatización del proceso de signaturización; luego, según las premisas de ISAD(G), a estas composiciones hay que añadir un prefijo que identifica el país y la institución que custodia los documentos, a modo de ejemplo se utilizará ES.ACCC.

También atendiendo a la dependencia que tienen los documentos de las obras y los eventos, habrá que controlar que cada documento aparezca como mínimo en una de las relaciones que implementan los tipos de relación N:M entre esta entidad y las mencionadas. Finalmente, considerando estos dos hechos, si está asociado en el nivel de Obra tendrá carácter de unidad documental compuesta y si lo está a un componente será de carácter simple.

También en este nivel hay que gestionar las posibles relaciones que puede tener una tupla de Documento en función del su tipo; concretamente:

- Si es texto y unidad documental compuesta se puede vincular a Literaria y/o una o más instancias de Evento.
- Si es texto y unidad documental simple se vinculará a Letra y/o Evento.

- Si es imagen y unidad documental simple se puede vincular a Disenio, Musica (en caso de las partituras) y/o Evento.
- Si es imagen y unidad documental compuesta se puede vincular a Grafica, Musical y/o a una o más instancias de Evento.
- Si es audio y unidad documental simple se puede vincular a Letra, Musica y/o Evento.
- Si es audio y unidad documental compuesta se puede vincular a Literaria, Musical y/o Evento.
- Si es video y unidad documental simple se puede vincular a Letra, Musica, Grafica y/o Evento.
- Si es video y unidad documental compuesta se puede vincular a Literaria, Musical y/o Evento.

DepositoLegal

Unicamente la integridad referencial, si se elimina un documento se eliminan las tuplas de esta relación que lo referencien; en esta ocasión por la relación de dependencia de existencia.

rol

En este caso se admite la duplicidad en la asociación entre persona y agrupación, pero solo para el subtipo Cargo; esta restricción se gestionará en el subtipo Cargo. La restricción a controlar aquí es que para un par Persona, AgrupacionAnio no se de una ocurrencia de tipo “Voz” si ya existe una de tipo “Orquesta”, y viceversa, puesto que son excluyentes el uno del otro; teniendo en cuenta que en esta abstracción se tiene el dato discriminador de la herencia.

Respecto a la integridad referencial, no se podrán eliminar tuplas de Persona ni de AgrupacionAnio si existen vínculos descritos en esta relación; por tanto, la desvinculación se hace sobre esta relación directamente.

Cargo, **Voz** y **Orquesta**

La restricción más importante en estas tres relaciones, es la multiplicidad de tipo semántico que se puede dar entre ellas, la cual ya ha sido controlada mediante la especificación de sus claves candidatas y el control de exclusión mutua entre Voz y Orquesta detallado en rol.

En lo referente a integridad referencial, se da en estos una peculiaridad debida al hecho de especializar una relación de cardinalidad N:M. En caso de eliminar una instancia de rol entre una persona y una agrupación, que según lo comentado en rol no se hará accidentalmente, se eliminará la instancia de cualquiera de estas relaciones que la referencie.

Orquesta → **Instrumento**

En cuanto a la integridad referencial, ante intentos de actualización o eliminación de un instrumento lo evitará, en caso contrario elimina en cascada.

usa

Para proteger la consistencia en cuanto al total de instrumentos asociados a una agrupación, se tendría que controlar que el mínimo sea el total de

instancias en rol Orquesta para la agrupación y el instrumento involucrados. Por simetría, se controlaría que en el rol Orquesta haya como máximo de tuplas el total aquí indicado, para el mismo instrumento; en consecuencia, de introducir roles a través de Orquesta podría actualizarse esta relación. No obstante, por el carácter paralelo, y meramente informativo, de Orquesta se descarta todo esto.

En cuanto a la integridad referencial, teniendo en cuenta que la eliminación de ocurrencias en AgrupacionAnio se controla en otras relaciones y niveles, de darse el caso de eliminación de agrupaciones se eliminará toda referencia a las misma contenidas aquí; mientras que, en lo referente a los instrumentos, se debe impedir la eliminación en Instrumentos de existir alguna tupla aquí. Esta decisión se basa en el hecho de que al eliminar una agrupación anual, si es que llega a darse el caso –puesto que muchos otros tipos de relación lo impedirán–, los instrumentos asociados sin componente son información accesoria, incluso parte intrínseca de la agrupación a eliminar; por tanto, no procede conservar una agrupación porque en ella se utilicen instrumentos, distinto es que la agrupación tenga premios de concurso o integrantes.

AgrupacionAnio → **Modalidad** y **Premio** → **Modalidad**

Por el carácter adjetivo de la entidad Modalidad respecto de las agrupaciones y los premios, de eliminarse una de estas se elimina la relación.

participa

Se debe controlar que el año y la categoría del concurso al que se asocia una agrupación, debe coincidir con el año de la agrupación, lo que podrá facilitarse en la futura interfaz; así mismo, para un año concreto y una modalidad no se pueden dar valores de premio, o clasificación, repetidos. La restricción de que las agrupaciones no se puedan presentar dos veces a un concurso se modelan en otro nivel, donde se permite la participación con diferente nombre y obra.

En lo relativo a la integridad referencial, por ser esta un tipo de relación asociativa, eliminar una agrupación o un concurso también está restringido, puesto que para eliminar una instancia de cualquiera de estos dos últimos conceptos habrá que retirar previamente cualquier tipo de información asociada, como es el hecho de participar. Cabría considerar, de forma análoga al caso de Agrupacion, si la interfaz de la BBDD debería facilitar la eliminación en cascada de conceptos que son dependientes de otros, como en este caso.

autoriaLetra (Letra ↔ Entidad), **autoriaMusica** (Musica ↔ Entidad) y **autoriaGrafica** (Disenio ↔ Entidad)

Las relaciones de autoría son obligatorias y no se les asigna un valor por defecto, para forzar la introducción; no obstante, se crearán dos entidades mínimas: la anónima y la desconocida. La integridad referencial, en este caso, restringirá la eliminación de entidades que se encuentren en condición de autor.

Empresa → **Actividad**

De nuevo, la única restricción es la de integridad referencial, como en el caso de la relación participa, si se elimina una Empresa se entiende que se deben eliminar las referencias a la misma de esta relación –considerando las actividades como adjetivos y que la eliminación de empresas se hará siempre que no tenga más información asociada por otras vías–.

crea

Unicamente la integridad referencial que evitará cualquier tipo de operación en las entidades que involucra esta relación, Disenio y Entidad.

Documento → Descriptor

En cuanto a la multiplicidad N:M, controlada mediante la creación de la clave candidata compuesta de los identificadores es estas entidades, hay que añadir que al menos un descriptor debe estar asociado al documento; por tanto, mínimo una instancia por cada documento debe existir en esta relación.

En cuanto a la integridad referencial, al igual que en otros casos de relaciones utilizadas para implementar cardinalidades N:M con carácter adjetivo, si se eliminan instancias en Documento se eliminarán las dependientes en esta; por el contrario, de tratar de eliminar un descriptor asociado a un documento se impedirá desde esta relación.

Documento↔ DepositoLegal

Respecto a integridad referencial, si se elimina un documento se eliminan las instancias de esta relación que lo referencien, por la relación de dependencia.

Documento↔ Evento

Un documento podrá asociarse a más de un evento y no es obligatoria la relación entre documentos y eventos; por tanto, en esta relación no debe haber un mínimo de instancias por cada documento registrado. Por otro lado, de no existir relación entre un documento y una obra o uno de sus componentes, esta relación debe contener como mínimo una instancia de dicho documento. En definitiva, por las premisas de diseño de nivel estructural y operativo que se han argumentado hasta este punto, las relaciones mínimas que debe tener una instancia de Documento ya han sido descritas, y deben ser controladas, dentro del ámbito y alcance de dicha entidad.

En cuanto a integridad referencial, de existir el tipo de relación dado por esta, no se eliminará ni el evento ni el documento; como en el resto de casos similares, habrá que desvincularlos en este nivel; es decir, específicamente mediante eliminación de la asociación.

Documento ↔ Obra

Mínimo debe haber una instancia por cada Obra en esta relación, la que representaría al documento que la describe necesariamente; es decir, debe haber como mínimo tantas instancias como instancias de Obra.

Por otra parte, según el tipo de documento hay que controlar con que subtipo de Obra debe esta vinculado; de nuevo, tal y como se acaba de comentar, se han detallado todas las particularidades dentro del nivel de la entidad Documento. Aunque, cabe añadir que, para todos los casos, sin obligatoriedad más allá del primero, por el que se introduce la Obra al sistema.

También cabe reiterar que todos los documentos contenidos en esta relación no tengan como valor de 3_1_4_NivelDescripcion, de la entidad Documento, “unidad documental simple”; exceptuando el tipo vídeo, donde podría darse la vinculación con obra gráfica por parte de una unidad documental simple –que sería una grabación de una pieza de audio concreta en formato vídeo–.

Documento↔**Letra**, **Documento**↔**Musica** y **Documento**↔**Disenio**

Debe existir una instancia en estas por cada una en Letra, Musica y Disenio, análogamente al caso del nivel de Obra; también como mínimo.

En cuanto a la integridad referencial, por la restricción de borrado y actualización que se da en estas relaciones, se evitará la eliminación del compuesto Obra o sus componentes sin antes deshacer el vínculo entre los mismos. Para eliminar el último documento tangible de un componente u Obra, puesto que hay una pareja de los mismos que tiene relación de existencia, habrá que retirar los datos que indican la URL o la localización del Documento, marcar como falsos los booleanos que indican si está digitalizado o en depósito y asignar como tipo de soporte “No disponible”. La gestión a nivel de interfaz de estas posibilidades, previa petición y confirmación consciente queda propuesta.

En cualquiera de los casos, todo esto lo gestionarán los procedimientos almacenados de baja. En principio, de eliminar un documento de un componente se eliminará la instancia de esta relación y posteriormente el documento, en caso de no tener más relaciones; mientras que, si se tratase de un documento único sin documento físico asociado, se restringirá porque se considera que lo que se intenta eliminar es la instancia del componente, sea Letra, Musica o Disenio; por tanto, se tendrá que utilizar la vía disponible para eliminaciones sobre dichas entidades.

ComposicionLiteraria, **ComposicionMusical** y **ComposicionGrafica**

La única decisión de diseño en cuanto a estas relaciones de composición es tener un identificador secuenciado que empiece por 1, para cada componente por cada uno de los compuestos; por ejemplo, para cada obra literaria se tendrán las piezas numeradas del 1 en adelante.

ComposicionAgrupacion

La única restricción es de integridad referencial, no se podrá eliminar una instancia de Agrupacion si existe alguna en AgrupacionAnio asociada.

GestionDocumento↔**Documento** y **GestionInformacion**↔**AgrupacionAnio**

Solo las interdependencias semánticas, ya comentadas a nivel de entidad.

Letra ↔ **Tema** y **Disenio** ↔ **Tema**

Solo de integridad referencial y con las peculiaridades de los casos que modelan tipos de relación de multiplicidad N:M con objetivos descriptivos; concretamente, de tratar de eliminar una tupla de la relación Tema se restringirá, por haber una letra o diseño que la utilice; mientras que, de eliminar una tupla en Letra o Disenio, se eliminarán aquí las tuplas que la referencien.

interpretaMusical (AgrupacionAnio ↔ Musical),
interpretaLiteraria (AgrupacionAnio ↔ Literaria) y
utilizaTipo (AgrupacionAnio ↔ Grafica)

Como restricción de integridad referencial habrá que evitar la eliminación de agrupaciones que interpreten o utilicen obras existentes, así como la eliminación de obras que estén vinculadas a alguna agrupación.

3.4 Modelo UML

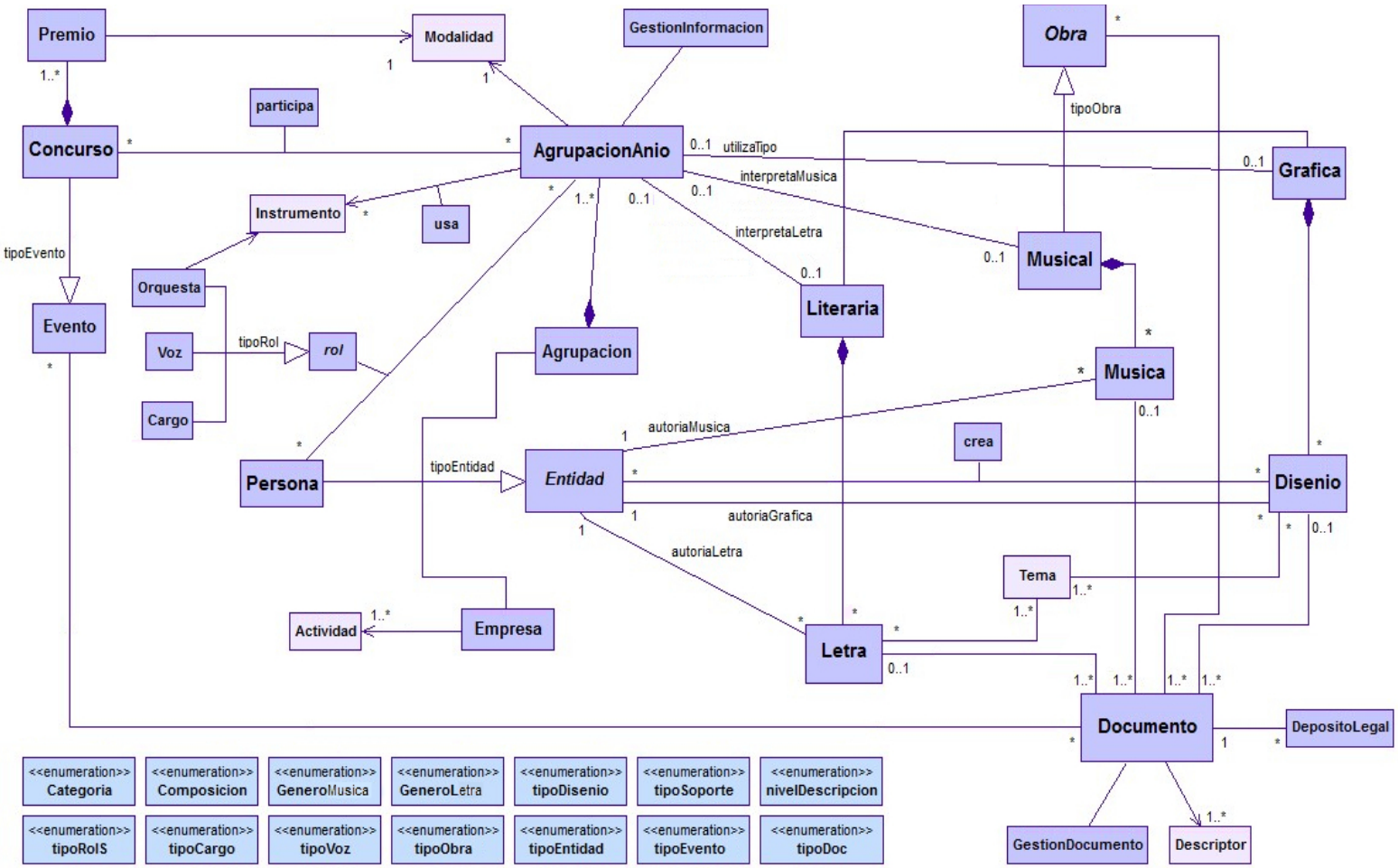

Figura 3. Modelo conceptual

3.5 Usuarios, perfiles y esquema externo

Para el alcance del presente proyecto solo se necesitan dos perfiles, indistintamente de los usuarios que se les asocien. Será a medida que se amplíe el SI, y esta BBDD, cuando será necesario ampliar este esquema externo en función a los perfiles de usuarios que surjan. Aun así, cabe recordar el contenido protegido por la LOPD que existirá en la BBDD, y ya se tiene en el conjunto de datos aportado por el cliente. La definición de estos dos roles y un par de usuarios de ejemplo se entrega en el *script* **04_roles.sql**.

En cuanto al esquema externo simplemente se crearán las vistas que el usuario solicita; tal y como se ha indicado, no se diferenciará entre el modelo accesible por los usuarios con rol de archivero y el accesible por los usuarios registrados en la AEPD, ya que queda fuera del alcance del presente proyecto. Además, se definirá un segundo esquema externo centrado en los estadísticos en este punto, aunque dichas funcionalidades aún no se han descrito.

3.5.1 Esquema Global

- Lista de Agrupacion junto con AgrupacionAnio, con todos los atributos.
- Lista de AgrupacionAnio junto con Persona por rol, con todos los atributos implicados excepto los de carácter personal.

- Lista de AgrupacionAnio junto con los tres subtipos Obra, más las instancias de Documento asociadas, también completas.
- Lista de Obra de tipo Literaria junto con sus componentes, más las instancias de Documento asociadas. Esta incluye la autoría de cada componente y los temas.
- Lista de Obra de tipo Musical junto con sus componentes y las instancias de Documento asociadas.
- Lista de Obra de tipo Grafica con los componentes y las instancias de Documento asociadas. Además de la referencia a la autoría de cada componente y la información contenida en la relación crea.
- Lista de Documento de tipo multimedia y su AgrupacionAnio, incluida la autoría y la dirección del grupo.
- Lista de Letra, Musica y Disenio junto con Tema.
- Lista de Evento con Documento.
- Lista de Concurso con AgrupacionAnio por participa, incluido Premio.
- Lista de Empresa con Actividad, más las creaciones mediante crea.

El código que genera estas vistas está contenido en el *script* **06_vistas.sql**.

3.5.2 Esquema *data warehouse*

- Lista de premios por concurso, agrupación y modalidad.
- Lista de premios por persona considerado el rol.
- Lista de dedicación oficial por persona, sin datos de carácter personal.
- Lista de temas recurrentes.
- Lista del estado de la documentación respecto a AgrupacionAnio y Obra.

Aunque se detallará en el apartado “6. *Data warehouse*” del presente documento, cabe hacer notar que se incluye como parte de esta funcionalidad la satisfacción del requisito RF04.7, que no fue identificado como de tipo resumen o precalculado inicialmente. Concretamente, mediante el último listado de este esquema, se satisface parte del requisito por el que se solicita disponer de información de soporte al proceso documental; de hecho, esta lista materializa dicha utilidad.

En este caso, el código por el que se crean estas últimas vistas se entrega en en el *script* **07_DWH_tablas_vistas.sql**.

4. Diseño lógico

Esta sección esta dedicada a la transformación del modelo conceptual al lógico, donde se da un breve comentario acerca de las decisiones de diseño consideradas de interés. Al igual que para el diseño conceptual, este modelo ha sido documentado al detalle de forma complementaria en el Anexo II.

4.1 Esquema

En el modelo lógico que se da a continuación se han reorganizado las relaciones atendiendo a la dependencia entre las mismas. En un solo caso se hace uso de claves foráneas a relaciones aún no definidas; concretamente, en AgrupacionAnio se apunta hacia las subclases de Obra. Esto es porque en el Anexo II se divide el modelo en dos zonas −Entidad y Evento, en primer lugar, y Obra y Documento, en segundo−, en la segunda reaparece dicha entidad y es remodelada conforme a tres tipos de relación de multiplicidad 1:M, donde AgrupacionAnio pasa a tener tres claves externas más de las que tenía en la primera zona descrita. En resumen, habría aquí una inconsistencia porque al declarar el total de atributos en AgrupacionAnio aun no se tienen todas sus dependencias.

Entidad (**ID**, **tipoEntidad**)

Agrupacion (**entidad**, **localidad, provincia,** alias, contacto, observaciones)
{entidad} es clave externa a Entidad (ID)

Modalidad (**modalidad**)

AgrupacionAnio (**entidad**, **nombre**, **anio**, **categoria**, **modalidad,** submodalidad, composicion, numComponentes, presupuesto, lugarEnsayo, observaciones, enPrensa, enDocumentoOficial, fuentes, observacionesGestion interpretaLiteraria, interpretaMusical, utilizaTipo)
{interpretaLiteraria} es clave externa a Literaria (obra)
{interpretaMusical} es clave externa a Musical (obra)
{utiliza} es clave externa a Grafica (obra)

Instrumento (**instrumento**)

usa (**instrumento, agrupacionAnio**, total)
{instrumento} es clave externa a Instrumento (instrumento)
{agrupacionAnio} es clave externa a AgrupacionAnio (nombre)

Evento (**ID**, **tipoEvento**, **anio, nombre, localidad, provincia**, ubicacion, descripcion)

Concurso (**evento**, **categoria**)
{evento} es clave externa a Evento (ID)

Premio (**ID, concurso, modalidad, premio**, cuantia, observaciones)
{concurso} es clave externa a Concurso (evento)
{modalidad} es clave externa a Modalidad (modalidad)

participa (**agrupacionAnio, concurso**, clasificacion, puntos, fase)
{agrupacionAnio} es clave externa a AgrupacionAnio (ID)
{concurso} es clave externa a Concurso (evento)

Persona (**entidad**, **nombre, apellido1**, apellido2, alias, DNI, telefono, domicilio, localidad, provincia, CP, profesion, fecNac, locNac, provNac, historia, observaciones)
{entidad} es clave externa a Entidad (ID)

rol (**ID**, **persona**, **agrupacionAnio**, **tipoRol**)
{persona} es clave externa a Persona (entidad)
{agrupacionAnio} es clave externa a AgrupacionAnio (ID)

Cargo (**rol, tipoCargo**, observaciones)
{rol} es clave externa a Rol (ID)

Voz (**rol**, **tipoVoz**)
{rol} es clave externa a Rol (ID)

Orquesta (**rol**, **instrumento**)
{rol} es clave externa a Rol (ID)
{instrumento} es clave externa a Instrumento (instrumento)

Empresa (**entidad**, **nombre, localidad, provincia**, CIF, domicilio, telefono, contacto)
{entidad} es clave externa a Entidad (ID)

Actividad (**actividad**)

ActividadEmpresa (**actividad**, **empresa**)
{actividad} es clave externa a Actividad (actividad)
{empresa} es clave externa a Empresa (entidad)

Documento (**ID**, **3_1_1_CodigoReferencia, 3_1_2_Titulo**, **3_1_3_Fecha_s, tipoDocumento, tipoSoporte**, ubicacion, deposito, estante, URL, {resto atributos ISAD(G)}, estaDescrito, estaDigitalizado, enDeposito, observacionesGestion)

Descriptor (**descriptor**)

DescriptorDocumento (**descriptor, documento**)
{descriptor} es clave externa a Descriptor (descriptor)
{documento} es clave externa a Documento (ID)

DepositoLegal (**documento**, **depositoLegal**, edicion, anio, generador)
{documento} es clave externa a Documento (ID)

DocumentoEvento (**documento**, **evento**)
{documento} es clave externa a Documento (ID)
{evento} es clave externa a Evento (ID)

Obra (**signatura**, **tipoObra**)

DocumentoObra (**documento, obra**)
{documento} es clave externa a Documento (ID)
{obra} es clave externa a Obra (signatura)

Literaria (**obra**, repertorio)
{obra} es clave externa a Obra (signatura)

Musical (**obra**, repertorio)
{obra} es clave externa a Obra (signatura)

Grafica (**obra**)
{obra} es clave externa a Obra (signatura)

Letra (**obraLiteraria**, **signatura**, **autoriaLetra, generoLetra**, transcripcion)
{obraLiteraria} es clave externa a Obra (signatura)
{autoriaL} es clave externa a Entidad (ID)
{modalidad} es clave externa a Modalidad (modalidad)

DocumentoLetra (**documento, letra**)
{documento} es clave externa a Documento (ID)
{letra} es clave externa a Letra (signatura)

Musica (**obraMusical, signatura**, **autoriaMusica, generoMusica**)
{obraMusical} es clave externa a Obra (signatura)
{autoriaM} es clave externa a Entidad (ID)

DocumentoMusica (**documento, musica**)
{documento} es clave externa a Documento (ID)
{musica} es clave externa a Musica (signatura)

Disenio (**obraGrafica, signatura**, **autoriaGrafica**, **tipoDisenio**)
{obraGrafica} es clave externa a Obra (signatura)
{autoriaG} es clave externa a Entidad (ID)

DocumentoDisenio (**documento, disenio**)
{documento} es clave externa a Documento (ID)
{disenio} es clave externa a Disenio (signatura)

Tema (**tema**, descripcion)

TemaLetra (**tema, letra**)
{letra} es clave externa a Letra (signatura)
{tema} es clave externa a Tema (tema)

TemaDisenio (**tema, disenio**)
{disenio} es clave externa a Disenio (signatura)
{tema} es clave externa a Tema (tema)

crea (**entidad, disenio**, coste, observaciones)
{entidad} es clave externa a Entidad (ID)
{disenio} es clave externa a Disenio (signatura)

A continuación se comentan las decisiones de diseño que se han tomado en esta fase, con respecto a las ya confirmadas para el modelo conceptual más las específicas de este. Previamente se comentan ciertas decisiones a nivel general, que son las siguientes:

- En cuanto a la denominación de las relaciones N:M que han surgido, se da el nombre anteponiendo el de la entidad vista como atributo[3]; por ejemplo, en el caso de la llamada ActividadEmpresa, se entiende Actividad como atributo o adjetivo de Empresa.
- Para las generalizaciones de Entidad y Obra se hace una transformación a relación por cada tipo de entidad descrita en el modelo conceptual; por un lado, por ser completa y disyunta y, por otro, porque el modelo se ha diseñado en este aspecto para poder vincular las superclases a otras entidades.

[3] "*Diseño de bases de datos*". Burgués, X.; Cabré i Segarra, B.; Casas, J.; Costal, D.; Juanola, P.; Ortego, S.; Pérez, J.; Rius, À.; Segret, R.

- Por último, los tipos de relación que involucran a Documento, que son todos de multiplicidad 1:N, donde un documento solo puede asociarse a una única tupla del extremo contrario de la relación, se modelan mediante relaciones intermedias. Se toma esta opción por no tener dentro de la relación Documento una clave foránea a cada posible relación de las que se han comentado, cuando en muchos casos no serán utilizadas. No obstante, se ha considerado la utilidad de este posible esquema para controlar los vínculos en función del tipo de documento indicado en tipoDocumento. En definitiva, se decidió esta estructura puesto que igualmente no se forzaría la existencia de todos los vínculos que implica cada tipo de documento.

Dicho esto se procede con el comentario de las decisiones específicas de cada relación resultante. Por claridad, como en casos anteriores, se omiten las que no implican complejidad en su transformación.

- Entidad (**ID**, **tipoEntidad**)
  - Simplemente se le añade el atributo discriminante de la herencia.
- Agrupacion(**entidad**,**localidad,provincia,**alias,contacto,observaciones)
  - La clave externa y primaria coinciden, por modelar la generalización.
- AgrupacionAnio (**entidad**, **nombre**, **anio**, **categoria**, **modalidad,**
  submodalidad, composicion, numComponentes,
  presupuesto, lugarEnsayo, observaciones, enPrensa,
  enDocumentoOficial, fuentes, observacionesGestion
  interpretaLiteraria, interpretaMusical, utilizaTipo)
  - El tipo de relación de composición, llamado en el modelo conceptual ComposicionAgrupacion, se modela con una clave externa hacia Agrupacion y la inclusión de un identificador propio del componente, en primera instancia será el atributo nombre teniendo en cuenta que es clave candidata en esta entidad por si solo.
  - La relación de esta con Modalidad se modela mediante una clave foránea, por ser un tipo de relación de multiplicidad 1:M donde la agrupación en un año solo se adscribe a una modalidad.
  - Por la dependencia de existencia y la obligatoriedad de tener una tupla en GestionInformacion por cada tupla de AgrupacionAnio, se decide modelar las dos entidades en una sola relación. En el modelo resultante se ha cambiado el nombre del atributo observaciones de la entidad GestionInformacion a observacionesGestion. Por otro lado, esta decisión de diseño no afecta en ningún caso al nivel de normalización de la relación –el atributo fuentes contendrá un bajo porcentaje de nulos porque la verificación es parte fundamental del proceso de catalogación–.
- usa (**instrumento, agrupacionAnio**, total)
  - Representa la transformación convencional de los tipos de relación con cardinalidad N:M, dos claves foráneas componen la primaria; donde, además, está el atributo total específico de la relación asociativa.

- Evento (**ID**, **tipoEvento**, **anio, nombre, localidad, provincia**, ubicacion, descripcion)
  - La primaria es de tipo autonumérica y tiene un atributo discriminante de generalización, análogo al definido en el concepto de Entidad; a parte, contiene una clave alternativa compuesta por los atributos estrictamente necesarios para la identificación unívoca de las tuplas.
- Concurso (**evento**, categoria)
  - La clave externa y la primaria coinciden como parte del modelo de generalización.
- Premio (**concurso, modalidad, premio**, cuantia, observaciones)
  - El tipo de relación de composición con Concurso se modela, desde el modelo conceptual al lógico, por la técnica común de transformación de las mismas. La clave primaria esta compuesta por el identificador de concurso, la modalidad y el premio, no es necesario a nivel semántico un identificador sintético porque este conjunto ya identifica unívocamente las tuplas.
  - El tipo de relación con respecto a Modalidad, de cardinalidad 1:M, donde cada premio pertenece a una única modalidad y que representa el valor de un atributo univaluado originalmente, se diseña mediante el uso de una clave foránea.
- participa (**agrupacionAnio, concurso**, clasificacion, puntos, fase)
  - Traducción convencional al esquema lógico de una relación de multiplicidad N:M. Cabría tener como clave candidata el compuesto de los atributos concurso y clasificacion, por no ser posible que dos agrupaciones queden en el mismo puesto; sin embargo, habría que añadir el concepto de modalidad para esto, el cual no está contenido en esta relación.
- Persona (**entidad**, **nombre, apellido1**, apellido2, alias, DNI, telefono, domicilio, localidad, provincia, CP, profesion, fecNac, locNac, provNac, historia, observaciones)
  - Se diseña como relación independiente vinculada a Entidad mediante clave foránea; la cual, además, es su clave primaria puesto que el objetivo inicial del modelo es que todas las tuplas de Entidad tengan un identificador único indiferentemente del subtipo.
  - Según los argumentos dados en el modelo conceptual, solo se considera clave alternativa al DNI que, aun admitiendo valores nulos, puede ser declarado de tipo único.
- rol (**ID**, **persona**, **agrupacionAnio**, **tipoRol**)
  - Se le añade el atributo discriminante de la generalización disyunta y completa, más un identificador sintético, para facilitar la vinculación de los subtipos de la que esta es superclase. Como relación de asociación de tipo N:M se tendría como clave primaria la compuesta por las referencias a las relaciones que vincula.

- El nuevo atributo autogenerado es más que paralelo al identificador compuesto por las dos claves foráneas; de hecho, estas no son clave alternativa por si solas, –por las razones expuestas en el modelo conceptual, permitir que una persona participe de una agrupación con más de un rol de tipo “Cargo”–, sino que requieren, además, del discriminante tipoRol para permitir que un par Persona y AgrupacionAnio puedan repetirse para distinto rol.

- Cargo (**rol**, **tipoCargo**, observaciones)
  - Como en el resto de relaciones de herencia, por heredar la clave primaria se garantiza su unicidad en caso de consultas de unión sobre los subtipos de rol.
  - La clave primaria es compuesta por los dos atributos, ya que se admite la repetición de la asociación dada por el identificador de rol, siempre que varíe el valor de tipoCargo.

- Voz (**rol**, **tipoVoz**)
  - La clave externa y la primaria coinciden como parte del modelo de generalización. Solo rol es clave primaria, de forma que un par Persona y AgrupacionAnio represente un único registro de voz.

- Orquesta (**rol**, **instrumento**)
  - Lo recién comentado para Voz. Por otra parte, el tipo de relación que identifica el instrumento utilizado por el rol Orquesta, se modela mediante una clave externa hacia Instrumento, por tratarse de un atributo univaluado.

- Empresa (**entidad**, **nombre, localidad, provincia**, CIF, domicilio,
  telefono, contacto)
  - Se le modelan las tres claves candidatas que tiene, siendo el identificador global heredado de la superclase Entidad el que actuará como clave primaria. Por otra parte, se tienen dos claves alternativas el conjunto nombre, localidad y provincia además del CIF, el cual admite valores nulos.

- ActividadEmpresa (**actividad**, **empresa**)
  - Se modela el atributo multivaluado actividad de las empresas como relación independiente; según la transformación convencional, se utiliza esta relación intermedia para modelar la cardinalidad N:M entre Empresa y Actividad, donde se hace uso de dos atributos que son claves foráneas a las mismas y componen la clave primaria.

- Documento (**ID**, **3_1_1_CodigoReferencia, 3_1_2_Titulo**,
  3_1_3_Fecha_s, tipoDocumento,
  tipoSoporte, ubicacion, deposito, estante, URL,
  {resto atributos ISAD(G)}, estaDescrito, estaDigitalizado,
  enDeposito, observacionesGestion)
  - Se cambian de posición los atributos que representan la clave primaria y los que son obligatorios, desplazándolos al principio de la

relación. Además, se le incluye un atributo numérico autogenerado a modo de identificador sintético, por eficiencia, ya que el atributo 3_1_1_CodigoReferencia es de tipo alfanumérico y previsiblemente grande como para actuar de clave primaria de forma óptima. Por otro lado, es el aludido para participar del último nivel en la signaturización contenida en el atributo anterior.

  - De nuevo en este caso, donde existe información complementaria dedicada a la gestión de la contenida en la entidad Documento original, se propone la inclusión de los atributos de la entidad GestionDocumento en esta; tanto por la dependencia en existencia como por la obligatoriedad y la cardinalidad 1:1. Al igual que antes se renombra el atributo observaciones a observacionesGestion.

- DescriptorDocumento (**documento**, **descriptor**)
  - Se traduce la relación N:M de forma común, mediante una relación auxiliar con claves externas que, a su vez, componen la primaria.

- DepositoLegal (**documento**, **depositoLegal**, edicion, anio, generador)
  - Esta relación de multiplicidad 1:N, que fuera un conjunto de atributos multivaluados e interdependientes incluidos en Documento, se transforma introduciéndole una clave externa a Documento.
  - La clave primaria está compuesta por la clave externa y el atributo depositoLegal, de manera que puedan existir varias tuplas de Depósito Legal por cada documento; además, este último atributo en sí se marca como clave alternativa puesto que no pueden existir dos números iguales de este concepto, aun refiriéndose a documentos diferentes.

- DocumentoEvento (**documento**, **evento**)
  - Traducción común de los tipo de relación de cardinalidad N:M.

- Obra (**signatura**, **tipoObra**)
  - Se le añade el atributo discriminador de la herencia disyunta y completa de la que esta relación es superclase; como en el caso de Entidad no admite nulos puesto que una Obra se crea directamente a través de sus subclases mediante la operativa interna.

- DocumentoObra (**documento, obra**)
  - Transformación típica de tipo de relación de multiplicidad N:M, mediante una clave foránea a la relación Documento y otra a Obra donde ambas conforman la clave primaria. Con este diseño se permite que un documento pueda asociarse a varias obras, lo cual es necesario para documentos de tipo multimedia con alcance a obra completa.

- Literaria (**obra**, repertorio)
  - Se le incluye el atributo obra, que es clave primaria y foránea, más representa la herencia en ambos casos.
- Musical (**obra**, repertorio)
  - Similar al anterior.

- Grafica (**obra**)
  - Similar a los anteriores.
- Letra (**obraLiteraria**, **signatura**, **autoriaLetra, generoLetra**, transcripcion)
  - Se incluye el atributo obraLiteraria como clave foránea para representar el tipo de relación de composición que esta relación tiene con la subclase Literaria. La clave primaria se compone de la anterior más una signatura propia.
  - Se crea atributo autoriaLetra para modelar el tipo de relación de multiplicidad 1:M, donde cada Letra debe tener un único autor y cada autor puede haber creado un número indeterminado de letras.
- DocumentoLetra (**documento, letra**)
  - Modela un tipo de relación de multiplicidad 1:M entre Letra y Documento; por esto se tiene como clave primaria únicamente la referencia a Documento. De esta forma, cada Documento puede referir a una única Letra, pero una Letra puede tener asociados varios documentos –necesario por los innumerables documentos de tipo multimedia que pueden asociarse a una sola letra–.
- Musica (**obraMusical, signatura**, **autoriaMusica, generoMusica**)
  - Se incluye el atributo obraMusical como clave foránea por el tipo de relación de composición respecto de las obras de tipo Musical. La clave primaria se compone de la anterior más la signatura propia.
  - Se crea atributo autoriaMusica para modelar el tipo de relación de multiplicidad 1:M, donde cada Musica debe tener un único autor y cada autor puede haber creado un número indeterminado de letras.
- DocumentoMusica (**documento, musica**)
  - Similar al diseño comentado para DocumentoLetra.
- Disenio (**obraGrafica, signatura**, **autoriaGrafica**, **tipoDisenio**)
  - Análogo a los diseños de Letra y Musica, con respecto a las obras de tipo Grafica, y con la clave externa hacia Entidad para modelar la autoría.
- DocumentoDisenio (**documento, disenio**)
  - Similar al diseño comentado para DocumentoLetra.

- crea (**entidad, disenio**, coste)
  - Análogo a la transformación de la entidad asociativa participa que se da entre AgrupacionAnio y Concurso; por otro lado, a diferencia de los tipos de relación que describen la autoría de las obras gráficas, en este caso la multiplicidad es N:M porque un solo diseño puede ser ejecutado por varias entidades, y una entidad puede ejecutar varios diseños; por tanto, se hace necesaria una relación intermedia.

4.2 Normalización

El proceso de normalización, ya aplicado al modelo lógico anterior, garantiza la corrección del mismo a nivel estructural y conceptual; concretamente, confirmará que todas las entidades están claramente definidas y diferenciadas del resto y que ninguna contiene conceptos ajenos a su semántica propia. La aplicación de las formas normales lo hacen comprensible, escalable y ausente de redundancias; además, y principalmente, evitan anomalías de edición; en contraposición, afectan a la eficiencia global del esquema, tanto por las combinaciones necesarias para consultas simples como por la complejidad que introduce para mantener la consistencia global.

A lo largo de todo el proceso de diseño se ha contado con estos condicionantes y, en ciertos casos, se decidió desde el comienzo no seguir las pautas de las diversas formas normales[3]; de hecho, se han diseñado algunos atributos semiredundantes de forma deliberada.

Se procede a comentar las premisas que conlleva aplicar cada una de las cinco formas normales, obviando que cada una implica la satisfacción de la anterior.

1FN Requiere que los atributos sean atómicos (indivisibles semánticamente) ni puedan ser una relación en si mismos –que implicaría una clave primaria adicional relativa la mismo–.

2FN Todo atributo que no pertenece a claves candidatas depende totalmente de ellas. Se hace más evidente donde las claves primarias no son compuestas.

3FN Todos los atributos dependen de la clave primaria y no hay dependencias transitivas; es decir, ningún atributo depende de otro, u otros, más directamente que de la clave primaria.

FNBC Todos los determinantes de todas las dependencias son claves candidatas de la relación. Suelen ser sensibles los casos en que se solapan claves candidatas, donde algunos atributos tienen mayor dependencia de una de las identificadas.

4FN Implica que no existen dependencias multivaluadas independientes. No debe haber ningún atributo atómico que contenga más de un valor (siendo, por tanto, divisible); y, menos aún, si afecta a otros atributos atómicos donde también se ha incluido más de un valor.

5FN Requiere que no existan dependencias de proyección-combinación, estando toda dependencia de unión basada en las claves primarias. Son más sensibles a esta forma normal las relaciones que involucran a más de dos entidades donde, por añadidura, hay simetría.

Excepto Persona, Empresa y Documento, todas las relaciones definidas satisfacen las cinco formas normales, incluida la de Boyce-Codd; todos los detalles se dan en el Anexo II, como parte del modelo lógico. Por otro lado, cabe mencionar que los campos dedicados a observaciones son considerados atómicos por su carácter incierto; y que, todos los que podrían haber sido aplanados o subdivididos, no lo han sido por considerarse accesorios.

A continuación se procede a detallar las razones por las que las relaciones anteriores no satisfacen las formas normales; en este comentario se desarrolla el análisis de las relaciones para todas las FN, a pesar de que la no satisfacción de alguna implique que las sucesoras dejan de ser aplicables.

---

[3] "*Diseño de bases de datos*". Burgués, X.; Cabré i Segarra, B.; Casas, J.; Costal, D.; Juanola, P.; Ortego, S.; Pérez, J.; Rius, À.; Segret, R.

- Persona

  Solo satisface la 2FN, 3FN, FNBC y 5FN. En cuanto a la 1FN, la mayoría de sus atributos son atómicos; excepto domicilio, teléfono, historia y observaciones, que por la semántica general del modelo y sus objetivos descriptivos, se consideran accesorios y autocontenidos –además, su nombre refiere claramente a su contenido y no se dan necesidades analíticas de tipo geográfico, más allá de la localidad–. Los atributos que no forman parte de las claves candidatas refieren a la persona en cuestión, por lo que satisface la 2FN. No hay dependencias transitivas, puesto que las dos candidatas tienen el mismo nivel de significación, más los dos determinantes complementarios podrían ser claves candidatas –que no lo son por la complejidad de incluir domicilio o solo hacer candidata a nombre y apellido1–; por tanto satisface la 3FN y la FNBC. Por otro lado, respecto a la 4FN se decide obviarla por no considerar telefono como multivaluado; se opta por la eficiencia que aporta eliminar la combinación que sería necesaria de modelarlo como multivaluado. Finalmente, todas las dependencias de unión están basadas en la clave primaria que, como se ha expuesto, es tan representativa de la relación como el resto de las claves candidatas.

- Empresa

  Ocurre algo similar, se toman las premisas ya expuestas respecto a los conceptos de contacto y teléfono; no obstante, se continua la evaluación teniendo en cuenta que estos defectos se asumen deliberadamente. Respecto a la 2FN, se dan tres determinantes y todos son claves candidatas, que a su vez son de igual nivel de significación; por tanto, no se dan dependencias transitivas. De nuevo, en cuanto a la 4FN, el atributo multivaluado actividad, que no se observa como componente de la relación, ya se ha modelado como relación intermedia de tipo N:M; por último, la dependencia de unión no puede generar redundancias por basarse unicamente en la clave primaria –sería el caso de unir todas las relaciones herederas de Entidad según sus atributos compartidos–.

- Documento

  Como se indicó en el modelo conceptual, esta relación por el carácter narrativo de sus atributos no cumple en muchos casos la 1FN, si cumple la 2FN, la 3FN y la 5FN, puesto que todos los atributos dependen de un único determinante y por ello, a su vez, no genera conflictos de combinación-proyección por tener una sola clave candidata con la que ser participe de estas operaciones. Respecto a la 4FN, se dan varios casos de valores multivaluados, como los productores, el historial archivístico, la documentación relacionada, etc.

Seguimiento y control

Aunque en el diagrama Gantt se refleja, cabe decir que: en este punto se inicia la creación de las tablas en el SGBD sobre el entorno de desarrollo ya en funcionamiento; y que, la finalización del diseño conceptual no se dio hasta estar finalizado el recién descrito; es decir, se ha refinado modelo conceptual durante todo el diseño lógico.

5. Diseño físico

El proceso completo de modelado, al igual que para los anteriores, se documenta con todo detalle mediante un documento externo, el Anexo III.

5.1 Nivel virtual

Teniendo en cuenta que la creación y gestión de los *tablespaces* se centra en cómo el SGBD hace uso del sistema de archivos físico de la maquina en que se ejecuta, se plantea tener las tablas de mayor porcentaje de campos de tipo texto de longitud indeterminada, concretamente Documentos y Letras, en dos *tablespaces* separados entre ellos y del resto de la BBDD, que estará en un tercero. Los dos primeros reciben el nombre de estas tablas y el tercero se denomina “global”. En el *script* de construcción de la BBDD, **01_database.sql**, están las sentencias DDL necesarias para crearlos.

Con este esquema de distribución de las tablas se pretende aumentar la eficiencia global mediante:

- La extracción de la tabla Documento, por el gran número de atributos nulos y todos de tipo texto de longitud indeterminada, propensos a contener un mínimo de un párrafo.
- La extracción de la tabla Letra que, aun solo conteniendo un campo de texto de gran longitud, concentrará la mayor parte de las búsquedas en este tipo de columnas.

De esta forma se paralelizarán los accesos a disco por parte del SGBD en caso de acceder a estas tablas tan pesadas. La ganancia de eficiencia no es tanto la introducida en dichas búsquedas, como la que se aporta al rendimiento global mediante el aislamiento del resto de tablas contenidas en el tercer *tablespace*.

Por otro lado, aunque fuera de la descripción de estos esquemas, se tendrá un *tablespace* para los índices y otro para las tablas del esquema de *data warehouse.*

5.2 Tablas

A continuación se hará un inciso acerca de las cuestiones consideradas de interés durante la definición de la tablas que corresponden al diagrama UML y, naturalmente, al modelo lógico. El detalle y las representaciones gráficas están contenidos en el Anexo III.

Se ordenan igual que en el modelo lógico; esto es así porque, a lo largo de esta descripción, el detalle que se da está centrado en las decisiones tomadas para la codificación SQL, y el *script* resultante, **02_tablas.sql**, necesita de dicho orden para que su ejecución no se vea afectada por las dependencias. Sin embargo, los comentarios de las tablas de *lookup*, por su simplicidad, y a las relaciones de cardinalidad N:M se comentan en conjunto.

En cuanto al uso de Camel y Pascal para los nombres, desde esta fase se desestima. Se convierte todo a minúsculas y, en los casos donde se componen de varias palabras, se utiliza un guión bajo como separador. Esto se decide porque, para mantener el esquema de nombres anterior, en PostgreSQL hay que utilizar comillas dobles, lo cual complica el uso y administración de la BBDD, así como resulta en un código muy ilegible.

Para definir las restricciones de integridad referencial se tienen ciertas condiciones que se hacen explícitas durante esta fase, pero que ya existían. Por lo general se tiene la restricción de actualización y borrado en las tablas referenciadas, excepto:

- Siempre que se cuente con tablas intermedias para las relaciones de tipo N:M se hará actualización y borrado en cascada cuando desaparezca el registro dominante; por ejemplo, en los casos en que las entidades débiles actúan de forma adjetiva.
- En las tablas Entidad, AgrupaciónAnio, Documento y Obra donde se admite la modificación del identificador, por la misma razón que el cliente solicita poder indicar dicho identificador. Toda columna dependiente de los identificadores de estas tablas será actualizada en cascada, más se continúa con la restricción en caso de borrado.

Entidad

El identificador hace uso de una estructura de tipo secuencia para generar los valores ordenados automáticamente; no obstante, de recibir alguno lo admitirá. Como ventaja se tiene que las estructuras de secuencia pueden ser actualizadas en cuanto al próximo valor a asignar.

Luego, desde este punto aclarar que todos los campos de tipo enumeración son obligatorios; para posibilitarlo, en los modelos anteriores, se proponían valores a modo de “se desconoce”. En ciertos casos podría flexibilizarse al respecto; pero, para los atributos discriminantes de herencia es vital que sean cumplimentados, aun siendo parte de la operativa interna de la BBDD.

Agrupacion

En este caso el identificador, entidad, se obtiene tras crear un registro en la anterior; por tanto, no hace uso de secuencias ni del tipo de dato *serial* de PostgreSQL, que es análogo a dicha estructura. Luego, el campo alias tiene una restricción de tipo *unique.*

Modalidad, **Instrumento** , **Actividad**, **Tema** **y** **Descriptor**

Los campos modalidad, instrumento, actividad, tema y descriptor, respectivamente, tienen una restricción de unicidad. Todas estas son tablas de tipo adjetivo y solo necesitan la restricción sobre el campo inicialmente propuesto como clave candidata. En cuanto a este último hecho, se dará por entendido que, en la declaración de las tablas, para todas las claves candidatas señaladas en el modelo lógico se define esta restricción.

AgrupacionAnio

Se crean las enumeraciones necesarias, llamadas enum_categoria y enum_composicion. Incluye una secuencia propia para su identificador, puesto que no va a ser compartido por varias entidades; además, esta se inicializa a alto valor para admitir la inserción por parte del usuario.

Por otro lado, las claves externas son entidad, para la herencia, y modalidad como adjetivo; asimismo, se tienen con restricción de unicidad el conjunto compuesto de nombre y entidad, más el nombre en si. Por último, se modelan restricciones de tipo *check* para los numéricos, principalmente porque son

todos naturales; más, en el caso del año, que tenga 4 dígitos –esto aplica para el resto de ocurrencias del concepto año que se utilizan más adelante–.

usa

Tiene la combinación de las claves externas definida como *unique*, por haberle introducido el identificador sintético; y, en cuanto al campo total, se establece que sea por defecto 0, más la restricción de que no sea negativo.

Evento

El identificador hace uso de una estructura de tipo secuencia, al igual que Entidad. Luego, para implementar la clave candidata se define una restricción de unicidad sobre anio, nombre y localidad.

Concurso

En este caso el identificador, evento, se obtiene al crear un registro en la anterior que, a su vez, es clave foránea; con esto, se modela la relación de herencia o generalización entre ambas entidades.

Premio

Se le añade un identificador sintético para reducir la complejidad, puesto que la clave primaria es compuesta del mismo más la externa, que modeliza la relación de composición respecto de Concurso.

Por otro lado, se crea una restricción de unicidad sobre el conjunto concurso, modalidad y premio, el cual es en realidad la clave identificativa de un registro.

participa

La clave primaria es compuesta de las externas y se añaden restricciones de tipo *check* sobre los numéricos, simplemente que sean positivos en este caso.

Persona

En esta tabla el identificador, entidad, se obtiene al crear un registro en la que este referencia; como es debido para toda generalización. Además, se crea una restricción de tipo *unique* sobre el DNI, que admite nulos y que es condición de este campo. En cuanto a la longitud de los tipos carácter se entendieron como suficientes.

rol

Se incluye un identificador complementario, de nuevo por la eficiencia que aporta el uso de una clave primaria simple y numérica respecto de una compuesta y alfanumérica. Naturalmente, esto implica que se defina una restricción de unicidad sobre el conjunto persona, agrupacion_anio y tipo_rol.

Cargo

Tiene como restricción de unicidad el identificador sintético de rol y el tipo de cargo, por esto se añade una clave primaria paralela; además, porque se permite la modificación del tipo de cargo, para ello hay que conocer el anterior aún sin recibirlo –esto hace necesario un identificador del registro independiente del par formado por rol y tipo_cargo.

Voz

Se utiliza el identificador heredado de rol como clave primaria y externa, porque no necesita admitir más de un registro por cada uno en rol.

Orquesta

Similar al anterior en cuanto a la clave primaria y la herencia; luego tiene el instrumento como clave foránea a la tabla que los contiene.

Empresa

La clave primaria la hereda de Entidad, generada por la secuencia asociada a la misma, y es clave externa para el modelado de la herencia. Además, tiene restricciones de unicidad sobre las que serían sus claves candidatas nombre, localidad y provincia más el cif, que además admite nulos. Cabe señalar que la admisión de nulos es una ventaja de las restricciones de tipo *unique*.

ActividadEmpresa

El par de claves externas se define como clave única, por la necesidad de crear un identificador sintético para la modificación de registros.

DepositoLegal

Su clave primaria es necesariamente compuesta por la relación de dependencia de multiplicidad 1:M; sin embargo, tal cual se expuso, requiere una restricción de tipo *unique* sobre el campo de deposito_legal a nivel de tabla. Por otro lado, se añade un identificador propio por no incluir a la clave primaria un campo de tipo texto.

Documento

En este caso la clave primaria hace uso de una secuencia inicializada en 100000, con objeto de admitir las signaturas que los usuarios del Aula de Cultura hayan ido asignando a los documentos a lo largo de los últimos años.

Al campo 3.1.1 Código de Referencia de la ISAD(G) se asignan 100 caracteres por la complejidad que se ha descrito en los modelos anteriores; así mismo, por razones obvias, tiene una restricción de tipo *unique.*

Obra

Al igual que Entidad, para controlar la unicidad de los identificadores de los subtipos de Obra se hace uso de una secuencia, además de imponer el valor del discriminante tipo_obra. También en este caso se inicializa a 100000 para permitir la introducción de identificadores de documentos ya catalogados.

DescriptorDocumento

Para facilitar la posibilidad de modificar los descriptores asociados a los documentos, se añade el identificador paralelo y se define el par de claves foráneas como clave única.

DocumentoEvento y **DocumentoObra**

En este caso se sigue el modelado natural de las relaciones de cardinalidad N:M, puesto que no admiten modificación, solo altas y bajas de registros; es decir, tienen como clave primaria el compuesto de las dos externas que involucran a Documento con Evento y Obra.

Literaria, Musical y **Grafica**

Por ser todas subtipos de Obra se diseñan de igual manera. Las claves primarias las heredan, generada por la secuencia asociada la signatura de Obra, y son clave externa para el modelado de la herencia.

Letra, **Musica** y **Disenio**

Aun teniendo diferencias semánticas, estos tres componentes de las Obras implican las mismas decisiones de diseño. Se introduce un identificador sintético para facilitar las referencias desde otras tablas; por otro lado, como parte del modelado de la relación de composición entre estas y Literaria, Musical y Grafica, respectivamente, se les crean restricciones de tipo *unique* sobre el par obra y signatura.

Respecto al valor de signatura, se establece el tipo *integer* sin secuencia asociada porque se gestionará la secuencialidad dentro de cada obra diferente; esto, como se expuso en el modelo conceptual, se gestionará mediante los procedimientos almacenados de alta. Cabe señalar que por esta misma razón ha sido necesario añadir un segundo identificador paralelo, para uso interno, al compuesto por el de la signatura de la obra y la propia del componente.

DocumentoLetra, **DocumentoMusica** y **DocumentoDisenio**

Estas relaciones vinculan las tres tablas recién comentadas respecto de la tabla Documento; la particularidad en este caso es que modelan relaciones de multiplicidad 1:N, por tanto la clave primaria no es compuesta, lo que controla que cada documento solo pueda apuntar a un componente de Obra.

TemaLetra, **TemaDisenio** y **crea**

Por último, se dan estas tres tablas intermedias para la implementación de relaciones de cardinalidad N:M; nuevamente, por haber incluido el identificador sintético no se tienen sus claves primarias compuestas por las foráneas, sino que se define el par de tipo único.

5.3 Índices

Este proceso se ha paralelizado con el de construcción y pruebas. La optimización esta centrada en la forma de construir las vistas principalmente; no obstante, el impacto global de cara a estas y al resto de posibles consultas está determinado por la correcta elección de los índices.

PostgreSQL crea automáticamente índices sobre las claves primarias y las columnas con restricción de unicidad –lo cual es común a muchos SGBD–; por esto, en cuanto al esquema de índices, solo resta definir estas estructuras en las claves foráneas, en campos de carácter temporal y en los que serán foco de búsquedas. Para decidir donde es pertinente utilizar índices es esencial encontrar el equilibrio entre:

- La carga computacional que introduce la gestión de estas estructuras.
- La eficiencia que aportarán en las consultas de combinación.

Finalmente, de forma explícita en el fichero **03_indices.sql**, se crean sobre: las claves externas, las columnas dedicadas a fechas, concretamente años, y las que serán foco de búsquedas; exceptuando las de tipo texto de longitud

indeterminada. Por otro lado, por ser clave primaria, no se crean índices en las claves externas que modelan las herencias. Cabe señalar que los índices extra generados sobre la tabla Documento y Letra son de gran utilidad por el tamaño y la densidad que se prevé que alcancen dichas tablas. Aún así, se dan excepciones atendiendo a las recomendaciones generales de definición de índices y a la probabilidad de búsqueda en los campos candidatos descartados.

5.4 Procedimientos almacenados

Durante el desarrollo de estos componentes de control surgieron diversas cuestiones que han llevado a tomar decisiones a cerca del uso global del modelo, las cuales se indicarán en el detalle de los mismos que se da seguidamente. Por otro lado, en ningún caso se ha hecho gestión de errores más allá de los que estos procedimientos tratan de controlar, que son los relativos a la semántica y el modo de uso del modelo; es decir, toda restricción ya implementada mediante el DDL convencional no se ha gestionado, por tanto se redirigen los errores tal cual los emite el SGBD y se gestionan los específicos del modelo.

Como resultado se tiene un conjunto por cada entidad editable y otro por cada atributo multivaluado, incluidas en estos últimos las relaciones de tipo N:M. Sin entrar al detalle de la implementación, que está contenido en el Anexo III, se comentarán las decisiones tomadas acerca del uso de la BBDD donde surgen.

Se procede a comentar la lista de procedimientos almacenados de alta, baja y modificación; más otros que, por un lado, se han hecho necesarios para permitir la satisfacción de ciertas necesidades generadas a lo largo de los procesos de modelado del sistema; y, por otro, los convenientes para la reducción y reutilización de código. Cada uno de los bloques que se listan está contenido en un *script* SQL identificado por el prefijo '**05_SP_**' y seguidos de un número que permite la ejecución ordenada de los mismos; luego, se identifican mediante dos guiones bajos los procedimientos que son de uso estrictamente interno –los cuales coinciden con los propuestos como forma de reutilización–.

En primer lugar, se desarrollaron los dedicados a las tablas de tipo *lookup*, contenidos en el *script* **05_SP_01_tablasLookup.sql**. Estos no conllevan ninguna complejidad, lo único mencionable es que se considera la capitalización de los valores para comprobar la unicidad de los mismos; se ha utilizado esta técnica también para los nombres en AgrupacionAnio.

- Alta_Modalidad, Modifica_Modalidad y Baja_Modalidad
- Alta_Descriptor, Modifica_Descriptor y Baja_Descriptor
- Alta_Tema, Modifica_Tema y Baja_Tema
- Alta_Instrumento, Modifica_Instrumento y Baja_Instrumento
- Alta_Actividad, Modifica_Actividad y Baja_Actividad

Por necesidad, el siguiente *script* desarrollado fue el de los procedimientos almacenados de la tabla Documento, denominado **05_SP_02_Documento.sql**, ya que muchos de los siguientes requieren de la gestión de documentos para verificarlos íntegramente. En este caso los procedimientos internos son todos

utilizados desde las funciones dedicadas a Obra y Evento, incluso desde los centrados en los componentes de obras.

- __Alta_Documento, __Modifica_Documento y __Baja_Documento
- Modifica_ID_Documento
- __Verifica_Codigo_Referencia y __Componer_Codigo_Referencia
- Alta_DepositoLegal_Documento, Modifica_DepositoLegal_Documento y Baja_DepositoLegal_Documento
- Alta_Descriptor_Documento, Modifica_Descriptor_Documento y Baja_Descriptor_Documento

Dentro del concepto de Documento se da una de las decisiones de implementación más notables, aunque se verá que análogas a las demás. Primero, se impedirá que los documentos puedan cambiar sus asociaciones entre los posibles conceptos a los que se vinculan –eventos, obras, letras, músicas y diseños–. Esto, a nivel técnico, solo implica omitir la definición de procedimientos almacenados de modificación sobre las tablas que relacionan en 1:N y N:M los conceptos mencionados; mientras que, a nivel lógico, requiere que el usuario de altas y bajas en las relaciones entre el documento y los objetos para realizar el cambio de vínculo sin la eliminación real del documento.

Segundo, se restringe la eliminación de documentos cuando el elemento al que se relaciona no tiene ninguno más, lo que fuerza la eliminación del elemento.

Por otro lado, como regla general en el uso de estas funciones internas, todos los que crean documentos gestionan la secuencia de identificadores de Documento y componen el Código de Referencia antes de la llamada, de forma que se controla la composición de signaturas según la CDU en función del tipo de entidad. En caso de recibir este valor se debe controlar que contenga las signaturas, en caso de lanzar el error se indican cuales se espera recibir; estas dos últimas necesidades se satisfacen mediante los dos procedimientos almacenados del tercer punto, referentes a composición y verificación.

Por no ser posible a otro nivel, se controla en el alta de documentos que se recibe un descriptor; luego, para el resto de descriptores o para modificarlos se hará uso de las funciones específicas. Así mismo, en la función dedicada a bajas de descriptores, se emite un error de ser el último asociado al documento indicado. Esta misma vía se utiliza para forzar el tema mínimo en letras y diseños o la actividad respecto a las empresas.

Finalmente, las funciones dedicadas a DepositoLegal y Descriptor tienen sentido adjetivo, en cierto grado, y la dedicada al cambio del identificador sencillamente corrige el Código de Referencia de ISAD(G) sustituyéndole el identificador anterior por el indicado. Con esto se gestiona el error humano debido al requisito de insertar el identificador del documento por el usuario en caso de existir.

Una vez implementadas estas funciones básicas se procedió con la jerarquía que parte de Entidad, contenida en el *script* **05_SP_03_Entidad.sql**. La herencia se implementa completamente en los subtipos de Entidad, mediante el uso de los siguientes procedimientos almacenados se controla que no haya actualización directa sobre dicha tabla; dichos procedimientos gestionarán el registro del que dependen, en la que representa a la superclase, y la secuencia asociada al identificador que comparten. Cabe señalar, a este respecto, que

PostgreSQL considera la ejecución de los procedimientos almacenados como una transacción, por hacer uso del bloque BEGIN/END; en consecuencia, se consideran actualizaciones atómicas.

- Alta_Empresa, Modifica_Empresa y Baja_Empresa
- Alta_Actividad_Empresa, Modifica_Actividad_Empresa y Baja_Actividad_Empresa

  En cuanto a estas, lo único a comentar es que se da la necesidad de incluir identificadores en las tablas N:M que permitan modificar las parejas de claves externas que forman estas relaciones. Concretamente, en los casos en que se consideran relaciones de tipo adjetiva, se admitirá solo el cambio del valor considerado adjetivo; y, por otro lado, solo se necesitará el identificador sintético añadido para admitir la edición. Esta decisión de diseño aplica a todos los casos de carácter adjetivo que se identifiquen.
- Alta_Persona, Modifica_Persona y Baja_Persona
- Alta_Agrupacion, Modifica_Agrupacion y Baja_Agrupacion
- Modifica_ID_Agrupacion

En lo referente a integridad referencial, por estar la autoría a nivel de Entidad, y según la atomicidad que PostgreSQL infiere a las funciones, se podrá eliminar cualquiera de los subtipos pero, en caso de haber relación por autoría, se interrumpirá al tratar de eliminar el registro en la tabla que representa la superclase abstracta.

Por otro lado, para la implementación de las restricciones semánticas se hace uso de estructuras de control básicas. Y, finalmente, se provee de la función necesaria para modificar el identificador de agrupación, el cual se implementa también en el nivel anual de estas; en sí no conlleva complejidad, solo ha sido necesario que las restricción de integridad tenga definida la actualización en cascada.

Seguidamente se implementaron los procedimientos almacenados necesarios para la jerarquía de Evento, contenidos en el *script* **05_SP_04_Evento.sql**. Una de las particularidades es que, desde el conjunto dedicado a la tabla Evento, no se pueden editar los eventos de tipo “Concurso”. La otra es que no se permitirá la eliminación en Premio si existe alguna alusión al mismo en participa, lo cual implica una restricción en una relación indirecta que es necesario forzar en este nivel –el resto están todas en el nivel del DDL–.

- Alta_Evento, Modifica_Evento y Baja_Evento
- Alta_Concurso, Modifica_Concurso y Baja_Concurso
- Alta_Premio_Concurso, Modifica_Premio_Concurso y Baja_Premio_Concurso
- Doc_Nuevo_a_Evento y Modifica_Documento_Evento
- Doc_Existente_A_Evento y Baja_Documento_Evento

  De estos cuatro últimos, los dos primeros implementan la inserción y actualización en Documento, a la vez que registran la relación con el evento indicado, y los dos siguientes gestionan únicamente vínculos definidos en la tabla DocumentoEvento.

Tal y como se ha introducido anteriormente, la idea es que no se pueda cambiar el vínculo de un documento entre objetos, sino a los objetos cambiarles los documentos. En resumen, habrá que dar alta en un objeto a un documento existente y, acto seguido, la baja en el objeto donde dicho documento no corresponda. Esto se haría mediante las dos últimas funcionalidades de la lista anterior.

A continuación, una vez probados los anteriores, se procede con los procedimientos centrados en las agrupaciones de carácter anual, contenidos en el *script* **05_SP_05_AgrupacionAnio.sql**. Al igual que para el concepto de Evento, este *script* contiene todas los procedimientos para la tabla principal y para todas las directamente relacionadas.

- Alta_AgrupacionAnio, Modifica_AgrupacionAnio y Baja_AgrupacionAnio

  Como facilidad para el uso de la BBDD se admite la creación de registros en Agrupacion desde este procedimiento de alta, concretamente de no recibir el identificador de la misma se requiere la localidad y provincia y con estas se da alta a una nueva agrupación. No obstante, no se considera una decisión tomada en esta fase; al igual que tampoco lo es la verificación de no participación en un concurso, por ejemplo a la hora de modificar el año, la categoría o la modalidad de la agrupación de carácter anual.
- __Verifica_AgrupacionAnio_Obra
- Baja_Literaria_AgrupacionAnio, Baja_Musical_AgrupacionAnio y Baja_Tipo_AgrupacionAnio
- Modifica_ID_AgrupacionAnio
- Alta_Agrupacion_Usa_Instrumento, Modifica_Agrupacion_Usa_Instrumento, Baja_Agrupacion_Usa_Instrumento
- __Alta_Rol
- Alta_Cargo_AgrupacionAnio, Modifica_Cargo_AgrupacionAnio, Baja_Cargo_AgrupacionAnio
- Alta_Voz_AgrupacionAnio, Modifica_Voz_AgrupacionAnio y Baja_Voz_AgrupacionAnio
- Alta_Orquesta_AgrupacionAnio, Modifica_Orquesta_AgrupacionAnio y Baja_Orquesta_AgrupacionAnio
- Alta_Participa, Modifica_Participa y Baja_Participa

No se han dado nuevas decisiones que afecten al modelo general durante el desarrollo de este conjunto de procedimientos almacenados, más allá de las puramente técnicas. Por ejemplo: la extracción del código necesario para las altas en la tabla rol que representa una clase abstracta; la necesidad de crear funciones para establecer el valor nulo respecto a las obras que se interpretan –porque la de modificación, al usar *coalesce* en las sentencias *update*, nunca introducirá un valor nulo en ningún campo–; y, aún implementando una restricción semántica, la dedicada a comprobar que el nombre de las obras asignada en interpretación contienen el nombre de la agrupación más las

necesarias para la correcta configuración de roles –que se controlan entre el DDL y estos procedimientos almacenados–.

En este punto se comienza con las funciones dedicadas a la zona derecha del gráfico que describe el modelo conceptual, según se dispuso en su momento; primeramente, se desarrollan las centradas en las obras y luego las dedicadas a cada uno de los componentes de cada subtipo de obra. Como en el caso de los documentos, se extrae todo el código común a un *script* independiente, denominado **05_SP_06_Obra.sql**.

- __Alta_Obra, __Modifica_Obra y __Baja_Obra

  En estos se gestiona la creación y vinculación del documento mínimo asociado a la obra, del tipo que sea; así como la corrección del Código de Referencia.

- __Modifica_Signatura_Obra

  Esta tiene más complejidad que las anteriores dedicadas a cambios de identificador, concretamente porque cada documento de cada componente contiene la signatura de obra que se trata de editar; por tanto, se recorre la tabla de documentos en busca de los Códigos de Referencia que contengan la signatura en cuestión para la sustitución de la misma. Luego la actualización de la signatura, donde sea necesario, se hace por la restricción de integridad referencial que contienen todas las claves foráneas que la utilizan como destino.

- __Doc_Nuevo_a_Obra

- __Alta_Documento_Obra y __Baja_Documento_Obra

  Se obvia la implementación de __Doc_Existente_a_Obra porque es exactamente la función que provee la primera de estas dos, ambas centradas en la tabla DocumentoObra.

- __Doc_Nuevo_a_Componente y __Baja_Documento_Componente

  Al igual que las anteriores, estas hacen uso de las funciones internas del *script* **05_SP_02_Documento.sql**; aunque de forma más limitada, porque cada componente de obra tiene su propia tabla para la relación N:M con Documento. Por otro lado, y obviamente, en las dos últimas relativas a las bajas de documento, se gestionan las restricciones pertinentes acerca de eliminar o no los documentos si quedan totalmente desasociados dentro de la BBDD, o tienen algún material tangible descrito a través del registro en concreto.

Seguidamente, por su similitud, se dan en conjunto las funciones desarrolladas para alta, baja y modificación de los tres subtipos de Obra, contenidas en los *scripts* **05_SP_061_ObraLiteraria.sql**, **05_SP_062_ObraMusical.sql** y **05_SP_063_ObraGrafica.sql**. La lógica es prácticamente idéntica en todos ellos, solo difiere la tabla sobre la que actúan; es decir, la implementación de cada procedimiento almacenado está casi completamente contenida en el *script* anterior.

- Alta_Literaria, Modifica_Literaria y Baja_Literaria

- Modifica_Signatura_Literaria

- Doc_Nuevo_a_Literaria, Doc_Existente_a_Literaria y Baja_Documento_Literaria
- Alta_Musical, Modifica_Musical y Baja_Musical
- Modifica_Signatura_Musical
- Doc_Nuevo_a_Musical, Doc_Existente_a_Musical, Modifica_Documento_Musical y Baja_Documento_Musical
- Alta_Grafica, Modifica_Grafica y Baja_Grafica
- Modifica_Signatura_Grafica
- Doc_Nuevo_a_Grafica, Doc_Existente_A_Grafica, Modifica_Documento_Grafica y Baja_Documento_Grafica

Finalmente, se desarrollaron las funciones dedicadas a los componentes, aunque aquí de nuevo se listan juntas, se entregan en otros tres *scripts*: **05_SP_071_Letra.sql**, **05_SP_072_Musica.sql** y **05_SP_073_Disenio.sql**. Como en el caso anterior son análogas entre ellas, a diferencia de las relaciones de que participan, todas con una implementación sencilla sobre tablas con formato N:M e incluyen las decisiones ya descritas a cerca de las relaciones de carácter adjetivo.

Por otra parte, como se indicaba en el detalle de las funciones __Doc_Nuevo_a_Componente, y __Baja_Documento_Componente, en cada uno de estos *scripts* se hace una mayor gestión de las restricciones de usuario o dominio y, por ende, un uso directo de las funciones contenidas en **05_SP_02_Documento.sql**. Por ejemplo, no todos estos componentes de las obras realizan las mismas verificaciones ante las altas y bajas de sus documentos asociados, principalmente atendiendo al tipo de documento.

- Alta_Letra, Modifica_Letra y Baja_Letra

  En las altas y modificaciones, por no ser definible en la declaración de la tabla, se controla que no exista más de un elemento de género ‘Presentación’ y otro ‘Popurrí’ por cada obra literaria.
- Alta_Tema_Letra, Modifica_Tema_Letra y Baja_Tema_Letra
- Doc_Nuevo_a_Letra, __Alta_Documento_Letra, Doc_Existente_a_Letra y Baja_Documento_Letra
- Alta_Musica, Modifica_Musica y Baja_Musica
- Doc_Nuevo_a_Musica, __Alta_Documento_Musica, Doc_Existente_a_Musica y Baja_Documento_Musica
- Alta_Disenio_Obra, Modifica_Disenio_Obra y Baja_Disenio_Obra
- Alta_Tema_Disenio, Modifica_Tema_Disenio y Baja_Tema_Disenio
- Alta_Crea_Disenio, Modifica_Crea_Disenio y Baja_Crea_Disenio
- Doc_Nuevo_a_Disenio, __Alta_Documento_Disenio, Doc_Existente_a_Disenio y Baja_Documento_Disenio

5.5 Construcción

La implementación del código detallado a lo largo de esta memoria se realiza sobre un servidor PostgreSQL virtualizado mediante el sistema de contenedores Docker. En el primer apartado del Anexo III se da un breve comentario acerca del uso de Docker, el mínimo necesario para poner en marcha el entorno de desarrollo. Por otro lado, como interfaz se ha utilizado la utilidad PGAdmin 4, la cual también se virtualiza en un segundo contenedor. Una vez iniciados los dos contenedores se hace uso de un navegador web para la ejecución del contenido del *script* de creación de tablas.

Seguimiento y control

La implementación de las restricciones de usuario han llevado a tomar las decisiones mencionadas, unas a cerca del uso global del modelo y otras de nivel técnico, las cuales no se habían considerado hasta el momento o resultaban implícitas.

Por otro lado, en cuanto a la gestión del tiempo se ha tenido que ajustar de forma gradual, así como se iban obteniendo los resultados y alcanzando los hitos. Esta problemática se vio fuertemente afectada por los procesos de pruebas incrementales. Por añadidura, en una segunda fase dedicada únicamente a pruebas, salieron a relucir aún más inconsistencias, de más alto nivel; concretamente, referentes a la interacción entre los conceptos.

La consecución de los productos descritos hasta este punto se da con varios días de atraso y un número de horas que casi duplica al estimado en la planificación inicial.

6. *Data warehouse*

El desarrollo de las funcionalidades estadísticas se basa en el conjunto de resúmenes de datos solicitados por el cliente, descritos como parte del requisito funcional RF05; y la definición, ya refinada, propuesta como tercer esquema externo en el apartado “3.1.3 Esquema *data warehouse*”. Al igual que para los modelos, el detalle completo del desarrollo del sistema de *data warehouse* se realiza sobre un documento complementario, el Anexo IV.

Dentro de este ámbito, cabe mencionar que las intenciones no están centradas en medir rendimientos, sino en acumular eventos determinados; por tanto, se hará uso de las denominadas tablas de cobertura[6].

Seguidamente se detalla la propuesta a nivel lógico de las tablas de hechos[7][8] y, en el siguiente apartado, se comentan los objetivos a satisfacer y las decisiones de diseño a partir de las vistas que componen el esquema externo.

6.1 Diseño lógico

Para la provisión de los 5 tipos de resúmenes de datos se plantean 4 tablas de hechos; aunque resultan en 5, ya que la denominada DWH_Entidad es utilizada a modo de *lookup* por la siguiente; por otro lado, la segunda satisface dos de los requisitos, dedicación y premios por entidad, puesto que ambos están centrados en las entidades.

DWH_PremiosConcurso (anio, concurso, localidad, categoria, modalidad, agrupacionAnio, autorLetra, autorMusica, localidadAgrup, premio)

DWH_Entidad (**ID**, tipoEntidad, /nombre, localidad, provincia)

DWH_FuncionesEntidad (entidad, funcion, primero, segundo, tercero, cuarto, veces, anioInicio, anioFin)
{entidad} es clave externa a DWH_Entidad (ID)

DWH_TemasRecurrentes (tema, tipoObra, anio, modalidad, categoria, total)

DWH_SoporteDocumental (agrupacion, tipoObra, tipoDocumento, total)

Decisiones de diseño y grado de desnormalización:

- DWH_PremiosConcurso. Se incluye la localidad del concurso por formar parte de la clave candidata de los mismos, más la de las agrupaciones por interés analítico. No se plantea extraer la redundancia que se generará en el par concurso, localidad porque como máximo se darán unas 50 repeticiones; esto último, derivado de la combinación entre categorías, modalidades y el total de premios del concurso en concreto.
- DWH_FuncionesEntidad. Con un atributo por cada premio, en lugar de una tupla por cada ocurrencia de premio y de participación, puesto que esta tabla de hechos satisfará los requisitos RF05.2 y RF05.3. En este caso se extraen las datos de las entidades a modo de dimensión[8] de la

[6] “*Diseño de un data warehouse: tabla de hecho*”. Curto, J.
[7] “*Mejores Prácticas para construir un Data Warehouse*”. Escarcega, B.
[8] “*3.4 Datawarehouse manager*”. Bernabeu, D.

tabla de hechos en DWH_Entidad, principalmente porque la lista mostrará concatenados los datos que componen la clave candidata del subtipo de entidad que sea; por tanto, si en la tabla de origen uno de estos datos fuera actualizado crearía inconsistencias en estos resultados. Por otro lado, una persona rara vez cambia más de tres o cuatro veces de rol, por lo que la combinación entre las relaciones basada en índice será eficiente.

- DWH_TemasRecurrentes. Este también contiene redundancia, el valor de tema se dará una vez por cada tipo de obra (Letra y Disenio), año, modalidad y categoría. Cabe recordar que las modalidades realmente utilizadas son un total de 5; de hecho, documentadas oficialmente 4; además, el valor tema es bastante corto.
- DWH_SoporteDocumentalAgrupaciones. Con esta tabla de hechos simplemente se pretende acceder a la tipología de documentos, correspondientes a obras, asociados a las agrupaciones. La redundancia deviene del atributo agrupación, que se repetirá tantas veces como tipos de documentos se gestionan en el dominio.

En definitiva, se tiene un grado de desnormalización bastante alto, lo único que se conserva es la atomicidad en los atributos, la cual disminuirá en la definición de las vistas.

6.3 Implementación y esquemas externos

Para la construcción se plantea el uso de un último *tablespace*, tanto por estructuración como por aislar los datos operativos de este nuevo conjunto. En esta ocasión se entrega un *script* SQL, **07_DWH_tablas_vistas.sql**, que contiene este componente, las tablas y las vistas; y, un conjunto con prefijo **08_DWH_SP_**, con los procedimientos almacenados y los disparadores de cada una de las funcionalidades que se describen continuación.

En cuanto a los *triggers* serán lanzados siempre tras la inserción, actualización o el borrado en las tablas que contienen los eventos de interés a registrar; estos únicamente ejecutarán los procedimientos almacenados que procesarán los datos antes de introducirlos en las tablas de hechos recién descritas. Se describe a continuación la vista final y el proceso de recolección asociado a cada una de las relaciones indicadas anteriormente. Cabe señalar que mediante este diseño físico basado en disparadores se aísla la funcionalidad *data warehouse* de la interfaz lógica definida mediante los procedimientos almacenados; es decir, de caer estos en desuso, aún continuaría en funcionamiento el modelo estadístico.

6.3.1 Agrupaciones ganadoras por concurso y año

Considerando que en la actualidad, cada año participan alrededor de 150 agrupaciones en el COAC, generar estos datos dentro del concepto de *data warehouse* es más que pertinente, más aun teniendo en cuenta la necesidad de acceder a las relaciones AgrupacionAnio, participa, Concurso y Premio; además de localizar la autoría predominante a partir del conjunto de Letras interpretadas por cada agrupación.

Implementación

Un primer disparador asociado a la tabla participa llamará al procedimiento encargado de poblar esta relación, el cual recibirá los valores recién registrados en dicha tabla. Si el valor de clasificacion es menor que 4 procederá a localizar el resto de valores, a partir de las claves externas agrupacion_anio y concurso. Del concurso se tomará el año, el nombre y la localidad; en cuanto a los premios, se deben gestionar los casos en que quedan desiertos; y, respecto a las agrupaciones se debe incluir el autor, para lo que se calculará cual es el principal por mayoría.

Habrá otros dos disparadores auxiliares que gestionarán la consistencia entre los datos de las agrupaciones y los eventos implicados en caso de cambiar en el origen.

Representación

La siguiente vista genera una lista ordenada por año y concurso, luego por modalidad y premio.

<anio, concurso, localidad, modalidad, premio, agrupacion, autores, localidad>

6.3.2 Premios por entidad y función

Este es un caso de información de interés estadístico de cara a la contratación, considerando el carácter anual de las agrupaciones, son las personas y agrupaciones de nivel de entidad quién acumulan los premios; además, permitiría averiguar el prestigio de los participantes en función de la actividad que desarrollan. En principio se tiene como suficiente gestionar contadores por entidad y rol específico para los premios 1º, 2º, 3º y 4º.

En el requisito funcional RF05.2 se hace alusión a las personas, agrupaciones y empresas; una vez realizado el diseño completo de la BBDD, se tiene la posibilidad de utilizar simplemente el identificador de nivel Entidad, puesto que las anteriores son subtipos de esta última relación. Por otra parte, se introduce el concepto de funcion para las entidades, dadas las posibilidades del modelo obtenido en este punto; se consideran funciones: el rol entre personas y agrupaciones, la autoría entre entidades y obras y la ejecución de los diseños. Respecto a esto último, hay un concepto no incluido en el alcance: el premio “Aguja de Oro”, que si será admitido por este resumen de datos cuando el modelo se amplíe a los eventos de tipo premio.

Implementación

El diseño de esta funcionalidad se basa en la estructura física que se comentará para la siguiente, centrada en la dedicación de las personas; es decir, compartirán la tabla de hechos por los campos y el tipo de contenidos que requieren ambos objetivos. Considerada dicha estructura, para gestionar la información relativa a premios se tienen tres contadores que serán incrementados por un disparador asociado a la tabla DWH_PremiosConcurso atendiendo al rol de la persona para la agrupación dada.

Con el *trigger* activado desde DWH_PremiosConcurso solo se recibirán registros con premios menores al quinto –que se gestiona en dicha tabla– y se reduce la carga en el lado de la BBDD operacional. El procedimiento

almacenado tendrá que localizar todas las entidades relacionadas con la agrupación premiada e incrementar el acumulador del premio correspondiente.

Por otro lado, en caso de dar alta roles respecto de agrupaciones ya premiadas, se debe de averiguar el premio de la misma al gestionar la nueva relación entre persona y agrupación. Esta acción será ejecutada como parte de la siguiente funcionalidad estadística centrada en la dedicación de las entidades a las fiestas de carnaval.

Representación

La consulta consistirá en combinar las dos relaciones implicadas, utilizando el campo nombre de la relación DWH_Entidad, que contendrá el concatenado del resto en función del tipo de entidad, y la proyección de los atributos indicados en esta estructura, que no implican más operaciones ni cálculos.

<entidad, funcion, primeros, segundos, terceros, cuartos>

6.3.3 Dedicación oficial al Carnaval por persona

La intención en este caso es conocer la trayectoria de de los participantes oficiales en el Carnaval de Cádiz, concretamente el porcentaje de participación. Para esto solo se considera necesario conocer el nombre, apellidos y alias más el número de veces que cada persona ha intervenido en el carnaval; por tanto, además de las instancias de rol que se creen, habrá que recoger las de autoría y las de creación de diseños. Por dar más coherencia al proceso analítico se introduce el valor del año de la primera aparición de la persona y el valor del último, teniendo en cuenta que se conservan datos desde en siglo XIX.

Cabe aclarar, con base en este detalle que la autoría contendrá como valor el total de letras, músicas y diseños considerándolos individualmente; de hecho, en este concepto es de interés el género y la modalidad de componentes de obra y agrupación, respectivamente, a la hora de componer y diseñar.

Implementación

Se crea la tabla DWH_Entidad para poder sincronizar los valores de la tabla Entidad de la BBDD y la de dimensión; por tanto, únicamente es necesario un *trigger* por cada subtipo de entidad que, cuando haya acciones de inserción o actualización en los mismos, envíen a esta los datos registrados en aquellas.

Luego, la tabla DWH_FuncionesEntidad se actualizará por cuatro vías; en primer lugar, por cada inserción en las tablas herederas de rol de la BBDD se creará en esta una tupla que indique el tipo de rol con el atributo conveniente al mismo, el año de la participación y el valor 1 en el atributo veces, esto si la tupla no existe; en caso contrario, se sumará 1 al valor de veces. En segundo lugar, al asignar una autoría a una Letra, Musica o Disenio se procederá de forma similar; y, en tercero, por cada nuevo registro en la tabla crea. La cuarta vía es la ya descrita para el registro de premios.

Una peculiaridad al respecto es que si una agrupación es premiada pero no se le conocen roles no se podrán registrar los premios; para gestionar este hecho se accederá a la tabla participa, en caso de registro de nueva entidad en DWH_FuncionesEntidad, para gestionar los premios que la misma ya tenga asociados por la participación. Cabe hacer notar que el acceso a participa

puede ser directo por la agrupación del rol o, indirecto, por la autoría del componente, siempre que la obra tenga interprete; es más, al asociar una obra a una agrupación requerirá de realizar esta comprobación.

Representación

Como consulta a la relación DWH_FuncionesEntidad es similar a la anterior, dedicada a premios, solo que centrada en otros campos y con otro sentido.

<entidad, funcion, veces, anioInicio, anioFin>

6.3.4 Temas recurrentes en letras por año y modalidad

Para los temas incluir el año es de interés por el contraste histórico, teniendo en cuenta que se pueden hacer consultas al mismo conjunto de datos obviando el año, se obtiene una doble utilidad. Por otro lado, hay que diferenciar entre los temas que etiquetan a las letras y los que etiquetan los tipos que adoptan las agrupaciones; así como la modalidad, en el caso de las chirigotas y los cuartetos se adopta un carácter humorístico, mientras que en las comparsas es más serio; por tanto, la modalidad es de interés para la consulta de letras en función del tema.

Implementación

Para este caso se necesita un *trigger* asociado a las tablas TemaLetra y TemaDisenio, que enviarán al procedimiento dedicado a esta entidad del *data warehouse* el identificador del tema, el componente de Obra etiquetado mediante el mismo más el tipo de componente en cuestión, Letra o Disenio. Luego el procedimiento almacenado se encargará de recuperar el resto de valores y crear el nuevo registro. Esto, como en el caso anterior basado en acumuladores, será si no existe ya una tupla con dichos valores, en caso contrario se aumentará en una unidad el atributo total.

Debido a la inclusión del año en el conjunto de datos existe dependencia del concepto de interpretación, un componente que no está asociado a una agrupación no dispone de modalidad ni fecha; en consecuencia, como en casos anteriores y en el siguiente, de no existir agrupación asociada no se gestionarán estos estadísticos. Esto se solventa activando un *trigger* en AgrupacionAnio, en caso que se establezca alguna obra literaria o gráfica, se recorren todos los componentes de dicha obra y se actualizan los contadores por cada tema.

Representación

<tema, tipoObra, anio, modalidad, categoría, total>

En esta consulta no es necesario ningún tipo de operación de combinación ni agregación; por tanto, con respuestas en tiempo constante igual a 1; sin embargo, por la composición de campos se darán posibilidades analíticas, por ejemplo agrupar por modalidad obviando el año o la categoría, etc.

6.3.5 Soporte al proceso documental de las agrupaciones

A lo largo del proceso de diseño se decidió adscribir esta funcionalidad de soporte, o aporte de información, al sistema de *data warehouse* por el carácter de resumen que tienen los datos implicados. Concretamente, se contará con acumuladores, por cada agrupación y tipo de obra, que indiquen para cada tipo de documento el total de estos registrados siempre y cuando el tipo de soporte esté indicado, lo cual implica que existe documentación representativa archivada.

Implementación

Mediante la tabla DWH_SoporteDocumental se recrea un último esquema basado en acumuladores como el descrito anteriormente. De nuevo se utilizan disparadores desde diversos lugares; por un lado, al editar la tabla AgrupacionAnio, de editarse los campos de interpretación y uso de obras habrá que hacer recuento de los documentos vinculados a las mismas.

La segunda vía de inserción se implementa a través de disparadores en las tablas intermedias que forman las relaciones de tipo N:M entre los documentos y las obras y sus componentes, DocumentoObra, DocumentoLetra, DocumentoMusica y DocumentoDisenio. Estos *triggers* llamarán a un procedimiento almacenado que actuará si la obra o el componente al que se asocia el documento tiene relación de interpretación, en caso contrario sería el disparador de la tabla AgrupacionAnio el que lograría la consistencia final de estas estadísticas.

Finalmente, desde la tabla Documento se activará un último disparador dedicado a gestionar esta tabla de hechos. En esta ocasión interesa el tipo de soporte del documento, además permitirá detectar la eliminación total o parcial de las relaciones del documento.

SoporteDocumentalAgrupaciones

<agrupacionAnio, tipoObra, tipoDocumento, tipoSoporte, total>

Al igual que la primera vista, AgrupacionesGanadorasConcurso, se trata de mostrar una lista ordenada, según el orden de los atributos indicados.

7. Pruebas

El desarrollo de esta fase se describe de forma detallada, por función y atendiendo a los dos entes diseñados, BBDD y DWH, en el Anexo V.

7.1 Enfoque y metodología

La metodología en este caso vuelve a ser iterativa e incremental, cada procedimiento lleva su propio conjunto de pruebas y se le han aplicado atendiendo a estos enfoques; asimismo, el mismo proceso de pruebas de cada uno ha llevado en ciertos casos a reiterar sobre el conjunto del que depende o al que afecta.

Según esta pauta, por cada procedimiento se centran las pruebas en el uso de la tripleta dedicada a alta, modificación y baja por cada tabla con objeto de forzar todas y cada una de las condiciones que estos tienen implementadas. De esta forma se avanza respetando las interdependencias entre tablas y los propios procedimientos almacenados. Cabe señalar en este punto que las funciones de uso interno fueron probadas en su momento de desarrollo al igual que las demás; esto es porque, como se exponía en el apartado dedicado a los procedimientos almacenados, toda la implementación se hace de forma gradual según las interdependencias entre tablas.

Durante todo el proceso de implementación y pruebas, como cabía esperar, se detectaron deficiencias en los modelos planteados; aunque, en la presente memoria, y en todos los anexos, todo está descrito de forma consistente al producto final. A continuación se listan los defectos detectados, o dicho de otra forma, las restricciones que no se hicieran explícitas.

Nivel técnico

- Se verifica que, por el hecho de que las funciones de PostgreSQL sean tratadas como transacciones, este SGBD no gestiona el estado de los tipos *secuence* al igual que el de los registros durante el *rollback.* Se decide en este punto gestionar el decremento en caso de *rollback* de las secuencias inicializadas a 100000, por preservar el conjunto disponible de identificadores.
- Se hace necesaria la inclusión de identificadores en las tablas que modelan relaciones de cardinalidad N:M; por ejemplo en las adjetivas y, particularmente, en la tabla crea porque admite que se le cambie cualquiera de los dos valores que son referencia a Entidad y Disenio.
- En cuanto a las relaciones denominadas adjetivas, también a lo largo del desarrollo se observa que no tiene sentido cambiar el descrito sino el descriptor; cosa que, aun así, no evita la necesidad del identificador.
- Desde este momento, y paralelamente, se empiezan a generar los diversos procedimientos almacenados dedicados a la tabla Documento. En principio se hizo una implementación básica para la tabla en si, más para las tablas de relación N:M en que participa; no obstante, a lo largo del desarrollo y las pruebas, estas funciones fueron creciendo para reducir el código redundante y centralizar validaciones en lo posible.

  Como consecuencia directa, por ejemplo, se concluye que estos procedimientos almacenados deben recibir los identificadores y el

Código de Referencia ya gestionados. Por otro lado, en relación a este último valor, se hacen ciertos ajustes en cuanto a su verificación, principalmente atendiendo a la posición de unas y otras signaturas de las que contiene.

- Las entidades referentes a agrupaciones, eventos y personas no llevaron a más conflictos que los propios de la codificación de condicionales y demás. Únicamente cabría aludir a la gestión de roles, concretamente durante las pruebas se detectó que en una misma agrupación se podían repetir ciertos roles si la persona no era la misma; por tanto, dentro del procedimiento de alta y modificación de Cargo había que consultar que ciertos roles no se hubieran registrado ya –aun teniendo la restricción de tipo *unique* sobre el compuesto entre persona, agrupación y tipo de cargo–.

  A raíz de esto se decidió controlar un hecho similar acerca de los géneros que admiten las obras; en concreto, que no debe haber más de una presentación y un popurrí por obra literaria, lo cual se controla en las funciones, y que no deben repetirse los géneros por obra musical, que se incluyó al DDL por ser representable mediante una restricción de unicidad en el compuesto entre signatura de la obra y género.

- De cara a la gestión de Obra, de forma similar al caso de Documento, se van centralizando las funcionalidades gradualmente en función de los resultados observados. Por ejemplo, para el tipo de documentos que pueden asociarse según el tipo de obra, la verificación se hace del lado de Obra, más es centralizable en una misma función –concretamente, en la que gestiona la tabla de relación entre Documento y Obra–.

  La deficiencia del modelado, en este caso, refiere a la forma de implementar la posibilidad de modificar la signatura de nivel de obra; para ello había que considerar el hecho de que, dicho valor, está contenido en todos los Códigos de Referencia de los documentos que describen los componentes de la obra en cuestión; en consecuencia, requiere de recorrer toda la tabla de documentos para dejar dichos códigos consistentes a la modificación. Naturalmente, verificar este comportamiento implica consultar el listado de documentos asociados a la obra, luego modificar la signatura de la misma y, finalmente, comprobar que se obtiene la misma lista con la actualización.

- Finalmente, en cuanto a los procedimientos propios de la BBDD, para las verificaciones acerca del tipo de documento que se puede asociar a cada componente de obra se alcanzó cierta reducción en la forma de implementar los condicionales

- Ahora bien, para el *data warehouse* se detectó que las comparaciones entre valores y nulos nunca funcionaban, aun siendo el caso válido. Esto se solventó mediante el uso de la función *coalesce* para convertir a numérico o a cadena, según el tipo a comparar.

  Por otro lado, en ciertos casos, hubo que cambiar el momento de la gestión de los eventos CRUD; por regla general se determinó hacerlos siempre en *after* excepto ante operaciones de eliminación, ya que en ciertos casos era necesario hacerlo antes de la desaparición de determinados registros, en caso contrario se hacía imposible localizar la

información de interés. Forzosamente, y como ejemplo, esto es necesario en el caso de eliminar letras y diseños, para reducir los contadores en temas recurrentes; esto es debido a que la eliminación de la relación entre temas y estos conceptos se hace en cascada por el carácter adjetivo de los temas.

Por último, dentro del ámbito técnico del *data warehouse*, se introducen los identificadores en algunos casos de conceptos ya incluidos; principalmente, para que las búsquedas fueran sobre numéricos.

Nivel semántico

- El primer cambio, en este aspecto, fue reubicar la restricción de *check* cadenas no vacías en la declaración de las tablas, en lugar de verificarlo en todos los casos de atributos de carácter con *not null.*
- En los modelos se desarrollan las condiciones sobre las claves candidatas y alternativas en Persona, pero no eran lo suficientemente concretas en cuanto al DNI; similar fue el caso del Código de Referencia de ISAD(G), al que se le refinaron las condiciones respecto al patrón inicial y el uso de barras y guiones.
- Luego, uno de los defectos más notables fue la falta de planificación del cómo debían gestionarse los documentos, al bajo nivel en que se trabajaba en ese punto. Si se había determinado que un documento no podría cambiar directamente de un objeto a otro, de entre los que es posible vincularlos; pero, se debe hacer notar que no es necesario eliminarlo porque esté asociado a un objeto erróneo. En definitiva, es suficiente con darle alta en el correcto y luego baja en el inicial.

  Como consecuencia de esto último, se llega al obstáculo en que se quiere dejar a un objeto descrito a través de un registro en Documento sin el mismo. Esto tiene una solución obvia tras lo argumentado, no obstante será el uso del SI final el que lleve al aprendizaje de estas peculiaridades, o artimañas, a seguir en pos de no perder un registro completo de Documento ya descrito.
- Por último, en cuanto al *data warehouse*, se da una circunstancia común a todas las funcionalidades propuestas –excepto en el caso de los premios por agrupación–: el hecho de no existir una agrupación relacionada, de forma directa o indirecta, con el hecho a registrar determina si se trata el hecho o no. Esto implica que, en caso de retirar una agrupación de uno de los hechos registrados mediante cualquiera de las tablas de *data warehouse*, o en caso de incluirla en una combinación que generaría uno de los datos objetivo, había que recuperar todos aquellos que no habían sido tratados, o que ya lo habían sido, en ese mismo instante.

  Otra posibilidad surgida durante las pruebas fue aprovechar la edición de la tabla DWH_PremioConcurso para el precálculo de los premios por persona, en lugar de realizarlo también a través del *trigger* asociado a Participa –que es el que alimenta a dicha tabla–. La solución es crear un disparador en DWH_PremioConcurso para gestionar los premios por persona, puesto que esta ya ha realizado los filtros pertinentes; por añadidura, se libera el lado operacional de la BBDD.

7.2 Rendimiento

Según las premisas y necesidades futuras del sistema, se hacían necesarias algunas mediciones del tiempo de respuesta ante consultas sobre campos de texto de longitud indeterminada. Por no ser de especial relevancia de cara a la defensa del proyecto se expone brevemente en que consistieron estas pruebas.

Considerando el objetivo, la preparación de la pruebas solo requería llenar la tabla Letra o Documento con alrededor de 20000 registros con el campo de texto libre cumplimentado con un alto número de caracteres. Esto consistía simplemente en crear un *script* con unas 100 llamadas al procedimiento de alta en la tabla Letras –por ser un contenido real disponible– y, posteriormente, ejecutarlo las veces suficientes. Luego se realizaron búsquedas simultáneas desde tres terminales físicamente independientes, se obtuvieron resultados de entre 20 y 25 segundos; mientras que, haciendo las consultas desde la misma maquina que ejecutaba el servicio PostgreSQL, no se alcanzaban los 10 segundos. En principio se considera bastante aceptable.

7.3 Baterías de pruebas y automatización

Directamente basadas en las premisas anteriores, acerca de la interdependencia entre contenidos, tablas y procedimientos almacenados. Se crea un conjunto de *scripts,* prefijados con el patrón '**10_pruebas**', que contienen un conjunto de llamadas a las diversas funciones. Se adjuntan también como parte del producto otro grupo de *scripts*, de prefijo '**09_inserts_**', que hacen un llenado inicial de la BBDD con objeto de permitir el uso libre de la misma; de hecho, se deben ejecutar estos últimos para que funciones los primeros, porque aquellos hacen uso de valores en las tablas de *lookup.*

Los *scripts* de pruebas hacen en cierto grado las mismas que las realizadas de forma manual, es decir, buscan que las condiciones generen error. Para ello, por cada tripleta de procedimientos, se realizan inserciones y luego consultas para localizar los identificadores necesarios de cara a la modificación y las bajas. En todos los casos, por cada procedimiento a probar, se inserta un registro a tal efecto; seguidamente, se obtiene su identificador, en los casos en que no puede indicarse, y se realizan sobre el todas las pruebas pertinentes; por último, se eliminan todos. En otros casos, cuando las inserciones dependen de la existencia de otros registros, se hacen consultas en las tablas de las que depende, tomando uno aleatoriamente.

Finalmente, con objeto de verificar el buen funcionamiento del *data warehouse*, se eliminan datos en sentido contrario al descrito; por esta razón, se entrega un último *script* de pruebas centrado en este aspecto. Cabe decir, que dentro de las automatizadas en los *scripts* anteriores no se tiene estado, puesto que todos son retirados; sin embargo, todas ellas componen este último sin las que gestionan el borrado de los registros añadidos, y sin los mensajes informativos.

8. Conclusiones

Considerando el orden temporal en que se desarrolla el proyecto, hay que empezar por resaltar la experiencia adquirida durante el contacto directo con un diverso grupo de usuarios, con diferentes niveles de interés, de conocimientos informáticos, etc. En este caso, tanto a nivel comunicativo como en el nivel de la gestión de requisitos y expectativas.

Seguidamente, se hace notar la necesidad de revisar repetidas veces los conceptos, a medida que se modela el dominio; concretamente, así como crece el nivel de detalle o, dicho de otra forma, como se baja en los niveles de abstracción. Probablemente estos dos hechos sean las experiencias más relevantes del conjunto del proyecto.

A continuación, y en consecuencia, la implementación de un código que contenga todas y cada una de las restricciones de usuario. Para esto se obvia totalmente el modelado lógico –el cual resulta prácticamente mecánico a partir del conceptual–, ya que un 40% del tiempo se ha dedicado a desarrollo y pruebas. La conclusión a este respecto es que ningún modelo es lo suficientemente mínimo o reducido; aunque a este, le falta mucho por crecer.

Finalmente, en cuanto el *data warehouse,* se partía de unos conocimientos y experiencia totalmente nulos. De nuevo en este caso se hizo una propuesta mínima que en la práctica también se tornaba desmesurada.

Respecto a los objetivos y su consecución, conseguidos están. La planificación siguió el orden pero requirió duplicar las horas, sobre todo en la fase de desarrollo y pruebas, más en toda la recta final. Otro inciso en este ámbito es que los imprevistos son lo que son; es decir, imprevisibles.

En cuanto a la metodología, el tono iterativo e incremental fue más notable durante la primera etapa, la centrada en el modelado conceptual; posteriormente, se vuelve a tomar este tono durante el diseño físico y la implementación. Aun así, se observa en general el desarrollo en cascada, pero en reiteradas ocasiones surgían decisiones de alto nivel que llevaban a una revisión completa del alcance, los modelos y la documentación; por otro lado, si se continuó con la paralelización de tareas complementarias. Resumiendo este aspecto, se considera que la metodología se ha seguido y que era la adecuada para el nivel de conocimientos de partida acerca del dominio y la tecnología.

A nivel técnico, ha resultado de especial interés la práctica relativa a las decisiones arquitectónicas del modelo, sobretodo el cómo se ha hecho presente la futura interfaz en ciertas decisiones. La forma en que afectará el producto a dicha interfaz o dónde deben estar las restricciones y validaciones, siempre con el objetivo de obtener un diseño modular consistente y altamente flexible e intercambiable. En definitiva, se entiende que proveer de una API en forma de conjunto de procedimientos almacenados provoca la seguridad e integridad del modelo según este fue definido, a la vez que lo convierte en un todo independiente.

Como conclusión general y a nivel personal, tras lo ya vivido durante mi relación con la UOC –y como cabía esperar de una y otra–, este proyecto vuelve a poner a prueba la parte de mis capacidades personales que me han llevado a finalizarlo, a él y al Grado.

9. Glosario

accésit: Primeros premios en concursos en la fase anterior a la última, o final.

AEPD: Agencia Estatal de Protección de Datos, entidad encargada de la gestión y el cumplimiento de la LOPD.

API: La *Application Programming Interface* es el conjunto de funciones que permiten la comunicación de *software* con su entorno a bajo nivel.

atomicidad: Característica de los SGBD por la que garantizan que las instrucciones son completamente ejecutadas o, en caso contrario, realiza las operaciones que restauran el estado inicial.

atómico, atributo: Característica de los atributos no divisibles semánticamente.

atributo discriminante: En relaciones de herencia puede utilizarse, en la clase de nivel más general, para diferenciar la pertenencia de las instancias a una u otra subclase.

Aula de Cultura del Carnaval de Cádiz: Institución vinculada al Ayuntamiento y al Patronato del Carnaval de Cádiz con fines burocráticos, legislativos y documentales

callejera: Tipo de agrupación carnavalesca que no asiste a concursos oficiales.

cardinalidad: (véase "multiplicidad")

catalogación, proceso: Conjunto de actividades descriptivas realizadas sobre los objetos de inventario en Centros de Documentación.

CDU: Clasificación Decimal Universal. Sistema para la signaturización jerarquizada por materias en gestión bibliotecaria.

Centro de Documentación: Ente dedicado a la proteger y conservar fondos documentales.

***check*, restricción:** Restricción a nivel de columna que permite verificar

chirigota: Modalidad de agrupación carnavalesca compuesto de entre 7 y 12 personas, con carácter humorístico.

clase abstracta: En relaciones de herencia o generalización, es la característica opcional de las clases en posición de general respecto de otras más especializadas, por la que se indica que no podrá ser instanciada directamente.

clave alternativa: En el diseño lógico representa la alternativa a las claves candidatas, con igual nivel de significación.

clave candidata: Atributo o conjunto de estos, en el nivel conceptual del diseño de BBDD, que permite identificar las tuplas de modo único.

COAC: Concurso Oficial de Agrupaciones Carnavalescas

***coalesce*:** Función de PostgreSQL para gestionar valores, en caso de recibir nulos establece el que reciba como último parámetro.

comparsa: Modalidad de agrupación carnavalesca compuesto de 15 personas, con tono crítico y pasional donde se da importancia al conjunto de voces.

componente (véase "composición, relación de")

composición, relación de: Relación entre entidades o clases donde una resulta estar compuesta de instancias en la otra, que actúa a modo de componente, mediante un tipo de relación de multiplicidad 1:N.

compuesto (véase "composición, relación de")

coro: Modalidad de agrupación carnavalesca compuesto de 45 personas, con tono humorístico donde se da importancia al conjunto de voces.

cuarteto: Modalidad de agrupación carnavalesca compuesto de entre 3 y 5 personas, con tono humorístico y carácter teatral.

Cuplé: Género literario y musical utilizado por comparsas y chirigotas, de tono humorístico.

***data warehouse*:** Es un almacén de datos independiente de la BBDD de la que se alimenta, dedicado a la colección de hechos y eventos de interés con fines estadísticos.

Depósito Legal: Exigencia legal de entregar un ejemplar de cualquier tipo de creación que pretenda distribuirse públicamente, como consecuencia la obra queda protegida por la LPI.

descripción descendente: Dentro del proceso de modelado conceptual de un dominio de interés refiere al modo de descripción que va de los abstracto a lo específico.

diagrama de clases: Representación gráfica de un modelo de *software* generalmente utilizado en el paradigma de la programación orientada a objetos. Consiste en un grafo donde los nodos representan a las clases y las aristas a las relaciones entre estas.

disparador: Evento asociado a las posibles operaciones de actualización de las tablas de una BBDD que inicia la ejecución de una función o procedimiento almacenado.

Fiesta de Interés Turístico Internacional: Distinción honorífica concedida por el Ministerio de Turismo.

FNBC: Forma normal de Boyce-Codd que solventa defectos de la 3FN.

formas normales (normalización): Conjunto de reglas aplicables a un diseño lógico de BBDD para reducir anomalías como la redundancia.

Gantt, diagrama: Gráfico utilizado en gestión de proyectos para organizar tareas conforme a una línea de tiempo.

generalización, relación de: Técnica del modelado conceptual por el que se define una jerarquía de clases, ubicando en la que actúa como nodo raíz las características compartidas.

generalización completa: Condición en las relaciones de herencia por la que no se permite hacer uso de clase padre; por tanto, esta última tiene la condición de abstracta.

generalización disyunta: Condición en las relaciones de herencia por la que no se permite que un subtipo herede de más de una clase padre.

herencia, relación de (véase “generalización”)

IDE: *Integrated Development Environment* son herramientas utilizadas en el desarrollo de *software* con diversas utilidades que facilitan este proceso, como marcadores de sintaxis y depuradores.

integridad referencial: Mecanismo de los SGBD por el que se controla el correcto uso de las relaciones declaradas entre tablas, atendiendo a las posibles configuraciones que esta admite.

ISAD(G): *General International Standard Archival Description* es la norma archivística que regulariza la forma en que se describen las unidades documentales.

LOPD, Ley Orgánica de Protección de Datos: Ley dedicada a la protección de datos personales cuando estos se entregan de forma reglada por la misma.

LPI, Ley de Propiedad Intelectual: Ley dedicada a la protección de los derechos de autor.

Multiplicidad: Refiere al rango en cuanto al número de instancias o elementos que puden participar de cada lado de una relación entre clases o entidades.

multivaluado: Característica o atributo de una entidad que admite más de un valor.

Pascal y Camel: Tipos de grafías utilizados para denominar variables y otros elementos en programación.

pasodoble: Género literario y musical utilizado por comparsas y chirigotas, de tono crítico y generalmente político o social.

Patronato: Es el órgano colegiado responsable de hacer cumplir el reglamento del concurso.

PMBOK: El *Project Management Body of Knowledge* es un marco de referencia en la gestión de proyectos de reconocimiento internacional.

Popurrí: Género literario y musical utilizado como último elemento del repertorio en las actuaciones, es una composición bastante larga que contiene una serie piezas diferentes.

Presentación: Género literario y musical utilizado como primer elemento del repertorio en las actuaciones, normalmente introduce la idea y el tipo que defiende la agrupación.

procedimiento almacenado: Conjunto de instrucciones SQL

relación 1:1, 1:N y N:M (véase “multiplicidad”)

requisito funcional: Necesidades que el sistema a modelar debe satisfacer, refieren a la semántica del dominio o contexto y al cómo desarrollas las operaciones.

requisito no funcional: Restricciones técnicas identificadas respecto a un sistema a modelar.

***rollback*:** Mecanismo de los SGBD por el que se deshacen las operaciones sobre datos que no han podido finalizarse con carácter atómico.

Romancero: Modalidad de agrupación carnavalesca compuesto de entre 1 y 4 personas, con tono humorístico y normalmente callejera.

***script*:** Fichero que contiene una secuencia de instrucciones interpretadas sin compilación.

Scrum: Metodología ágil de carácter iterativo e incremental que da una alta importancia a los procesos comunicativos periódicos.

SGDB: Sistema gestor de bases de datos, normalmente abarca al servicio y los procesos.

SI: Sistema de Información, conjunto organizado e integrado de aplicaciones *software.*

signatura: Identificador único, numérico o alfanumérico, utilizado en archivśitica.

subclase: En relaciones de herencia, esta especializa a la clase declarada como general; por tanto, hereda los atributos que compartidos con otros y añade los propios.

subtipo (véase “subclase”)

superclase: En relaciones de herencia contiene las características comunes a las subclases que heredan de ella.

tango: Género literario y musical de carácter crítico utilizado por los coros.

tipo: Es el conjunto de diseños utilizado por las agrupaciones para caracterizarse.

trampas de diseño: Son errores en el modelado de BBDD que responden a patrones conocidos; son ejemplos el abanico, el corte, la pérdida de afiliación y la aridad.

transacción: Mecanismo de los SGBD que permite definir y ejecutar un conjunto de instrucciones como si fuera una única instrucción, con carácter atómico.

trigger (véase disparador)

UML: El *Unified Modeling Language* es un estándar para el modelado de sistemas de *software*, le cual establece reglas para definir distintos tipos de diagramas.

univaluado: Característica o atributo de una entidad que admite un solo valor.

10. Bibliografía

[1] **Pradel, J.; Raya, J.** (2011) *"Ingeniería del software".* Editorial FUOC – Eureca Media, SL. ISBN: 978-84-693-9187-7. Depósito legal: B-4.230-2011

[2] **Rodríguez J. R.; Mariné, P.** (2010) "*Gestión de proyectos*". Editorial FUOC – Eureca Media, SL. ISBN: 978-84-693-4292-3. Depósito legal: B-29.661-2010

[3] **Burgués, X.; Cabré i Segarra, B.; Casas, J.; Costal, D.; Juanola, P.; Ortego, S.; Pérez, J.; Rius, À.; Segret, R.** (2015) "*Diseño de bases de datos".* (4ª ed.) FUOC - Oberta UOC Publishing, SL. Depósito legal: B-9.858-2015

[4] **Consejo Internacional de Archivos** (1999) *"General International Standard Archival Description".* [Norma en línea] <www.mecd.gob.es/cultura-mecd/dms/mecd/cultura-mecd/areas-cultura/archivos/recursos-profesionales/normas-archivisticas/isad.pdf> [fecha de consulta: 20 de Marzo de 2017]

[5] **Curto, J.** (2007) *"Diseño de un data warehouse: definiciones".* [Artículo en línea]. <josepcurto.com/2007/10/16/diseno-de-un-data-warehouse-definiciones> [fecha de consulta: 15 de Abril de 2017]

[6] **Curto, J.** (2008) *"Diseño de un data warehouse: tabla de hecho".* [Artículo en línea]. <josepcurto.com/2008/07/12/diseno-de-un-data-warehouse-tabla-de-hecho> [fecha de consulta: 15 de Abril de 2017]

[7] **Escarcega, B.** (2008) *"Mejores Prácticas para construir un Data Warehouse".* [Artículo en línea]. <gravitar.biz/datawarehouse/mejores-practicas-para-construir-un-data-warehouse> [fecha de consulta: 15 de Abril de 2017]

[8] **Bernabeu, D.** (2009) *"3.4 Datawarehouse manager".* [Artículo en línea]. <www.dataprix.com/datawarehouse-manager> [fecha de consulta: 16 de Abril de 2017]

Recursos web utilizados

www.wikipedia.org

www.postgresql.org

www.postgresql.org.es

11. Anexos

- **Anexo I: Diseño conceptual**

  Contiene todo el detalle y las implicaciones semánticas que han llevado a generar el modelo de datos resultante. Contiene un apartado por cada actividad descriptiva asociada a este nivel de modelado: entidades, claves candidatas y restricciones semánticas, o del dominio.

- **Anexo II: Diseño lógico**

  Análogo al anterior, pero centrado en el proceso del diseño lógico del esquema de datos. El documento se divide en dos partes, correspondientes a las dos mitades en que se puede dividir el gráfico del modelo anterior. Cada parte comienza por la porción de relaciones que caen dentro de la zona, o mitad, que le corresponde; seguidamente, se comenta al detalle cada relación: las decisiones de diseño, el nivel de satisfacción de las formas normales y las restricciones lógicas y semánticas que habrá que implementar en la siguiente fase.

- **Anexo III: Diseño físico**

  En este caso se inicia el documento con un apartado dedicado a la preparación del entorno de desarrollo, luego se hace un inciso acerca de las decisiones tomadas para la estructura del nivel virtual. Tras estos breves detalles se procede con la descripción de los procedimientos almacenados, incluyendo decisiones de nivel técnico y de uso de los mismos.

- **Anexo IV: Data Warehouse**

  Este documento se centra en las funcionalidades *data warehouse* implementadas. El documento se organiza por funcionalidad, luego para cada una se desarrollan los diversos niveles de modelado; es decir, conceptual, lógico y físico, finalizando con las vistas de cada una de estas funcionalidades.

- **Anexo V: Pruebas**

  El último de los anexos generados describe el proceso de pruebas, estructurado en función a los *scripts* de la BBDD; paralelamente, se va haciendo seguimiento del estado de las tablas de hechos del *data warehouse*, así como las pruebas de la BBDD van sucediéndose.