

Propuestas para Enseñar el Diseño de Estructuras de Datos en Bases de Datos Relacionales

Luis Esteban Damiano¹, Roberto Miguel Muñoz¹, Calixto Maldonado¹, Soledad Romero¹, Andrea Guevara¹, Matías Bueno¹, Sergio Quinteros¹, Juan Peretti¹

¹ Departamento Ingeniería en Sistemas de Información, Universidad Tecnológica Nacional, Facultad Regional Córdoba, Maestro Marcelo López esquina Cruz Roja, Ciudad Universitaria, Córdoba, Argentina.
rmunoz@frc.utn.edu.ar
{luis.damiano, calixtomaldonado, romeroma.soledad, andrezza77, matiasbueno, ser.quinteros, peretti.juan}@gmail.com

Resumen. En la cátedra de Gestión de Datos de Ingeniería en Sistemas de Información, de la Universidad Tecnológica Nacional, uno de los contenidos mínimos por Resolución Ministerial es Normalización. Por ello un grupo de docentes de la cátedra busca la manera de obtener una metodología de trabajo, en el ámbito educativo, que favorezca el análisis de datos y acompañe al proceso de aplicar las reglas, llamadas formas normales, para obtener una estructura de datos deseable en la base de datos de una organización. Reforzando la aplicación de las reglas y considerando las propuestas de este trabajo se evitará la obtención de las estructuras basadas sólo en la experiencia personal y en un proceso puramente deductivo. A partir de las experiencias de docentes en la cátedra y el trabajo en el proyecto de investigación, que se generó a tal fin, el grupo propone definiciones, principios y recomendaciones, que marcan el desarrollo de una técnica de normalización, la cual pueda ser utilizada como recurso didáctico dentro del ámbito académico.

1 Introducción

En el marco curricular de la carrera de Ingeniería en Sistemas de Información, que se dicta en la Universidad Tecnológica Nacional (UTN), la asignatura Gestión de Datos del tercer año tiene como eje central a las Bases de Datos. En lo que respecta a Bases de Datos Relacionales se desarrollan los conceptos del modelo, integridad, seguridad y se incluye a la normalización como técnica de diseño de una base de datos.

En la Facultad Regional Córdoba (FRC), un conjunto de docentes de esta asignatura trabaja en la temática Base de Datos desde proyectos de investigación en la UTN. Estos docentes no sólo se dedican a las actividades de docencia e investigación, sino que reúnen importante experiencia en el desarrollo profesional, tanto en empresas privadas como públicas. El proyecto de investigación (PID), que da sustento a este artículo, no es el primero que el grupo lleva adelante y por ende ha reunido experien-

cia docente y profesional en los conceptos propios al modelado de datos y normalización de modelo de datos para base de datos relacionales.

El actual proyecto de investigación se denomina: “Buenas prácticas en el diseño de estructuras de datos en bases de datos relacionales”. Este proyecto está orientado al desarrollo de buenas prácticas, como lo menciona el nombre, no tan sólo en el diseño de estructuras de datos, sino también en el desarrollo de técnicas y métodos explicativos que mejoren la didáctica en la enseñanza de los conceptos teóricos y prácticos para el diseño de estructuras de datos relacionales. Este PID fue homologado por la Secretaría de Ciencia, Tecnología y Posgrado de la UTN, con el código SIUTN-CO0004979 y ejecución en el período del 1° de enero de 2018 al 31 de diciembre de 2019. Se encuentra dentro de la Unidad Científico-Tecnológica Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (CIDS) de la UTN-FRC.

Los antecedentes de proyectos de investigación en la misma línea de conocimientos en los que han participado algunos de los docentes del proyecto referido, son los siguientes:

1. TecnoDB - Administrador de Base de Datos Relacional (2007) [1]
2. PROMETEO - Desarrollo de un método y una herramienta para el aprovechamiento de Metadatos de Base de Datos Relacionales (2010) [2]
3. Análisis y aplicación de metodologías para la generación de consultas complejas utilizando esquemas OLAP (2010) [3]
4. Generador Automático de Modelos de Datos Normalizados en Bases de Datos Relacionales (2014-2015) [4]
5. MultiDB. Plataforma Web para acceder a diferentes Sistemas de Gestión de Bases de Datos Relacionales (2016-2017) [5]

1.1 Origen de la Investigación

El proyecto surge de haber notado discrepancias en la forma de entender los aspectos prácticos que se deben considerar para la normalización de estructuras de datos. Las disparidades no están dadas en el texto ortodoxo de las definiciones teóricas, existentes en la bibliografía clásica de autores reconocidos, sino en la observación de los Arquitectos de Datos al momento de fundamentar la acción y la práctica de normalizar un modelo de datos.

Los autores de libros de base de datos, de renombre global y local, han publicado una gran variedad de libros que dan fundamentos teóricos sobre el tema que se plantea en el proyecto. En dicha bibliografía se describe el modelo relacional en forma minuciosa y, en términos generales, todos los autores describen las misma temáticas sin significativas diferencias, pero no se encontró, por ejemplo, respuesta en ninguno de ellos al siguiente interrogante: ante un requerimiento de diseño para base de datos relacionales, proveniente de distintos Arquitectos de Datos, ¿Puede ser única la estructura de datos normalizada solución? Tampoco encontramos una propuesta de técnica o de método para desarrollo de modelos de datos para alcanzar una forma normal determinada.

Para definir sintéticamente los conceptos principales del tema del que estamos hablando se debe entender: 1- bases de datos relacionales son las estructuras que brindan persistencia a los datos en el modelo relacional, muy extendido en el mercado y ampliamente estudiado en la academia; 2- el modelo relacional es un modelo matemático, basado en teoría de conjunto, y lógica de predicados, creado por el Dr. Edgar Frank Codd en el año 1970; 3- el sustento más significativo para normalizar una estructura de datos son los conceptos de dependencia funcional, dependencia funcional completa, dependencia funcional multivaluada, dependencia funcional de reunión, junto a la definición de las formas normales; 4- las formas normales a su vez tienen sustento en las dependencias funcionales y la definición de las 15 reglas de Codd para el modelo relacional; 5- las reglas de Codd son el fundamento de la definición del modelo relacional; 6- las formas normales son cinco (5), mencionada la primera como primera forma normal con sigla 1FN, y así hasta la quinta forma normal con sigla 5FN, más una especialización de la tercera forma normal (3FN), denominada forma normal Boyce- Codd, con sigla FNBC; 7- en general se entiende en el ambiente académico y profesional de la informática que una estructura de datos bien formada se encuentra en tercera forma normal (3FN); 8- la normalización es la acción que nos permite desarrollar el modelo de datos conceptual para una base de datos relacional; 9- el Arquitecto de Datos es quien ha logrado un nivel de experticia suficiente para producir una normalización; y 10- el modelo conceptual para estar normalizado a 3FN tiene que tener todas sus estructuras en esa forma normal.

Respecto de la bibliografía accedida, podemos destacar los siguientes tres autores que plantean a la normalización, como técnica para crear estructuras de bases de datos relacionales en estado óptimo y deseable (normalizado), evitando redundancia de datos y futuras inconsistencias. C.J. Date en su libro “Introducción a los Sistemas de Base de Datos” [6], que ha enfocado su atención en aspectos técnicos que marcan una tendencia clara en el conocimiento y opinión de los Arquitectos de Datos, como «...En otras palabras, normalizada y 1FN significan exactamente lo mismo; aunque debe tener presente que el término normalizado se usa a menudo para indicar uno de los niveles más altos de normalización (por lo regular a la tercera forma normal, 3FN)...» (página 349), o como expresa Ramez Elmasri y Shamkant B. Navathe en su libro “Fundamentos de sistemas de bases de datos” [7] «...Así, el diseño de base de datos, tal y como se realiza en la actualidad en la industria, presta especial atención a la normalización hasta la 3FN, la BCNF o la 4FN...» (página 300). Los autores expresan aproximadamente los límites prácticos del concepto “Normalización”, desde lo observado en el campo de la Industria del Software. No hay una estricta coincidencia en el alcance, pero sí expresan una conceptualización con algún grado de aproximación, en función del criterio de cada uno, en ambos casos los autores coinciden que se logra la Normalización a partir de la 3FN, como mínimo. A su vez Silberschatz trata el concepto Normalización con la misma visión de Date y Elmasri, en el capítulo 7 de su libro “Fundamentos de Bases de Datos” [8], pero también se permite expresar un singular punto de vista sobre la “Normalización” que expresa: «... Cuando se define con cuidado un diagrama E-R, identificando correctamente todas las entidades, las tablas generadas a partir del diagrama E-R no necesitan más normalización. No obstante, puede haber dependencias funcionales entre los atributos de una entidad...» (página 183), y en la misma página más adelante, también expresa «... De manera alternativa,

la normalización puede dejarse a la intuición del diseñador durante el modelado E-R, y puede hacerse formalmente sobre las relaciones generadas a partir del modelo E-R...». En el trabajo de este autor, Silberschatz, no se percibe que acentúe su mirada pragmática desde la observación de la industria informática sobre el tema normalización, como lo hicieron los otros dos autores y que les permitió expresar que la Normalización de una estructura de datos requiere que ésta alcance al menos la 3FN (tercera forma normal). Pero cuando se permite expresar su singular punto de vista, respecto de la Normalización, entiende que los aspectos intuitivos de un Arquitecto de Datos pueden llevarlo al diseño Normalizado de una estructura de datos. Este tipo de conceptos vertidos por los autores Date y Elmasri ha sido asimilado como una pieza de conocimiento que no tiene discrepancia, por la gran mayoría de los especialistas que diseñan estructuras de datos, no suscitando controversia. No se pudo encontrar aún, entre los autores estudiados, que algunos de ellos haya orientado su análisis y explicación de la Normalización, sobre seguir una metodología que permita realizar esta acción, demostrando que la estructura resultante es única, partiendo de una relación universal (entidad primera que reúne todos los atributos observados) como lo expresa Silberschatz, en el “El enfoque de la relación universal” en el libro ya citado «... El segundo enfoque del diseño de bases de datos es comenzar con un solo esquema de relación que contenga todos los atributos de interés y descomponerlo...» (página 183).

El equipo de trabajo postula, en función de su experiencia de campo, profesional y en la docencia, demostrar que el modelo conceptual para un problema de almacenamiento de datos es único, en el marco de base de datos relacionales y cumpliendo con las reglas de la normalización. Dicho de otra manera, la solución realizada por un Arquitecto de Datos, que se expresa en un modelo conceptual para una base de datos relacional, debiera ser única, en función a un problema correctamente documentado, claramente delineado y limitado, esto, ya fue expuesto en el libro “Base de Datos” [9]. Este es el disparador que da inicio al proyecto de investigación, pero en el desarrollo de semánticas para fundamentar esta hipótesis se han observado conceptos que dan fundamento a un conjunto de definiciones que sustentan una técnica para desarrollar una estructura de datos normalizada, las que se explicitan en el artículo.

Se ha realizado una intensa búsqueda bibliográfica para encontrar opiniones, discusiones, afirmaciones o negaciones del problema planteado. No se hallaron resultados que conformen un antecedente en los materiales abordados y que figuran en la referencia bibliográfica.

Hay que destacar que existe unanimidad conceptual en el material de todos los autores sobre las definiciones referidas al modelo relacional. Algunos con mayor énfasis en aspectos de principios, otros con más desarrollo en las características de las formas normales, pero en términos generales sin desacuerdos importantes.

Las búsquedas también incluyeron anales de congresos Argentinos (WICC, CoNa-IISI y JAIIO) y sitios de internet, sin hallar resultados favorables sobre la temática específica. Es fundamental destacar que en congresos relevantes de Informática, en los que integrantes del grupo han podido participar, cuentan con simposios de base de datos y un porcentaje significativo de los trabajos presentados se refieren a estudios sobre bases de datos relacionales.

Cabe aclarar que los integrantes del equipo de investigación, que como ya se dijo anteriormente, son docentes de la Cátedra de Gestión de Datos, trabajan los conceptos y aplicaciones de las bases de datos, que conforman los contenidos mínimos de la asignatura. Los conocimientos son homogéneos entre los integrantes, sobre los conceptos inherentes a base de datos, modelo relacional y normalización de una estructura de datos. Esta tarea docente ya ha sido desempeñada durante más de veinte años, trabajando con estos conocimientos, y han presentado diversos artículos en congresos de alcance nacional, sobre distintas temáticas que involucran estos saberes. Esta participación involucra la actitud atenta al análisis del tema de hipótesis, siendo un factor importante para aseverar que el problema a develar no está siendo tratado en este momento y no presenta antecedentes expuestos en el pasado.

2 Recursos para las Buenas Prácticas

2.1 Un Tema no Tratado

El trabajo en equipo durante años nos hizo observar y cuestionar los aspectos metodológicos prácticos, que hacen falta para realizar la normalización de una estructura de datos. El trabajar en una línea vertical de investigación, nos ha llevado a pensar y repensar conceptos que parecen terminados, desde el punto de vista de la definición o semántica, encontrando en este proceso, definiciones complementarias no tratadas en la bibliografía habitual de estudio, de capacitación y actualización para los temas de base de datos relacionales.

Uno de los tópicos observados y tratados es la cantidad probable de estructuras de datos solución, que representan satisfactoriamente un problema de almacenamiento de datos. Cuando un estudiante lee la bibliografía actual, no encuentra ninguna sugerencia de la cantidad de estructuras posibles de solución, que se podrían llegar a modelar para un problema determinado. Al no haber ninguna definición al respecto, se genera un entendimiento tácito, que hay una variedad posible de soluciones que satisfacen el problema tratado, por un lado, y por el otro, que la solución posible para ese problema es única. Este tema es realmente controvertido, pues al no encontrar el tratamiento del asunto, el conocimiento de cada Arquitecto de Datos queda enmarcado en el del saber hacer de cada uno, conocimiento que se funda en su práctica cotidiana. Estas prácticas habituales del Arquitecto de Datos son las que le llevan a considerar uno u otro punto de vista como válido.

2.2 Nuestra Observación del Problema

En términos generales, como expresan algunos autores, y este grupo de investigación concuerda, una estructura de datos en estado deseable debe encontrarse en tercera forma normal 3FN. Como ya se mencionó, las acciones técnicas y métodos prácticos, para lograr que una estructura quede en 3FN, no está definida explícitamente.

Cada Arquitecto de Datos ha desarrollado una forma de constituir el diseño solución, que posiblemente tenga influencia de su maestro, junto a su experiencia personal para lograr las soluciones. Por experiencia en la cátedra, si observamos a dos docentes, de distintos espacios académicos o no, encontraremos que ambos tratan los mismos temas teóricos en forma similar o igual, respecto de la normalización de datos, pero que sus abordajes prácticos pueden ir desde una visión no estrictamente igual a una totalmente disímil. Pero sí podemos aseverar que ambos docentes terminan mostrando y demostrando estructuras de datos en 3FN.

Al no observar un método y técnica que guíe el razonamiento de conversión de un conjunto de datos no normalizado a un diseño de una estructura normalizada, las formas de llegar al modelo solución pueden ser varias. Ahora, si las formas de llegar a un modelo solución pueden ser varias, ¿es correcto pensar que los modelos solución a un problema de datos determinado, acotado y correctamente delineado, también son varios? La experiencia marca un camino de conocimiento del equipo de trabajo, que permite enunciar expresiones que se basan en su propio contenido, las que se desarrollan a continuación.

2.3 Recomendaciones para las Buenas Prácticas en Normalización

Las recomendaciones para las buenas prácticas en la normalización de estructuras de datos, son un conjunto de definiciones y principios que se proponen para ser transmitidas a quien está siendo capacitado en normalización de datos, para bases de datos relacionales. Ninguna de estas recomendaciones está por delante de las definiciones propias del modelo relacional, las cuales no serán tratadas pues son una constante en toda la bibliografía de base de datos. Sí se aportarán ideas, principios y definiciones que surgen de la interacción del equipo en los distintos ámbitos en que se desempeña.

En primera instancia se definen en forma general los conceptos o términos utilizados, actores y componentes que interactúan en la construcción de una estructura de datos, luego se profundiza en conceptos específicos que hacen a la observación de un resultado o modelo solución.

La estructura de datos. Cuando se habla de normalización de datos, se pueden observar una variedad de conceptualizaciones aceptables para el desarrollo de un modelo. Nuestro interés está limitado sólo a las estructuras basadas en entidades donde cada una de ellas representa un concepto delimitado y caracterizado por una colección de atributos que establecen las propiedades que dan base conceptual a la entidad.

El Experto. La estructura de datos surge del conocimiento suministrado por un experto, que es quien da las características fundamentales para el desarrollo del mismo. El experto puede estar en la figura de una persona que es quien tiene el mayor conocimiento del negocio o dominio del problema, del que se debe desarrollar un modelo solución; o también en un documento que tiene en sí el contenido de todo el problema. El concepto del experto como interlocutor válido del problema, es esencial, pues es

quien fija los límites claros e inalterables del problema al que se debe ajustar la solución.

El Arquitecto de Datos. El modelo es desarrollado por un especialista que domina los principios que rigen a la normalización, este especialista es denominado Arquitecto de Datos.

Los principios de la Normalización. Normalizar una estructura de datos es poder establecer, dentro de la misma, las características y restricciones de forma que establecen las Formas Normales. Las Formas Normales son un conjunto de reglas restrictivas, cuyas características establecen un estado del modelo para cada una de ellas, con la condición de que cada Forma Normal de mayor grado, requiere que el modelo todo cumpla con la Forma Normal inmediatamente anterior. Un modelo bien formado, cumple con las condiciones establecidas en las formas normales. Para poder tener un correcto entendimiento de las formas normales se debe tener previamente el conocimiento de los conceptos de dependencia funcional en el modelo relacional.

El dominio del problema. Todo modelo tiene origen en un problema, que tiene una dimensión que debe saberse delimitar, es fundamental que el Arquitecto de Datos tenga esta habilidad. El dominio del problema está dentro del dominio del conocimiento general del tema a tratar y no es el dominio de todo el conocimiento posible del tema, sino una reducción bien acotada que representa con claridad el problema, y éste es el que se debe representar con el modelo solución.

El entendimiento del Problema. Es imperativo que el Arquitecto de Datos pueda mantenerse objetivo en el entendimiento del problema, sin caer en interpretaciones, ya que estas tienden a producir estructuras no del todo útiles a la solución esperada. Una buena práctica es recurrir al experto tantas veces como sea necesario, para esclarecer lo necesario, hasta lograr el entendimiento apropiado de este conocimiento. El quedarse con la primera observación del problema, realizando deducciones provenientes del conocimiento del Arquitecto de Datos, no se considera el camino correcto. Indudablemente todos los Arquitectos de Datos poseen en términos generales conocimientos previos del problema que tratan, esto no es una ventaja para el diseño. Entender el problema objetivamente, fuera de toda dudosa interpretación, conlleva a lograr un buen modelo de datos solución.

El Modelo Conceptual. El modelo conceptual es la representación lógica del modelo, es el diagrama representativo de la solución que se propone, para el problema tratado. En definitiva el modelo conceptual es el diagrama representativo de la solución del dominio del problema entendido y este diagrama del modelo se encuentra normalizado.

Las estructuras dentro del Modelo Conceptual. Es recomendable estandarizar la representación del modelo conceptual. Nuestra propuesta es la siguiente: a) Escribir el nombre de las entidades en plural, con mayúscula sostenida y subrayado o resaltado; b) Escribir el nombre de los atributos en minúscula, y si es compuesto unir las palabras con un guión bajo; c) Declarar la clave primaria con la indicación PK; d) Ubicar la indicación PK a la izquierda de los atributos que la conforman. Si la clave es compuesta por varios atributos se recomienda posicionarlos en forma contigua, luego tra-

zar una línea a la izquierda de los mismos, que cubra todo el largo de la combinación de atributos, y a la izquierda de esta línea escribir PK; e) Declarar a las claves foráneas o ajenas con FK, f) Ubicar FK a la derecha del atributo o de los atributos que la componen; h) Para las claves foráneas compuestas se propone seguir el mismo formato que para las PK compuestas, con la salvedad, como dijimos, se deben representar a la derecha de los atributos que la componen y con FK; i) Evitar el uso de claves artificiales por las limitaciones que ofrecen para los controles de datos, en definitiva se propone trabajar con los datos que naturalmente se reconocen en el problema. Se incluye un ejemplo en la Figura 1.

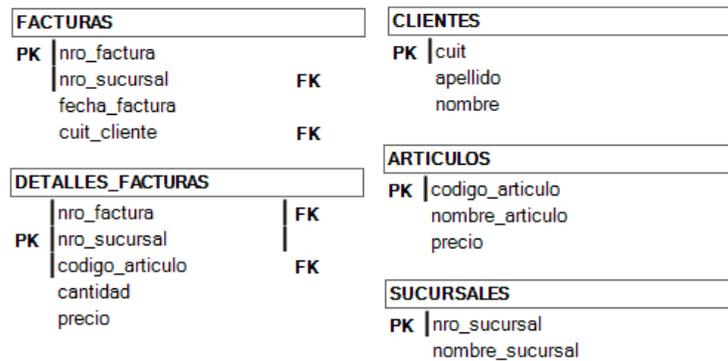


Fig. 1. Ejemplo de estructuración de una parte de un modelo de datos.

2.4 Tipificación de las Estructuras de Datos.

Para concretar el desarrollo de un modelo solución que satisface los requerimientos de un problema, se debe realizar una combinación de estructuras menores al modelo en sí y relaciones entre éstas, que en la suma terminan constituyendo el modelo final. La expresión estructura en su conceptualización de menor tamaño es equivalente a decir Entidad.

Se explicarán las estructuras tipificables, que en su variedad finita, terminan siendo las estructuras utilizables para conformar una solución válida. .

Estructura de Referencia. Las estructuras, o entidades de referencia son aquellas que cumplen la función de clasificación de un dominio establecido. Por lo general una estructura de referencia tiene un dominio claramente limitado y es mínima la cantidad de atributos que requiere para su definición. La mayoría de las estructuras de referencia se nombran con la expresión, tipo del dominio (ej. TIPOS_DOCUMENTOS), cuentan con una clave primaria, que por lo general es una secuencia numérica, y un atributo descriptor, que es en sí quien proporciona la posibilidad de mostrar los integrantes del dominio que la estructura representa, como en el ejemplo de la Figura 2.

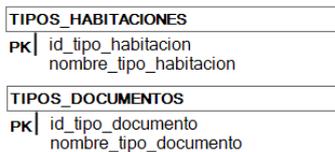


Fig.2. Estructura de referencia.

Estructuras Maestro. Las estructuras que llevan la calificación de maestro son estructuras con mayor cantidad de atributos y describen el dominio que representan. Estas estructuras son de una única entidad y constituyen una definición de mayor gravitación conceptual que una estructura de referencia. Ver Figura 3, donde se representa una estructura Maestro de clientes.

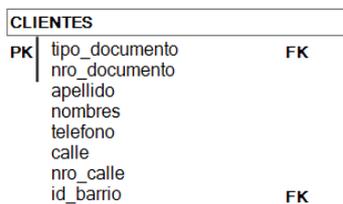


Fig. 3. Estructura maestro.

Estructura Maestro Detalle. Es una estructura que involucra a dos o más entidades que interactúan enlazadas entre sí por una clave foránea, que se ubica en la entidad detalle y que la vincula con la entidad maestro. Es una variante de la estructura maestro o ampliación de la estructura maestro. Requiere necesariamente de una entidad adicional para completar el concepto que se está modelando. Por ejemplo si estamos construyendo la estructura que permite cargar los datos de un cliente como se ve en la figura 3, pero se requiere cargar múltiples teléfonos de este cliente, cuando el número de ellos es indeterminado, se requerirá de una entidad auxiliar para completar el concepto. La estructura estará conformada por la entidad maestro de clientes y la entidad detalle, en este caso teléfonos por cliente. Ver Figura 4.



Fig. 4. Maestro detalle.

La estructura Maestro Detalle puede crecer más allá de la utilizada para mostrar el concepto, por ejemplo se puede pensar en una estructura de maestro con múltiples detalles, al concepto ya utilizado como ejemplo se le puede agregar la necesidad de

Principio de Correcta Dimensión de la Solución de Normalización. La correcta dimensión del problema es también entendible con poder fijar un límite a la solución, para que represente claramente el problema. Que no tenga estructuras faltantes y que no tenga estructuras que no estén documentadas en el texto del ejercicio o requerimiento del experto.

Principio de Elocuencia. El estudiante que se está formando para ser Arquitecto de Datos, tiene que internalizar que los modelos son una representación de una realidad acotada a la organización, y mientras más elocuente es el diseño, menos explicación adicional deberá tener. Las entidades y atributos deben ser definidos bajo este criterio, de manera que no causen duda cuando un tercero lea su diseño.

Principio de Nombre de Dominio Explícito. Una entidad está compuesta por una colección determinada de dominios, que surgen del conocimiento suministrado por el experto. En tiempo del diseño conceptual un dominio sólo debe representar con certeza el concepto que simboliza. Por ello mientras más detallado sea el nombre del dominio más evidente se volverá su contenido. Por ejemplo si estamos en la base de datos de un registro civil, en la entidad de personas pueden convivir varios dominios que tienen su fundamento en fechas, por ello si se los detalla con claridad no debería haber confusión.

Principio de Clasificaciones Contenidas. Este principio es un aditamento especial de principio de elocuencia. El modelo de datos debe contener las clasificaciones más obvias, por más que el dominio de estas clasificaciones se limite a unas pocas como el tipo de documento. El modelo de datos se expresa por sí mismo al ser observado por un tercero y no tendrá que ser explicado especialmente por este motivo.

Principio de Clave Primaria Natural. El criterio de utilización de los atributos que conforman una entidad para la declaración de la PK es primordial. La disyuntiva que se presenta entre la utilización de PK naturales o artificiales, entre los Arquitectos de Datos formados, no se debe focalizar en este momento de aprendizaje, donde el interés primero es que el estudiante internalice como construir una estructura. Luego, en su desarrollo profesional tendrá suficiente tiempo para optar por el camino natural o artificial de las PK, por ejemplo al considerar la performance u optimización. Si deben conocer que las Claves Primarias naturales ofrecen un mecanismo de control de integridad que no ofrecen las artificiales.

Principio de Clave Primaria Artificial. En algunas entidades no queda más opción que utilizar PK artificiales, que por lo general son una secuencia numérica. Estas entidades suelen ser las que entran en las estructuras de referencia, como se detalló en la clasificación de estructuras.

Principio de Igualdad de Clave Primaria Natural de dos Entidades. Si dos entidades de un modelo de datos tienen la misma PK, se debe reconsiderar minuciosamente la construcción de estas entidades, pues pueden ser la misma o estar incluyendo algún error en el diseño. Se debe analizar si los atributos de ambas entidades, no terminan estableciendo las características conceptuales de lo mismo.

Principio de Atributo de Valor Repetitivo. Todo atributo que presenta una repetición de datos en forma sistemática con una variedad de dos (2) o más valores de dominio que no sean numéricos, se debe clasificar en una tabla de referencia. Este principio está basado en evitar la presentación de atributos descriptivos que pueden ser clasificables y que dentro del modelo no pertenecen a una tabla de referencia.

Principio de igualdad estructural de dos Entidades. Lo siguiente debe tomarse como una alternativa pero no obligatoria. Al observar dos entidades que representan claramente conceptos distintos, pero que sus estructuras son exactamente iguales, se podrían unir en la misma entidad compartiendo un nombre, que represente a ambos conceptos. Por ejemplo si en un modelo de datos, los clientes y los proveedores terminan teniendo estructuras de entidad exactamente iguales, se podrá pensar en una entidad Personas que comparten la estructura.

Principio de Pertenencia Natural del Atributo. En algunos modelos un atributo en particular puede ser alojado en más de una entidad en forma alternativa. El concepto de pertenencia natural de un atributo a una entidad es poder establecer en qué entidad cumple una mejor funcionalidad dentro del modelo. Expresado de otra forma sería en qué entidad el atributo produce menos esfuerzo de recuperación y mantenimiento. Si dentro de un modelo que representa la solución a un problema de administración de un congreso científico se requiere almacenar el artículo enviado y quien lo envió, junto a la nómina de autores del artículo, una solución podría ser pensar una entidad de artículos y otra de autores de artículos, y en la entidad artículo estar el atributo que referencia a quien lo envió, pero, también quien lo envió podría almacenarse en la entidad autores por artículos, teniendo un atributo que indique esto, y no, alojarlo en la tabla artículos. ¿Cuál es el almacenamiento natural del atributo?

Principio de Dicotomía de Atributos Calculables. Hay una dicotomía conceptual respecto de la representación o no, de datos calculables dentro de un modelo solución. En esta instancia enriquece al estudiante el análisis de cada situación, porque en realidad es redundancia y complejiza el mantenimiento, pero sí debe considerarse la complejidad o el costo de recalcular el dato cada vez que desee ser accedido. La cátedra mantiene la consigna de Evitar la redundancia, pero no de eliminar la redundancia sin un análisis previo.

Referencias

1. Gastañaga, I., Maldonado, C., Martínez Spessot, C. I., & Hintermeister, E.: TecnoDB una Base de Datos Relacional y Prometeo un método de aprovechamiento de Metadatos y Generador de Consultas. In VIII WICC. (2006)
2. Marciszack, M., Maldonado, C., Martínez Spessot, C. I., Muñoz, R., Navarro, A., Peretti, J. P., Roggero, L.: Prometeo, una herramienta para el aprovechamiento de metadatos de base de datos relacionales. In XI WICC. (2009)
3. Vaca, P. A., Abrutsky, M.: Herramienta para consultas complejas orientada a usuarios finales. In XI WICC.(2009)

4. Paz Menvielle, M. A., Cuevas, J. C., Damiano, L. E., Muñoz, R., Quinteros, S.: Generador automático de modelos de datos normalizados en bases de datos relacionales. In XV WICC. (2013)
5. Muñoz, R.M., Maldonado C., Damiano L.E., Romero M.S., Cuevas J.C., Quinteros S.R., Guevara A.M., Carrasco A.: MultiDB, Plataforma Web para acceder a diferentes Sistemas de Gestión de Bases de Datos Relacionales. XVIII WICC (2016)
6. Date, Christopher: Introducción a los Sistemas de Bases de Datos, Volumen 1, Quinta Edición, Edit. Addison Wesley Iberoamericana – Estados Unidos, ISBN: 0-201-51859-7. (1993)
7. Elmasri, R. y Navathe, S.: Fundamentos de Sistemas de Bases de Datos, 5ta Edición – Edit. Pearson- Estados Unidos, ISBN: 978-84-7829-085-7. (2007)
8. Silberschatz y otros : Fundamentos de Bases de Datos , Quinta Edición, Edit. Mc Graw Hill, Estados Unidos , ISBN: 84-481-4644-1. (2006)
9. Reinoso, E.; Maldonado, C.; Muñoz, R.; Damiano, L.; Abrutsky, M. : Bases de Datos , Edit. AlfaOmega Editores , Argentina , ISBN: 978-987-1609-31-4. (2012)