

**Modelado de Datos en el Mundo Real:
Teoría vs. Práctica**

*Clase 2:
Modelado de Datos: Teoría vs. Práctica*

Laura Rivero, Viviana Ferraggine
2010 Grupo Base de Datos y Procesamiento de Señales

Modelado de Datos: Teoría vs. Práctica

- Las reglas del negocio:
 - su entendimiento
 - sus objetivos.
- La incorporación del factor humano:
 - la gente como un recurso,
 - hablando con la persona correcta,
 - obteniendo la información apropiada.
- Lidiando con escenarios desfavorables.
- Conversiones de sistemas legacy.

Reglas del Negocio

- El proceso de modelado de datos incluye la elaboración de **los modelos de las operaciones de la organización**.
- Las reglas del negocio a veces son una **zona gris** porque tratan de definir como opera una organización.
- No puede haber diferencia entre la interpretación del usuario y la del modelador de la base de datos.
 - Los usuarios **finales** conocen "todos los hechos".
 - Los usuarios **gerenciales** desean la información resumida.
 - Por ejemplo al empleado le interesa conocer el informe diario del saldo en las cajas de cada sucursal, al gerente con el informe semanal por región le alcanza y al gerente general solo le interesa saber el saldo de la empresa →

Diferentes enfoques de una misma información

Reglas del Negocio

- Obtener la **información correcta** de las **personas adecuadas** es fundamental para el proceso de pre-diseño.
- Las reglas de negocio son una parte esencial del modelo de datos.
 - agregar **valor semántico** a los objetos de datos
 - es imperativo trabajar con ellos de una **manera estructurada**.
- La mayor parte del tiempo empleado en el diseño de un modelo de datos debería gastarse en **descubrir y documentar**, con precisión (sin ambigüedades) las reglas de negocio → **esto pueden marcar la diferencia entre el éxito y el fracaso de un SI**.

Entender las reglas y los objetivos

- Un modelo de datos se crea utilizando pasos metodológicos que contengan objetos claramente definidos.
 - esto no garantiza que se ajuste a las reglas que la organización posee.
 - P. ej. un modelo relacional brinda la posibilidad de controlar ciertas reglas básicas como son las relaciones entre los datos → esto no garantiza que se cubran **todas** las reglas del negocio de la organización
- La **teoría de conjuntos** combinada con el **cálculo de predicados** son un modo formal para capturar modelos de datos y sus Reglas de Negocio (restricciones).

Que se define como Regla del Negocio?

- Para una organización son esencialmente **procesos y flujos** de todo lo que está involucrado con el trabajo diario de esa organización → la operación de su actividad empresarial y las decisiones adoptadas para ejecutar los procesos operativos.
 - ¿Que tiene que hacer la organización para funcionar?
 - ¿Cuál es su tarea?
 - ¿Cuál es el significado y propósito de su existencia?
- En esencia, las **Reglas de Negocio** representan cualquiera de los siguientes aspectos:
 - **Políticas** de la organización
 - **Cálculos o fórmulas** (por ej. los cálculos de amortización del préstamo para una compañía de préstamos hipotecarios).
 - **Normas** (por ej. los requisitos estándar de la industria)

Que se define como Regla del Negocio?

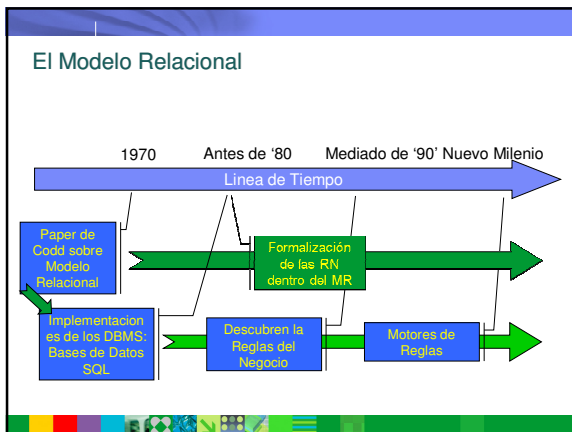
- Las reglas del negocio referidas a modelos de datos van desde
 - implementaciones simples como relaciones entre objetos a
 - implementaciones complejas codificadas en aplicaciones o en procesos de la base de datos.
- Las restricciones son los requisitos que se imponen a los valores almacenados en una base de datos.
 - P.ej. los valores de un campo en los registros de una tabla pueden ser validados haciendo que sólo puedan contener la letra M o F para representar Hombre o Mujer.

Que son las Reglas del Negocio?

- Algunas definiciones ...
 - “Una sentencia que define o restringe algún aspecto del negocio. Su propósito es asegurar la estructura del negocio o controlar o influir en el comportamiento del mismo.”
 - “Un requisito precisamente articulado soportado por una computadora y de fácil lectura”
 - “Una restricción ejercida con el fin de mantener la exactitud de los datos persistentes”
 - “Un conjunto de condiciones que rigen un evento del negocio para que ocurra de un modo aceptable.”
 - “Las RN incluyen reglas de presentación, de bases de datos y de aplicación”

La importancia de las Reglas del Negocio

- No se puede decir que un MD no cuenta con reglas del negocio, ni tampoco decir que esta sobrecargado de ellas, ambas afirmaciones son falsas.
 - Un MD NO puede evitar definir reglas de negocio, por ejemplo la simple existencia de relaciones entre objetos, el permitir o prohibir valores nulos en campos individuales o cualquier tipo de procesamiento de validación compleja.
 - Otras veces se utilizan:
 - procedimientos almacenados (stored procedures) para realizar chequeos complejos
 - triggers para disparar acciones específicas ante determinados eventos.
- Nota:** su uso debe ser cuidadoso para no degradar el rendimiento.



La incorporación del factor humano

¿Qué se entiende por factor humano?

- No se puede crear un modelo de datos sin la participación de las personas, sean éstas técnicas, programadores, usuarios finales, o ejecutivos → Hay que tener en cuenta que los empleados saben COMO funciona la empresa.
- Diseñadores experimentados aprovechan las experiencias del pasado ante cada nuevo desafío, pero ... No hay que olvidar que existen situaciones extraordinarias en cualquier empresa !!! → con la experiencia solamente no basta → Se recurre a la metodología.
- Se puede encontrar lo que se necesita para el modelo hablando con las personas que lo utilizarán → Hablar con los desarrolladores y los usuarios finales.

La gente como recurso

- La primera tarea es hablar y escuchar al personal de la empresa → ellos puede aportar muchos detalles sobre los objetos del modelo y sus relaciones.
 - Hay que recordar que los usuarios son por lo general personas no técnicas y no conocen sobre modelos de datos → Se debe tener cuidado con que herramientas se habla de los datos con los usuarios finales → NO siempre es posible hablar a la luz de un ER con los usuarios.
- El diseñador de un MD aplica sus conocimientos técnicos a las declaraciones tomadas de los usuarios (ya sean verbales o escritas).
- El diseño de un MD es en gran medida un proceso de planificación, y pensar diferentes opciones es una gran parte de ese proceso de planificación.

La gente como recurso

- El diseñador de un modelo de datos **debe**:
 - **Organizar** todo en secuencias de pasos lógicos, matemáticos.
 - **Obtener** las piezas para colocarlas en un modelo lógicamente estructurado.
 - **Averiguar** todo sobre la organización lógica.
- Los **usuarios finales conocen todo de su negocio**, pero probablemente lo piensan en términos de sus prioridades y de cómo reaccionan ante diferentes situaciones.
- Todo es subjetivo en esta etapa → El diseñador de un MD debe ser **objetivo, crítico, analítico, y sobre todo exacto !!**

La gente como recurso

- Los usuarios finales ven todos los detalles → los diseñadores de un MD deben buscar elementos comunes para la abstracción.
- Un MD debe tener en cuenta todas las circunstancias especiales → pero debe seguir siendo abstracto.
- No se debe
 - **perder el foco de lo abstracto,**
 - **dejar de representar información relevante**
- ¿Alguna vez se encontraron con un diseño de un MD en una empresa, que cubre una pared entera? → esta es una situación que asusta !!!!!

Hablando con la persona adecuada

- Existen situaciones en las que se debe entrevistar a personal:
 - Técnico o administrativo
 - De niveles gerenciales medios
 - Ejecutivo
- ¿Cuál es la persona con la que un modelador de datos debe hablar?
 - los administrativos o técnicos → conocen todos los detalles
 - los ejecutivos → grandes organizaciones vs. organizaciones medianas
 - gerentes de alto nivel → poseen una buena visión global
- Un diseñador de un MD debe hablar con diferentes tipos de personas en múltiples niveles y con múltiples habilidades dentro de la empresa.

Hablando con la persona adecuada

- En general, el nivel de detalle requerido depende de la complejidad de la organización y qué parte de esa complejidad va a ser informatizada.
- Cuanto mayor es la complejidad de un MD, mayor es la cantidad de preguntas que debe hacer un diseñador → más gente puede que sea necesario involucrar...!!!
- Un MD debe tener la simplicidad adecuada → el trabajo de un modelador es hallar la simplicidad a través de la abstracción → es necesario integrar los diferentes aspectos de las operaciones de la organización en un conjunto conciso de tablas y relaciones.

Hablando con la persona adecuada

Solo hablar con:

- Técnico o administrativo: pueden existir situaciones en la que los técnicos planteen muchos obstáculos → en general la consulta a los usuarios finales da una nueva perspectiva
- Ejecutivo: pueden dejar de lado detalles cruciales → ellos deben señalar a los empleados adecuados.

Obteniendo la información correcta

¿Cuál es la información correcta?

- Existe una ventaja al consultar a las personas que no están involucradas directamente en una tarea para que no se sientan amenazadas por la introducción de cambios...!!
- Los ambientes difíciles son comunes → A veces, un consultor bien puede conducir el proceso, o ayudar a eludir obstáculos. Por el contrario, existen situaciones en las que la gente de la organización es la mejor opción.
- Si una situación es totalmente negativa, tal vez usted debería tratar de utilizarla para su propio beneficio al hacer la tarea de análisis.

Obteniendo la información correcta

- Para obtener la **información adecuada** es necesario hablar con las **personas adecuadas**, dado que ellas saben cómo funciona la organización
- Cuanto mayor es la cantidad de gente que se entrevista mayor es el detalle que se obtiene → Riesgo: **versiones encontradas que se deben conciliar**.
- La **combinación de las habilidades** de un modelador de datos con el conocimiento de otros puede ayudar a construir un excelente diseño.
- Es necesario **equilibrar** las diferentes perspectivas y opiniones para encontrar la solución mejor.
- La postura del diseñador debe ser **firme** pero debe escuchar todas las posibilidades con **humildad**.

Lidiando con escenarios desfavorables

- En cualquier escenario siempre es necesario **obtener la mayor información posible** sobre lo que el nuevo modelo se supone debe lograr. → tomar la información desde los sistemas legacy, planillas, archivos en papel, etc. → en general es necesario partir desde abajo.
- **Informatizar de un montón de papeles: Sistemas basados en papel:**
 - Puede ser el problema más fácil a resolver o el más difícil, estos sistemas en general son muy **precisos y detallados** y es necesario tomar en cuenta todas las posibles eventualidades que han sido almacenadas durante muchos años.

Lidiando con escenarios desfavorables

- No ha sido meticulosamente diseñado y ha ido creciendo crecido de acuerdo a la medida de las necesidades (de manera **caótica**). Existen entonces:
 - diferentes tipos de papeles,
 - con una gran cantidad de información duplicada,
 - contradictoria,
- En esta situación extrema, pueden **utilizarse los formatos básicos** de documentos en papel y establecerse los **aspectos operativos**.

Lidiando con escenarios desfavorables

Como encontrar la persona que en una organización sabe cómo funciona todo?

- Deben averiguarlo los modeladores...
- Puede que no exista tal persona (situación extrema de confusión total) → hay que construir el modelo con cautela y obtener algún tipo de verificación de personas que conozcan el funcionamiento operativo de la empresa
- **Ventaja:** transmitirán sus conocimientos, podrán señalar otros factores a tener en cuenta (campos a agregar, casos especiales, etc.)

Lidiando con escenarios desfavorables

- La manera más fácil es juntar tanto material impreso como se pueda, clasificarlo (si es con la ayuda de personal de la empresa mejor)
- La informatización de un montón de papeles **no es una tarea desalentadora** → porque están tratando de liberarse de ese montón de papel que constantemente tiene que llenar y controlar.

Lidiando con escenarios desfavorables

Conversiones de sistemas legacy

- Puede ser la más difícil de las tareas porque:
 - Son **parcialmente inaccesibles** o de difícil acceso.
 - Tienen **modelos de datos extremadamente grandes y complejos.**
 - **Encontrar a la persona** que posee el conocimiento de todo puede resultar imposible → el modelador debe asumir la tarea de analizar y verificar la estructura → confirmar la funcionalidad operativa del modelo descubriendo si hace en realidad lo que debe hacer.
 - **El modelo se construyó de manera incorrecta,** o las necesidades han cambiado → es necesario encontrar a alguien que le haga evidentes esa diferencias al modelador.

Nota: Es mejor buscar un experto, se obtendrá mas información y más rápidamente haciendo preguntas.

Lidiando con escenarios desfavorables

Integración de bases de datos heterogéneas

- Un sistema heterogéneo es un sistema compuesto por partes diferentes.
- En términos de modelos de BD, algunos motores poseen herramientas que hacen posible la integración homogénea de bases de datos heterogéneas. Es posible vincular de forma transparente múltiples tipos de bases de datos.
- En otros casos esta integración se debe hacer manualmente, con procesos ad-hoc.
- Además, todo el software informático, incluyendo BD y sus modelos correspondientes tiene un ciclo de vida útil con límites aceptables en el costo de su efectividad → se llega a un punto en el que el sistema "viejo" es demasiado caro de mantener o puede ser reemplazado fácilmente, entonces no hay razón para no volver a escribir el software.

Lidiando con escenarios desfavorables

Conversión desde hojas de cálculo

- En general son más fáciles que desde sistemas legacy o papel.
- Al principio es divertido, porque se ven como archivos planos, pero es posible encontrar todo tipo de complejidades con fórmulas en múltiples niveles e incluso páginas relacionadas.
- Se requieren de análisis y especificación de requisitos de lo que debería existir.

Lidiando con escenarios desfavorables

Reordenamiento de una base de datos inconsistente

- Significa que el modelo de base de datos es un completo desastre.
- Se encontrarán datos no válidos, registros sin vincular, y otros maravillosos problemas!!!.
- Hay que establecer primero lo que se necesita antes de entrar en el terreno de los datos.
- Puede parecer una tarea desalentadora, pero tiene dos etapas bien diferenciadas.
 - Primero, establecer y crear la estructura correcta.
 - Segundo copiar los datos a través de nuevas tablas.

Tips y sugerencias

- Aunque varios de los consejos y sugerencias que se ofrecen son el resultado de largas experiencias y conocimientos de expertos de profesionales de modelado de datos, estos no son balas de plata.
- En cada proyecto de modelado de datos, habrá casos especiales y situaciones inusuales. Cuando usted se enfrente a tales circunstancias, estos consejos le ayudarán a resolver problemas especiales y encontrar soluciones eficaces.
- Las sugerencias se refieren sobre todo a las fases importantes del desarrollo de modelado de datos: el **diseño conceptual y lógico**. La modelización de datos físicos no está incluida.

Tips y sugerencias

- **No hay dos situaciones en diferentes proyectos de modelado de datos que sean exactamente iguales** → las **soluciones también pueden diferir...!!**
- Los tips y sugerencias que se detallan son a nivel general.
Cómo utilizarlas?
 - Cada proyecto es distinto → las condiciones varían → las situaciones son diferentes... pero... cada sugerencia puede apuntar a una forma de enfrentar una situación similar.
 - Sin embargo, resulta útil mantener un inventario general de la gama de consejos.

Definición de requisitos

- La definición de requisitos es un esfuerzo repetitivo → hay esfuerzo repetitivo en el modelado de datos → mala definición de requisitos → modelo de datos incorrecto.
- Algunos consejos prácticos:
 - **Entrevistas:** método útil para recopilar requisitos
 - No realizar entrevistas sin la preparación adecuada.
 - Todas las entrevistas deben tener una lista publicada de temas a tratar.
 - El usuario también debe estar preparado y listo para discutir.
 - Durante las entrevistas no debe haber interrupciones.
 - Mantener el tamaño del grupo de usuarios a ser entrevistados menor a 3.
 - Entender claramente la información recibida de los usuarios.
 - Entrevistar a todos los usuarios de todos los niveles (ejecutivos, gerenciales y operacionales).
 - Documentar los procedimientos de cada entrevista y obtener el documento revisado y confirmado por el usuario en particular.

Definición de requisitos

- **Sesiones de Grupo:** suelen ser útiles para confirmar requisitos parciales de entrevistas individuales.
- **Algunos consejos prácticos:**
 - Deben tratarse como una reunión formal, teniendo agendas publicadas con suficiente antelación.
 - Establecer tiempo límite y procurar que los actores clave estén presentes.
 - Es difícil organizar sesiones de grupos de muchas personas → mas comunes: mesa redonda vs. presentación.
 - Establecer métodos para llegar a un consenso del grupo.
 - Asignar responsabilidades y plazos. Nombrar representantes de los usuarios para el seguimiento.
 - Documentar acciones en un resumen de formato de discusión y distribuir.

Definición de requisitos

Grupos geográficamente dispersos:

- Ser consciente de las lenguas y culturas locales.
- Los métodos de entrevistas y sesiones de grupo deben ser los mismos; la coordinación y organización es mas compleja.
- Tratar de utilizar el idioma local.
- Tratar de que no se formen minigrupos conversando en el idioma local → establecer normas y directrices aceptables desde el principio.

Definición de requisitos

- **Documentación:** es esencial, constituye la base para el modelo de datos.
- **Algunos consejos prácticos:**
 - Debe ser la justa: "ni poco, ni demasiado"
 - Debe estar bien organizado. Subdividirlo de acuerdo a departamentos, funciones, ubicación geográfica, etc.
 - No esperar hasta el final de la fase de definición de requisitos para iniciar documento.
 - Acompañar el diagrama de modelo de datos con la información de los requisitos del documento de definición.
 - Prever la aceptación de los cambios y revisiones que se realizan en la documentación.
 - Por un período de tiempo, es probable que se superpongan la recopilación de requisitos y el modelado de datos.
 - No siempre es factible congelar los requisitos al iniciar la etapa de modelado
 - En desarrollos ágiles, los cambios se deben tener en cuenta lo más pronto posible sin extender el tiempo del proyecto.

Definición de requisitos

- **Gestión del Cambio:** en desarrollos iterativos, es necesaria a medida que se avanza en el desarrollo (se aleja de los enfoques tradicionales) aceptar la idea de prever los cambios desde el inicio del proyecto.
- Algunas sugerencias:
 - Establecer un procedimiento de gestión del cambio definitivo → informárselo a los clientes y convencerlos de la necesidad de usarlo con prudencia.
 - Discutir claramente con los usuarios cómo van a solicitar un cambio y cómo serán incorporados al modelo.
 - Documentar todos los cambios al inicio, su razón y la incorporación del cambio.
 - Establecer un proceso de aprobación para filtrar los cambios.
 - El modelado es incremental → hay que referirse a cada versión del modelo de datos → hay que documentar el estado de cada versión.
 - Debe haber puntos de corte: establecer puntos de corte de cada modelo parcial o establecer puntos de corte por iteraciones.
 - Evitar el esfuerzo excesivo es importante, pero no es todo el proyecto.

Definición de requisitos

- **Notas para el modelado de datos:** porque son útiles por separado? Son mantenidas por modeladores de datos para su uso específico.
- Algunos consejos:
 - Asegurarse de que las notas están sincronizadas con los requisitos.
 - Anotar todos los cambios al modelo de datos a medida que se solicitan.
 - Puede ser un buen método dividir las notas en secciones separadas para los tipos de entidades, atributos identificadores y relaciones y anotar los casos especiales en cada sección.
 - Tener una sección separada para dudas. Tomar nota de las resoluciones.
 - Si se está modelando para una organización global es necesario tener la información de contacto con los números de teléfono resulta ser muy útil.

Participación de los Stakeholders

Cualquiera sea el grupo de TI que tenga un interés directo en el resultado del desarrollo de software, deben estar incluidos en esta amplia categoría.

- Son personas expertas en el contexto del desarrollo.
- Son los que continuamente pueden dar su opinión sobre los requisitos.
- Se consideran cuatro factores distintos de participación:
 - como organizar,
 - quien es el usuario de enlace,
 - interacción continua de los usuarios,
 - qué hacer cuando los interesados se encuentran en múltiples lugares distantes.

Participación de los Stakeholders

Modelado Iterativo: El modelado iterativo e incremental van de la mano.

- Se divide el modelado en partes manejables → se trabaja de manera incremental en una parte a la vez → se integra la parte anterior a la actual.
- Se establece un corte inicial del modelo, se presenta la versión inicial a los usuarios, se obtienen sus comentarios, se perfecciona y se produce la próxima versión. Se definen entonces:
 - **Ciclos para cada iteración:** incluye una fase de creación, revisión y retroalimentación de fase, y una fase en la que se refina el modelo.
 - **Incrementos lógicos:** interacción entre los requisitos de definición y el modelo de datos

Participación de los Stakeholders

Es necesario tener en cuenta:

- Baja cantidad de Interacciones (2 o 3 son prácticas) y definir el propósito de cada ciclo.
- Preparar y usar una lista de tareas para cada etapa.
- Establecer las responsabilidades en cada fase y definir claramente la participación del usuario.
- Determinar la duración de cada fase para cada fragmento del modelo.
- Evitar fases largas, si es necesario rearmar el tamaño de los fragmentos.
- Incrementos: como definirlos? Seleccione el tamaño del modelo de fragmentos de acuerdo a la experiencia de los modeladores y los usuarios

Bibliografía

Beginning Database Design. Gavin Powell (Chapter 2.)

Data Modeling Fundamentals. A Practical Guide for IT Professionals. Paulraj Ponniah (Chapter 12)
