

Organizar datos de negocio con modelado de datos

La creación de modelos de datos puede ayudar a los negocios a organizar y acceder mejor a los datos. En este curso, va a conocer los tipos y tareas básicos de modelos de datos y de qué forma los tres niveles de modelado de datos pueden proporcionar un plano detallado de los datos de una empresa.

Tabla de contenido

1. [Organizar datos de negocio con modelado de datos](#)
2. [Introducción al modelado de datos](#)
3. [El poder del modelado de datos](#)
4. [El proceso de modelado de datos](#)
5. [Tipos comunes de modelos de datos](#)
6. [Estilos de modelado de datos](#)
7. [Construir su modelo de datos](#)

Organizar datos de negocio con modelado de datos

Las organizaciones se basan en varios tipos de datos para comprender plenamente su negocio. Para hacerlo, se basan en datos de fácil acceso y comprensión. El modelado de datos juega un rol crítico en ambos aspectos.

En este curso, conocerá las características del modelado de datos, así como los estilos y tipos comunes de modelos de datos usados. También aprenderá estrategias comunes para crear modelos de datos, y cómo de esa forma mejorar las operaciones diarias de una organización.

Introducción al modelado de datos

El modelado de datos es un tema complejo. Aunque en términos simplificados, consistiría en el proceso de estructuración y organización de un conjunto de datos. Y el modelo de datos representa las interacciones entre los diferentes objetos de datos. Considere el proceso de construcción de una casa; el diseño, dimensiones, materiales, etc., son los datos que utilizan los constructores para la construcción. El plano es el modelo de datos que muestra cómo interactúan los datos. Este tipo de acceso a datos fácilmente legibles en un entorno de negocio juega un rol crítico a la hora de ayudar a usuarios finales y organizaciones a acceder a los datos para tomar decisiones de negocio impactantes y bien fundamentadas. El término "modelo de datos" hace referencia a dos cosas distintas.

La primera es la propia descripción de la estructura de datos. Y la segunda aborda la organización de los datos dentro de la estructura. El modelo de datos puede integrarse

fácilmente en una base de datos. Aunque recuerde que en función del sistema de gestión de base de datos que use una organización, el modelo de datos podría establecer ciertas limitaciones en los datos contenidos dentro del propio modelo. Los desarrolladores también usan modelos de datos para registrar y organizar datos empresariales de tal forma que se facilite la comunicación entre los elementos de datos.

Estos modelos ayudan a los patrocinadores del negocio y desarrolladores a acordar con precisión los datos que se representan y qué significan, y a compararlos con su comprensión. Los modelos promueven el consenso entre interesados. Aunque un modelo de datos es abstracto, suele representarse con un diagrama que se usa para planificar cómo se van a almacenar los datos y al desarrollar software. Un modelo de datos de una base de datos ayuda con el diseño, análisis, registro y transferencia de sus estructuras de base de datos elementales. No hay ningún modelo de datos estándar, y sin importar cómo se expresen los datos en un sistema informático, cada modelo de datos establece sus propias reglas de lenguaje para el dominio que contiene los datos.

Esto hace que los datos de dos modelos estructurados de forma diferente resulten difíciles de integrar. Vamos a considerar de qué forma los modelos de datos van a representar y organizar datos. Pensando en la casa que queremos construir, organizamos por clases de entidades, que son las cosas que representa la información, como dormitorios, cuartos de baño, materiales y personal especial necesario. A continuación, el modelo desglosa las entidades en atributos. Por ejemplo, atributos para dormitorios incluyen especificaciones de uso y dimensiones. Los atributos para materiales pueden incluir panel de yeso, suelo, componentes eléctricos y fontanería. Y atributos para personal serían forjadores, fontaneros, electricistas y albañiles. El modelo de datos muestra la relación entre entidades, es decir, la relación entre dichas personas, habitaciones y materiales.

Finalmente, el modelo de datos ilustra la relación entre atributos. Por ejemplo, la relación entre personal especializado y sus materiales. Ya que el modelado de datos organiza datos de forma tan efectiva, ayuda a los negocios a funcionar con más eficiencia. Por ejemplo, el modelado de datos permite a los negocios acceder a información de productos, clientes e ingresos para ayudar a personalizar el desarrollo de sus productos e iniciativa de marketing con el fin de dirigirse a los clientes de forma efectiva, ofrecer productos actualizados y mejorar el soporte al cliente.

El poder del modelado de datos

Dirigir un negocio es algo desafiante. Existen restricciones de presupuesto, requisitos de recursos y la necesidad de actualizarse con nuevas tecnologías que se introducen casi todos los días. Para ayudar a mantener la ventaja competitiva de su organización, los responsables de la toma de decisiones deben ser capaces de usar modelos de datos para evaluar y mejorar distintos aspectos del negocio, en función de datos empresariales explícitos.

En particular, puede que las organizaciones grandes tengan numerosas bases de datos que contengan información que no necesite integrarse necesariamente con otras bases de datos. Y las necesidades de los usuarios de estas diferentes fuentes de datos pueden ser diferentes,

lo que podría requerir un modelo de datos diferente. Al tratar datos como un activo estratégico, no solo obtiene un retorno de la inversión, sino un retorno de la oportunidad, que aumenta significativamente cuando lleva esos datos al negocio y hace que sean de fácil acceso. Ese es el motivo por el que cuando se crean modelos de datos para cada aspecto de las operaciones del negocio, proporcionan una vista precisa de 360 grados de la organización y de sus unidades individuales.

De la misma forma que no intentaría construir una casa sin planos, una base de datos no debe construirse sin un modelo de datos en el que basarse. Un modelo de datos contiene varias capas de información. De una descripción general de alto nivel a la distribución detallada más básica. Es una ilustración gráfica que proporciona un plano para diseñadores de bases de datos y varios interesados que extraen diferente información de él. El modelado de datos no solo proporciona un lenguaje preciso y común para comunicarse con usuarios finales, permite que los interesados accedan a datos significativos relacionados con el negocio, independientemente de dónde se ubiquen.

Estos datos se usan para encontrar y asimilar fuentes de datos actuales para su análisis, y determinan los siguientes pasos para áreas de mejora del negocio. El sistema que usa un departamento de ventas es muy diferente del sistema que usa RR. HH. Por lo tanto, los interesados usan modelado de datos de la organización para diseñar e implementar fuentes de datos de calidad superior a través de sistemas dispares. Y ya que el modelo es gráfico, es más fácil realizar un análisis comparativo visual y coordinarse con varios modelos de datos a través de plataformas. Además, y en función de la necesidad del negocio, varios usuarios finales usan modelos de datos para gestionar procedimientos y equipos de modelado de datos y explorar opciones de modelado de datos de autoservicio como, por ejemplo, servicios en la nube.

Adicionalmente, el modelado de datos se usa para documentar configuraciones de metadatos y feedback de interesados a través de un bucle de feedback. Un bucle de feedback captura datos, convierte los datos en información utilizable y actúa sobre la información. Por ejemplo, si un cliente compra un artículo específico en un sitio web, un bucle de feedback ayuda al sitio a sugerir otros artículos en función de los datos de compra previos. Una mejor comprensión conduce a mejores decisiones y tomar mejores decisiones permite a las organizaciones estar por delante del resto.

El uso de todos los datos básicos que estén disponibles debe ser una parte natural del proceso de toma de decisiones de su organización. El modelado de datos permite a los negocios eliminar silos organizacionales y conectarse con interesados para mejorar la fluidez de los datos, responsabilidad y colaboración. Aquí es donde comienza la transformación digital, el primer paso en la redefinición de su negocio, mercados y relaciones con los clientes.

El proceso de modelado de datos

El proceso de modelado de datos implica varias tareas, siendo la primera identificar entidades o tipos de entidades. Un tipo de entidad describe un grupo de objetos similares, como un grupo de personas, lugares o cosas, cada uno de ellos con su propio concepto diferenciador. Para una empresa de fabricación, estos podrían ser clientes, productos, materiales, I+D e instalaciones. Cada entidad puede tener más de un atributo. Por lo tanto, la siguiente tarea en el modelado de datos es identificar los atributos de cada tipo de entidad. En una base de datos relacional, una columna representa el atributo de datos. Por ejemplo, en un modelo de datos lógico básico, la entidad cliente, bajo el encabezado de columna, tendría atributos como el nombre y apellido.

En un modelo de datos físico básico, se llamaría la columna TCliente, por ejemplo. Y las columnas tendrían NOMBRE_CLIENTE y APELLIDO_CLIENTE en su lista de atributos. Posteriormente se aplican convenciones de nomenclatura. Tenga en cuenta las diferencias entre las convenciones de nomenclatura de modelos de datos lógicos y físicos. En el modelo lógico, el encabezado de columna es Cliente. Pero en el modelo físico, el encabezado de columna es TCliente. El modelo lógico busca la legibilidad del usuario final, mientras que la convención de nomenclatura física refleja especificaciones técnicas. Las directrices y estándares de nomenclatura deben cumplir los protocolos de codificación de la organización y deben estandarizarse. La siguiente tarea implica identificar relaciones entre entidades.

Por ejemplo, un cliente se ubica en una dirección y tiene una cuenta. "Vive en", "Usa" y "Tiene", describen relaciones entre entidades. Una tarea adicional consiste en aplicar patrones de modelos de datos. Aunque los modelos de datos son exclusivos para las necesidades de una organización, tienen patrones comunes que pueden reutilizarse a través de varias plataformas. Por ejemplo, las facturas para materiales de construcción son similares a facturas para clientes. Independientemente de que los datos traten sobre una persona o producto, la clasificación del patrón sigue siendo la misma.

A continuación, asignamos claves naturales y subrogadas a las tablas. Una clave es un paradigma de almacenamiento de datos diseñado para almacenar y recuperar fragmentos de datos. Una clave natural hace referencia a uno o más atributos particulares para el concepto del negocio. La clave subrogada, que es solo una nueva columna, hace referencia a una clave sin valor específico. En nuestro ejemplo, una clave natural sería "número de cliente". Sin embargo, las direcciones no tienen una clave natural. Cada línea de la dirección necesitaría formar una clave por sí misma.

Generalmente, la dirección se usa como una clave subrogada. La siguiente tarea consiste en reducir la redundancia de los datos. Supone un desafío que los mismos datos se almacenen en varios lugares en una base de datos relacional. Por esa razón, los desarrolladores de aplicaciones se centran en la normalización de los datos, una técnica para organizar los atributos de los datos dentro de un modelo de datos. Por lo tanto, los tipos de entidades están más conectados, y se evita la duplicación. La tarea final consiste en mejorar el rendimiento. Ya que los esquemas de datos normalizados se centran en reducir la redundancia de los datos más que en rendimiento, pueden esperarse algunos problemas de rendimiento, como por ejemplo una ralentización del acceso a los datos. Para acelerar el

proceso y mejorar el rendimiento, podría considerarse la desnormalización de algunas partes del esquema de datos.

No obstante, debe asegurarse de que los beneficios compensen cualquier perjuicio. Ya que se duplican algunos datos, la desnormalización de datos requiere más espacio, y, algo que preocupa más, introduce la posibilidad de anomalía en los datos, ya que estos pueden alterarse en más de un lugar.

Tipos comunes de modelos de datos

No importa lo bien que estén modelados, los datos tienen que vivir en algún sitio para permitir que los usuarios los obtengan. Necesitan una base de datos. Existen varios modelos de bases de datos que pueden elegir las organizaciones en función de sus necesidades y de lo que admita su sistema de gestión de bases de datos actual. La mayoría se construyen teniendo en cuenta un modelo de datos particular y requieren que los usuarios adopten dicho modelo.

En el modelo relacional, los datos se disponen en columnas y filas, como en una tabla, denominada relación. Cada columna contiene un atributo de una entidad específica como, por ejemplo, empleado, departamento o proyecto. El grupo de atributos en una relación se llama dominio. Cada fila contiene datos en relación con una instancia particular de la entidad específica, como por ejemplo datos de un empleado particular o información de proyecto. En la base de datos, las tablas pueden normalizarse, lo que hace que sean flexibles, adaptables y escalables. Cuando se normalizan, cada pieza de datos se desglosa en piezas útiles más pequeñas.

En cambio, el modelo de base de datos de documentos almacena todos los datos sobre un objeto específico en una instancia individual. Los objetos almacenados no necesitan seguir una estructura establecida; cada objeto puede ser único. Esto simplifica el trazado de la base de datos y beneficia a aplicaciones que se actualizan de forma regular, y en casos donde la velocidad de despliegue es una prioridad. En el modelo jerárquico, los datos se disponen como en un organigrama.

Para el almacenamiento de datos, el orden físico emula el mismo orden en el que se disponen registros hermanos, cada uno de ellos con una rama matriz. Este modelo es adecuado para ilustrar relaciones en el mundo real, pero no se suele usar debido a ineficiencias operativas. El modelo de red es similar, y contiene conjuntos de registros que están relacionados; cada registro puede incluirse como un miembro en numerosos conjuntos. Es adecuado para relaciones "muchos a muchos" complejas, como por ejemplo tareas de mantenimiento y materiales, donde un material puede ser utilizado por más de una persona para más de una tarea.

Los modelos de bases de datos orientadas a objetos definen una base de datos como una colección de objetos o elementos de software reutilizables con sus características y métodos asociados. Existen varios tipos de bases de datos orientadas a objetos, incluyendo modelos híbridos. Por ejemplo, una base de datos de hipertexto permite la vinculación de cualquier

objeto con otro, y una base de datos multimedia integra archivos multimedia. Una base de datos de modelo relacional de objetos incluye lo mejor de los dos mundos. Toma las funciones progresivas del modelo de base de datos orientada a objetos y las combina con la simplicidad del modelo relacional.

Esto permite que los diseñadores asimilen objetos en una estructura de tablas. Por otra parte, y similar al modelo de red, el modelo de base de datos de relaciones de entidades traza relaciones entre entidades reales. Se usa para el diseño conceptual de una base de datos, donde el dominio se compone de atributos de entidades, como contactos de clientes. Una forma común de tipo de relaciones de entidades se denomina esquema en estrella. En esta disposición, múltiples tablas dimensionales se conectan a una tabla central en forma de estrella. Finalmente, está el modelo de entidad-atributo-valor, o EAV. En este caso, la entidad representa el elemento de datos que se está describiendo, es decir, un empleado.

El atributo representa datos que describen la entidad, como por ejemplo un nombre y número de empleado. El valor es el valor de ese atributo, como define la empresa. EAV ofrece flexibilidad, aunque menor eficiencia, cuando se recuperan los datos de forma masiva. Cada modelo tiene sus propias fortalezas y debilidades. La elección de un modelo depende completamente de las necesidades de la empresa.

Estilos de modelado de datos

Imagine una obra de construcción en sus etapas iniciales. Hay camiones hormigonera, vigas de acero, madera, material de fontanería, cable eléctrico y una multitud de obreros. Ahora imagine que a estos obreros se les dice simplemente que pongan manos a la obra y que conviertan ese material en un edificio de oficinas. ¿No hay planos? No importa, solo hay que construir algo. Ahora imagine crear una base de datos sin un plan. El modelado de datos es como un plano para crear una base de datos real, y estos modelos pueden clasificarse en tres estilos principales. Normalmente, los requisitos de datos se documentan primero en un modelo conceptual. Desde una perspectiva del negocio, los modelos conceptuales ofrecen la descripción general de alto nivel de un problema que los datos pretenden abordar.

Se suelen usar para explorar conceptos con interesados. Los modelos conceptuales de alto nivel suelen crearse como parte de la visualización de los requisitos iniciales. Y se usan para explorar conceptos y estructuras de negocio estáticas de alto nivel, ya que están limitados a requisitos de datos iniciales. También se suelen crear antes de generar un modelo de datos lógico, o se usan como sustituto de alguno. De esto se desprende que el segundo modelo es el modelo de datos lógico o LDM, que convierte el modelo conceptual en una estructura más definitiva. Estos modelos se usan para explorar los conceptos de dominio y las relaciones.

Además, pueden usarse para un único proyecto o para toda una empresa. En algunos casos, un modelo de datos podría incorporar varios modelos de datos lógicos más pequeños prestando poca atención a cómo se implementarán realmente los datos en el modelo físico. A diferencia del modelado conceptual, el modelado lógico proporciona detalles exhaustivos

sobre datos, como por ejemplo todas las entidades, todas las relaciones entre entidades, y todos los atributos, claves principales y claves foráneas para entidades especificadas.

A nivel lógico se resuelven las relaciones "muchos a muchos", donde una tabla que contiene múltiples filas hijas se relaciona con una fila padre de otra tabla, y es donde tiene lugar la normalización. Un paso final adicional implica transformar el modelo de datos lógico en un modelo de datos físico, o PDM. El modelado físico proporciona una guía beneficiosa sobre la implementación de datos en una base de datos. Los modelos de datos físicos ilustran nombres de columna, tipos de datos y restricciones de datos en tablas, relaciones entre tablas, además de claves principales y foráneas.

Para diseñar un modelo de datos físico, un equipo debe convertir entidades en tablas, atributos en columnas, relaciones en claves foráneas, y considerar las restricciones físicas antes de actualizar el modelo físico. Considere cómo la complejidad y nivel de detalle aumentan en cada etapa del desarrollo de base de datos. Ese es el motivo por el que el modelado de base de datos suele empezar con un modelo contextual para obtener esa comprensión de alto nivel de varias entidades y cómo se relacionan. A continuación viene el modelo lógico para comprender todos los detalles. Y finalmente viene el modelo físico para mostrar los aspectos específicos de cómo implementar el modelo de datos en una base de datos elegida.

Las organizaciones pueden considerar el uso del Data-based Employment Manager, una herramienta de diseño basada en datos de software libre que permite la diferenciación entre estructuras de una base de datos e identifica las relaciones entre ellas. Juntos, estos modelos crean una base de datos integral en la que se basan usuarios finales, desarrolladores y responsables de la toma de decisiones para tomar decisiones críticas para el negocio.

Construir su modelo de datos

Construir una base de datos es como construir una casa. Si el propietario dice que quiere una casa grande con cinco dormitorios y un garaje para tres vehículos, el arquitecto tendrá una idea básica de las necesidades del propietario y puede empezar a dibujar los planos. De la misma forma, un modelo de datos de alto nivel o conceptual especifica lo que es importante para un negocio o proyecto.

Un modelo de datos conceptual es un medio para comunicar entidades o conceptos de negocio, y las relaciones entre dichas entidades con los usuarios finales. Este modelo suele ilustrarse en un diagrama de relaciones de entidades o ERD. A continuación nos movemos al nivel lógico.

Para el arquitecto, significa trazar la distribución de la casa. Implica hacer encajar todas las habitaciones y cómo hacer corresponder la habitación con su función. En el modelado de datos, significa determinar la distribución, tamaño y tipo de datos y las relaciones entre ellos. En el nivel físico se planifican detalles más técnicos, como por ejemplo la fontanería de la casa para el arquitecto y los detalles de la estructura de la base de datos para el

modelado de datos. Ya sea una casa o una base de datos, el modelo físico debe tener la capacidad de respaldar el modelo conceptual.

Una estrategia de modelado de datos común, el método de arriba a abajo, empieza con el nivel conceptual, se mueve al nivel lógico y finaliza con el nivel físico. Este modelo es apropiado para construir nuevos sistemas desde cero. Se centra en necesidades del negocio, tiene en cuenta la perspectiva general e incluye todas las áreas independientemente de si pueden implementarse.

Por el contrario, el método de abajo a arriba es aquel donde el enfoque se encuentra en construir sobre sistemas actuales. Este método resulta ideal si el objetivo es comprender la aplicación existente. Existen muchos beneficios del modelado de datos conceptual. Transmite los principios de la organización y conceptos complejos en un lenguaje no técnico que los usuarios finales pueden comprender y con el que pueden relacionarse. Este ejercicio común es crítico para comprender las reglas del negocio e información que rigen su organización.

A diferencia del método en cascada, el método estándar más tradicional para el desarrollo de software, el modelado de datos ágil deja espacio para la flexibilidad en caso de cambiar los requisitos del proyecto, y ayuda a gestionar dichos cambios. Considere el ejemplo de la casa. Con el método en cascada, los arquitectos solo consultarían al propietario antes de construir, y luego procederían a construir la casa de acuerdo con las instrucciones iniciales. Pero ¿qué ocurre si el propietario quisiera añadir otra habitación o cambiar una característica? Aquí es donde el modelado de datos ágil sería útil. Permitiría actualizaciones y mejoras sobre la marcha, lo que mejora la satisfacción del cliente.

Las organizaciones pueden usar múltiples estrategias para construir su modelo de datos. En definitiva, los modelos de datos deben proporcionar a toda la organización una forma de comunicar y abordar sus necesidades de negocio de forma fácil y efectiva.