

4.2.4 Integración de vistas

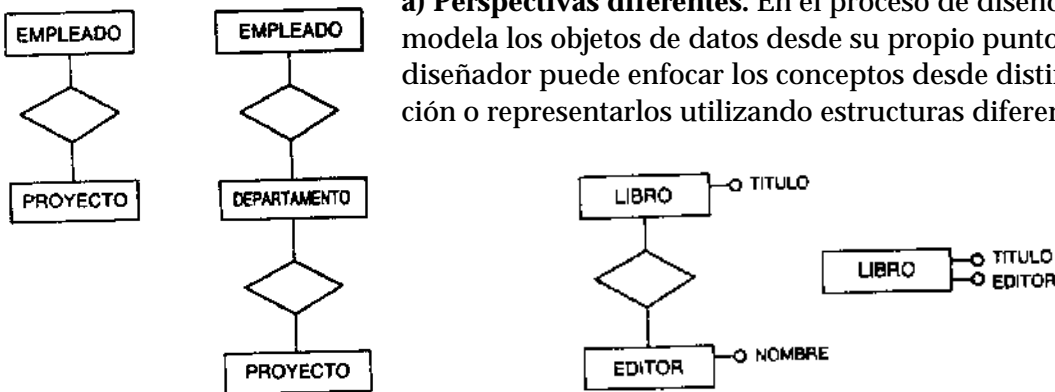
La integración de vistas es el proceso en el que se diseñan varios esquemas conceptuales individuales (vistas), y posteriormente se funden o combinan en un *esquema conceptual global*, que representa todos los requisitos de información del sistema. Este enfoque suele utilizarse con bases de datos de gran tamaño y complejidad.

El principal objetivo de la integración de vistas es encontrar todas las partes de los esquemas conceptuales (EC) de entrada que se refieren a la misma porción de la realidad, y unificar su representación. Esta actividad se llama integración de esquemas y resulta muy complicada, ya que una misma parte de la realidad suele estar modelada (representada) de forma distinta en cada esquema.

Aspectos de la integración de vistas

Una ventaja de este enfoque es que las vistas suelen ser pequeñas, por lo que su diseño resulta bastante sencillo. Sin embargo, la mayor dificultad de la integración de vistas radica en descubrir las diferencias y correspondencias que existen entre los esquemas que se van a fusionar.

Las diferencias de modelado pueden tener las siguientes causas:



a) Perspectivas diferentes. En el proceso de diseño, cada diseñador modela los objetos de datos desde su propio punto de vista. Cada diseñador puede enfocar los conceptos desde distinto nivel de abstracción o representarlos utilizando estructuras diferentes.

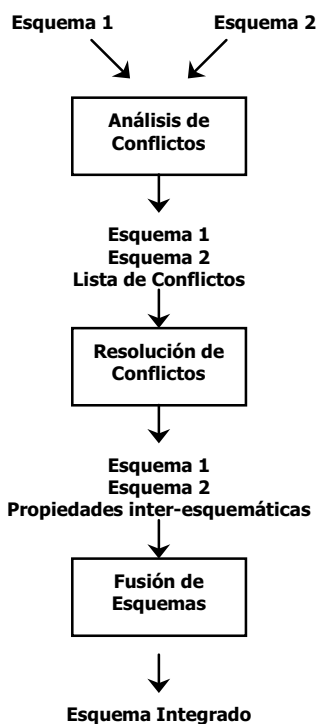
b) Equivalencia entre las construcciones del modelo. Los modelos conceptuales tienen una rica variedad de estructuras de representación, por tanto permiten distintas representaciones equivalentes de la misma realidad. Por ejemplo, la asociación entre un libro y su editor se representa en un esquema mediante una interrelación de las entidades LIBRO y EDITOR; y como un atributo de la entidad LIBRO en el otro. De forma similar, la partición de las personas en hombres y mujeres se puede representar con una jerarquía de generalización/especialización entre las entidades PERSONA, MUJER y HOMBRE en un esquema, y con una entidad PERSONA con un atributo *sexo*, en el otro esquema.

c) Especificaciones de diseño incompatibles. Los errores durante el diseño de vistas relacionados con nombres, estructuras y restricciones de integridad, pueden producir entradas erróneas para la de integración, por tanto, durante la misma, tales errores se deben detectar y corregir. Por ejemplo, el primer esquema de la figura de la izquierda indica que cada empleado *siempre* se asigna a un *único* proyecto; sin embargo el segundo esquema indica que cada empleado trabaja en múltiples proyectos. Los dos esquemas parecen correctos por separado, pero uno de ellos es erróneo de cara a su combinación.

Cada una de estas causas puede provocar un *conflicto*, es decir, representaciones diferentes de los mismos conceptos, o representaciones iguales para diferentes conceptos. La presencia de conflictos que se influyen mutuamente hace que el proceso de integración no sea trivial.

Es necesario un método para abordar la integración de las vistas en un esquema de base de datos global. Así, la integración de esquemas puede dividirse en las siguientes tareas:

1. Identificar y analizar los conflictos y las correspondencias entre los esquemas
2. Modificar las vistas para ajustarlas entre sí
3. Combinar las vistas
4. Reestructurar el esquema global.



Durante el **análisis de conflictos** se identifican los conflictos y las correspondencias; es importante identificar los conflictos cuanto antes.

Durante la **resolución**, se modifica uno de los esquemas, o ambos, para resolver o eliminar cada conflicto; esto debe hacerse conjuntamente con el diseñador.

Durante la **fusión de esquemas**, se combinan los esquemas y se obtiene un esquema preliminar integrado. Puesto que todos los conflictos ya han sido resueltos, esta actividad suele ser muy sencilla.

1. Identificación y Análisis de conflictos

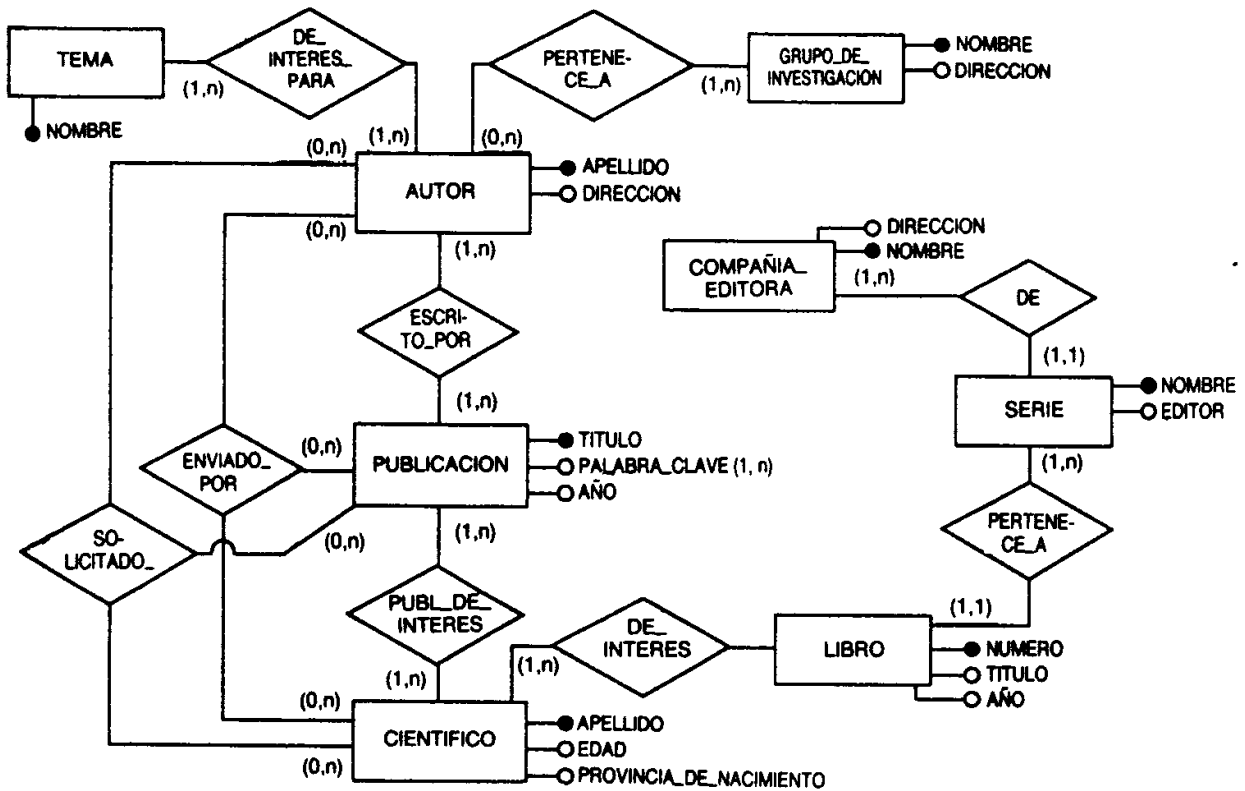
Utilizaremos un ejemplo relativo a la administración de una biblioteca para ilustrar el proceso de integración de un par de esquemas o vistas. Los esquemas de entrada se muestran en la figura de la página siguiente. El esquema 1 corresponde a la vista del científico y describe la estructura de la biblioteca privada de un investigador. El esquema 2 corresponde a la vista del bibliotecario y describe la estructura de la biblioteca central de un departamento. Las tablas 1 y 2 describen algunos de los conceptos de los esquemas 1 y 2, cuyos significados no son obvios.

Tabla 1. Conceptos del esquema 1 (esquema del científico)

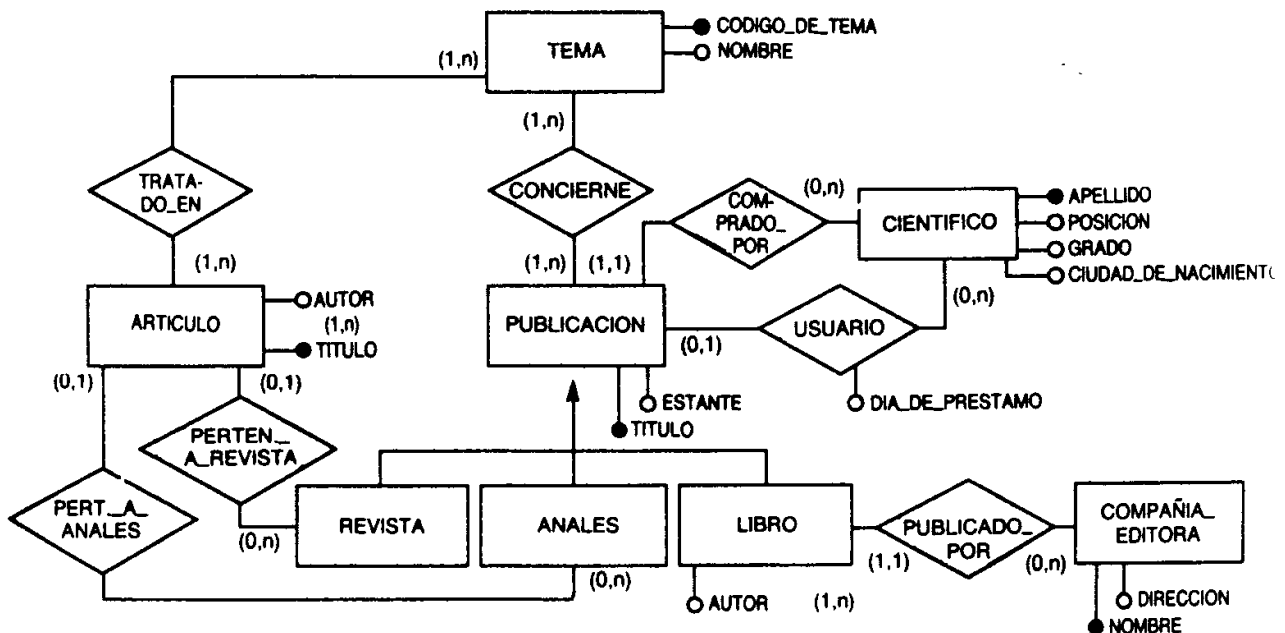
Nombre	Descripción
AUTOR	Los autores de las publicaciones de interés para los científicos.
PUBLICACIÓN	Las publicaciones guardadas por los científicos en sus estantes privados; los científicos las obtienen, por lo general, directamente de los autores.
TEMA	Las áreas de investigación de interés para los autores.
SOLICITADO_POR	Conecta los artículos solicitados por algún científico, con el autor a quien ha sido hecha la solicitud.
ENVIADO_POR	Conecta los artículos que han sido enviados por los autores, con los científicos que los han solicitado.

Tabla 2. Conceptos del esquema 2 (esquema del bibliotecario)

Nombre	Descripción
PUBLICACION	Las publicaciones que actualmente se conservan en la biblioteca.
ARTICULO	Los artículos publicados en revistas o anales guardados en la biblioteca.
TEMA	Temas de los artículos.
COMPRADO_POR	Indica qué científico es responsable de la subvención usada para adquirir la publicación.



Esquema 1. Esquema del científico.



Esquema 2. Esquema del bibliotecario

Durante esta tarea, es posible descubrir conflictos de varios tipos: conflictos de nombres o conflictos estructurales.

Conflictos de nombres

Existen dos tipos de conflictos de nombres: los sinónimos y los homónimos.

Un **sinónimo** ocurre cuando el mismo objeto (concepto) del dominio de aplicación se representa con diferente nombre en cada esquema; un **homónimo** ocurre cuando dos esquemas usan el mismo nombre para describir diferentes objetos del dominio de aplicación.

Para descubrir sinónimos y homónimos se puede utilizar la noción de *semejanza* y de *discrepancia* entre conceptos.

La **semejanza de conceptos** surge cuando **conceptos con nombres diferentes** poseen **algunas propiedades y restricciones comunes** en los esquemas. La semejanza de dos conceptos indica que pueden ser **sinónimos**.

La **discrepancia de conceptos** surge cuando **conceptos con el mismo nombre** poseen **propiedades y restricciones distintas** en cada esquema. La discrepancia entre dos conceptos indica que pueden ser **homónimos**.

Los términos **propiedad** y **restricción**, en estas definiciones significan lo siguiente: las propiedades de un concepto son todos los demás *conceptos vecinos* en el esquema; las restricciones son reglas o condiciones sobre conceptos del esquema.

Tabla 3. Propiedades y restricciones

Elemento del Esquema	Propiedades (conceptos vecinos)	Restricciones
Entidad	Sus atributos, interrelaciones adyacentes y jerarquías de generalización	Cardinalidades mínima y máxima de las interrelaciones donde la entidad participa; identificadores
Interrelación	Sus atributos; entidades participantes	Cardinalidades mínima y máxima de las entidades participantes
Atributo	Entidades o interrelaciones a las que pertenece	Cardinalidades mínima y máxima; conjunto de valores; identificadores que incluyen el atributo

Sin embargo, el que exista semejanza (o discrepancia) entre conceptos no siempre significa que sean sinónimos (u homónimos). Por ejemplo, considérese la entidad CIENTIFICO que en el esquema 1 desempeña el papel de interesarse y conseguir artículos y libros; en cambio en el esquema 2, participa en la compra de publicaciones. A pesar de la discrepancia, las dos entidades sí corresponden al mismo concepto de la realidad (es decir, son discrepantes pero NO son homónimos).

Conflictos estructurales

Una vez analizados y solucionados los conflictos de nombres (véase el apartado siguiente), se consigue la *unificación de los nombres*, de forma que en este punto se supone que dos conceptos (atributos, entidades o interrelaciones) con igual nombre, representan el mismo concepto de la realidad.

Durante el análisis de conflictos estructurales, se comparan los conceptos que tienen el mismo nombre en los esquemas de entrada, para determinar si pueden fusionarse.

Se pueden dar las siguientes posibilidades:

1. *Conceptos idénticos*: con exactamente la misma estructura de representación y propiedades vecinas.
2. *Conceptos compatibles*: poseen diferentes estructuras de representación o propiedades vecinas, pero no contradictorias. Un ejemplo de compatibilidad de conceptos sería el uso de una entidad y un atributo para representar el mismo concepto (entidad TEMA en el esquema 2 y atributo TEMA en el esquema 1). Otro sería el uso de un atributo con diferente dominio en cada esquema.
3. *Conceptos incompatibles*: poseen restricciones contradictorias entre sí. En este caso, es necesario eliminar las fuentes de incompatibilidad antes de fusionar los esquemas. En el modelo EER, algunas posibles incompatibilidades son las siguientes:
 - a) *Cardinalidades diferentes* para el mismo atributo o entidad,
 - b) *Identificadores (claves) diferentes*: un atributo identificador en un esquema no lo es en el otro,
 - c) *Jerarquías de generalización/especialización inversas*: el tipo de entidad A es subtipo del tipo de entidad B en un esquema, mientras que en el otro esquema B es subtipo de A.

2. Resolución de conflictos

Resolución de conflictos de nombres

Resolver un conflicto de nombres supone cambiar el nombre de algún concepto, y posiblemente también añadir alguna *propiedad interesquemática*.

Se cambia el nombre de un concepto siempre que se detecta un sinónimo o un homónimo. Los sinónimos deben eliminarse para evitar la ambigüedad. Por ejemplo, cuando dos conceptos como USUARIO y CLIENTE son sinónimos, se selecciona uno de ellos, por ejemplo USUARIO, y se cambia el nombre de CLIENTE a USUARIO.

En el caso de un homónimo, supóngase que REGISTRO, en un esquema se refiere al proceso de registrar a una persona que alquila un coche, mientras que en el otro esquema significa hacer una reserva de un coche. En este caso, el nombre del concepto del segundo esquema se debe cambiar a RESERVA.

Las propiedades interesquemáticas expresan restricciones mutuas entre los conceptos semejantes (o discrepantes) que aparecen en diferentes esquemas. Por ejemplo, CANDIDATO_A_ANALISTA en una vista, semejante a ESTUDIANTE en otra vista, puede estar limitado a ser un subconjunto de éste último. Estas propiedades deben anotarse como *extensiones* de los dos esquemas; y posteriormente se utilizarán en la integración de los esquemas.

En los esquemas del ejemplo existen ejemplos de semejanza y de discrepancia entre conceptos.

1. Ejemplos de discrepancia de conceptos:

TEMA (Esq. 1): interés de un autor → AREA_DE_INVESTIGACION

TEMA (Esq. 2): lo que tratan las publicaciones

PUBLICACION (Esq. 1): objeto solicitado a los autores y enviado por éstos..... → ARTICULO

Interrelación PUBL_DE_INTERES..... → ARTICULO_DE_INTERES

PUBLICACION (Esq. 2): objeto comprado (o solicitado en préstamo) por un científico

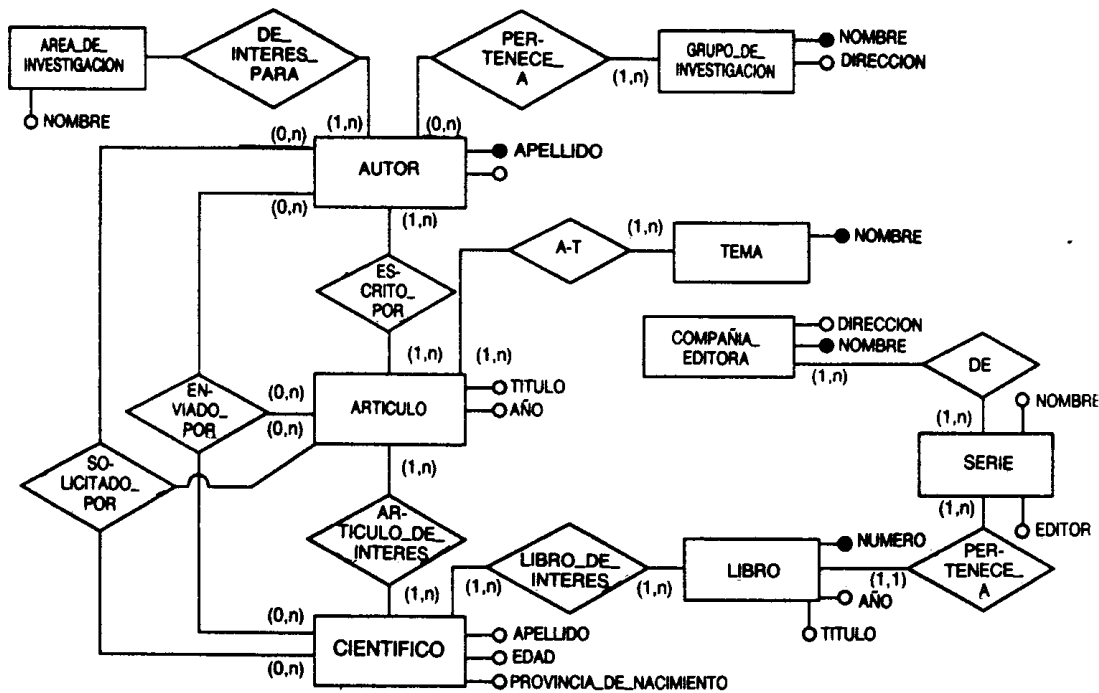
3. Ejemplo de semejanza entre conceptos:

El atributo PALABRA_CLAVE de la entidad PUBLICACION en el primer esquema, y la entidad TEMA en la interrelación CONCIERNE con PUBLICACION en el segundo esquema, tienen las mismas cardinalidades mínima y máxima. Este conflicto se resuelve cambiando el nombre del atributo PALABRA_CLAVE por el de TEMA en el primer esquema (una vez hecho el cambio del ejemplo 1).

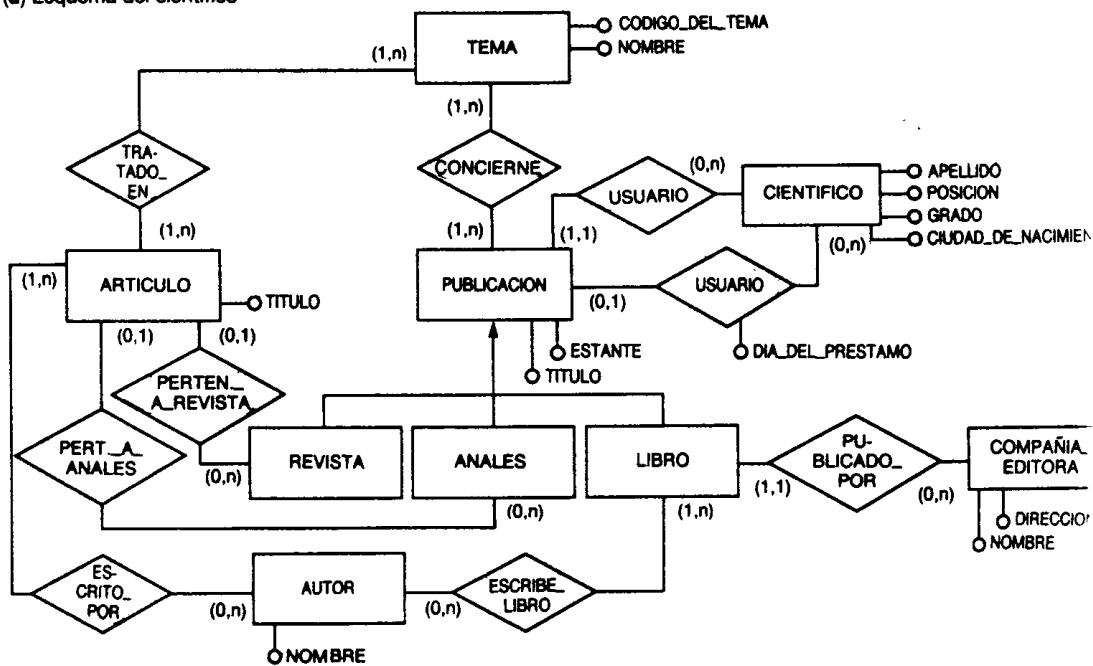
Resolución de conflictos estructurales

1. Los conceptos idénticos no suponen ningún conflicto.
2. Los conflictos de compatibilidad se resuelven con facilidad, alterando una de las dos representaciones para igualarla a la otra.
3. Las posibles soluciones para los conflictos de incompatibilidad incluyen la selección de una representación en vez de otra, o bien la construcción de una representación común tal que todas las restricciones de los dos esquemas estén incluidas en el esquema integrado.

A continuación se muestran los esquemas obtenidos una vez resueltos los conflictos de nombres, así como los estructurales indicados después:



(a) Esquema del científico



(b) Esquema del bibliotecario

Resolución de conflictos estructurales en el ejemplo:

1. CIENTIFICO es una entidad en los dos esquemas que, aunque posee diferentes atributos e interrelaciones, representa el mismo objeto y no necesita ninguna actividad de reestructuración.
2. AUTOR y TEMA poseen diferentes estructuras de representación en los dos esquemas. AUTOR debe transformarse en una entidad en el esquema 2 y TEMA debe transformarse en una entidad en el esquema 1. En ambos casos se introducen las interrelaciones adecuadas.
3. Surgen dos casos de incompatibilidad de conceptos. El primero se refiere a la cardinalidad mínima de la entidad COMPAÑIA_EDITORA en la interrelación PUBLICADO_POR con LIBRO: es 0 en el esquema 2, pero es 1 (a través de la entidad SERIE) en el esquema 1. Es decir, la primera representación incluye compañías editoras que no han publicado ninguno de los libros que están actualmente en la biblioteca, mientras que en la segunda se excluye a estas compañías editoras. Se elige la primera alternativa porque es menos restrictiva.
La segunda incompatibilidad se refiere al concepto AUTOR. El esquema 1 incluye cualquier autor (como entidad), mientras haya un científico que se interese por sus publicaciones (aunque éstas no estén en la biblioteca). El esquema 2 incluye sólo los autores (como atributos de ARTICULO) que tienen al menos un artículo disponible actualmente en la biblioteca. De nuevo, se elige la primera alternativa, por ser menos restrictiva

4. Fusión (combinación) de vistas (esquemas)

La *fusión de vistas* actúa sobre los esquemas de entrada y produce un único esquema que incluye todos los conceptos representados en los esquemas de entrada.

En esta etapa del diseño ya se han resuelto todos los conflictos, así que suele consistir en una simple “superposición” de conceptos comunes.

- Las entidades que coinciden por completo se superponen directamente. Por ejemplo COMPAÑIA_EDITORA.
- Las entidades correspondientes al mismo concepto, pero que poseen diferentes atributos, se superponen para dar una entidad cuyo conjunto de atributos es la unión de atributos; el conjunto de identificadores (claves) es la unión de identificadores.
- Si las jerarquías de generalización/especialización se han representado como propiedades interesquemáticas, se añaden al esquema.
- Cuando las entidades están completamente superpuestas, en el esquema resultante aparecen todas las interrelaciones y generalizaciones de los esquemas de entrada.
- Las propiedades interesquemáticas pueden llevar a reestructurar, añadir o suprimir conceptos en el esquema resultante. Un ejemplo es el de las entidades TEMA Y AREA_DE_INVESTIGACION en la aplicación del bibliotecario: un área de investigación puede dividirse en temas, que representan una subdivisión más fina de las disciplinas. Así, se puede añadir un enlace lógico entre las entidades AREA_DE_INVESTIGACION y TEMA, introduciendo una interrelación 1:N entre ellas.

El esquema conceptual global obtenido se muestra en la figura de la página siguiente.

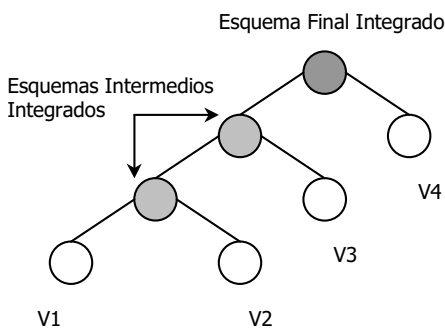
5. Reestructuración del Esquema Conceptual global o integrado.

Este es un paso opcional, en el que el esquema global se analiza y reestructura para eliminar redundancias o complejidades innecesarias.

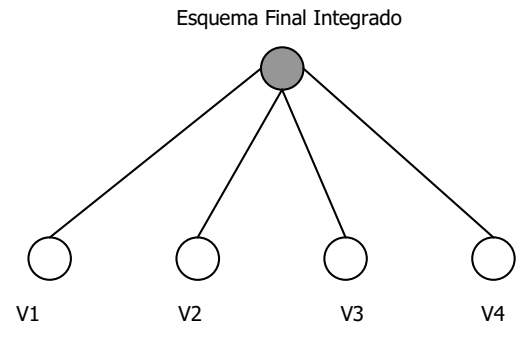
Hasta ahora hemos considerado la existencia de **dos** esquemas, sin embargo, suelen ser algunos más los esquemas que han de combinarse para obtener el esquema global.

Algunas **estrategias** para el proceso de **integración de vistas** que se han propuesto son las siguientes:

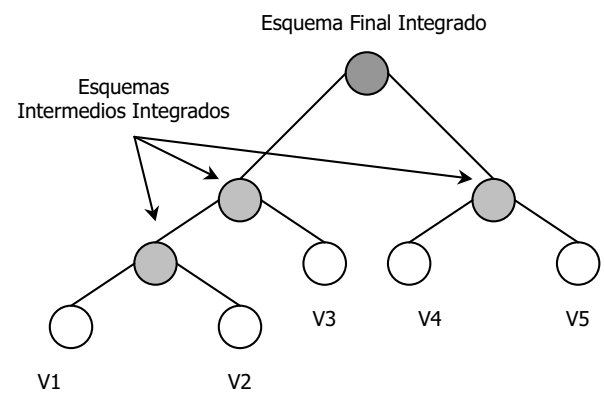
1. **Integración de escalera binaria:** primero se combinan dos esquemas que sean muy similares. El esquema resultante se integra con otro esquema y así sucesivamente hasta que todos estén integrados. El orden de los esquemas para la integración puede basarse en alguna medida de la similitud de los esquemas que establezca el diseñador. Esta es la estrategia adecuada para la integración manual, ya que se realiza paso a paso.
2. **Integración n-aria:** todas las vistas se combinan a la vez, después de analizar y especificar sus correspondencias. Esta estrategia necesita herramientas software y se aplica para problemas de diseño grandes.
3. **Integración binaria balanceada:** primero se integran pares de esquemas; luego se combinan dos a dos los esquemas resultantes y así sucesivamente hasta obtener el esquema global final.
4. **Integración mixta:** los esquemas se dividen en grupos según su similitud, y cada grupo se integra por separado; los esquemas resultantes se agrupan de nuevo, se combinan los componentes de cada grupo y así sucesivamente.



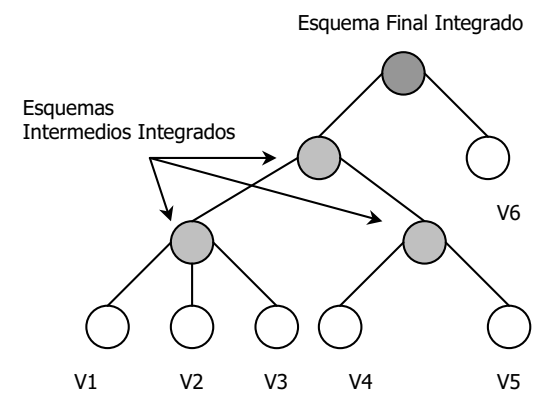
Integración de Escalera Binaria



Integración n-aria



Integración Binaria Balanceada



Integración Mixta