

2. Diagramas Entidad-Relación

Los diagramas Entidad-Relación son un estándar actual en el diseño de bases de datos. De hecho, se trata de una herramienta sin la que, posiblemente, las bases de datos tal y como las entendemos actualmente no existirían.

Los diagramas Entidad-Relación son herramientas gráficas clave en el diseño de bases de datos. Su confección debe ser sistemática y rigurosa si se desea obtener un sistema válido y eficiente, ya que será a partir de estos diagramas que se desarrollará toda la implementación de bases de datos en los sistemas gestores de bases de datos concretos que corresponda a cada empresa u organismo.

2.1 Diseño de bases de datos

El diseño de BD se estructura, fundamentalmente, en tres grandes etapas:

- Diseño conceptual
- Diseño lógico
- Diseño físico

Aunque nos centramos en el estudio y en la práctica del diseño conceptual, no deben perderse de vista las otras fases del diseño, que también son importantes, para obtener una visión de conjunto de todos estos procesos.

Conocer cómo se estructura el modelo Entidad-Relación es importante. Y también lo son cuestiones que deben permitirnos aprovechar la tecnología proporcionada por las BD y los correspondientes sistemas gestores, como los siguientes:

- Qué entidades debe incluir una BD determinada.
- Qué interrelaciones deben considerarse.
- Qué atributos deben existir y en qué entidades o interrelaciones deben incorporarse.
- Qué claves primarias ya se pueden establecer en la fase de diseño conceptual.

2.1.1 Fases del diseño de BD

Diseñar BD no es tarea sencilla. Aunque la porción del mundo real que se quiera modelizar en un caso concreto sea relativamente pequeña, las estructuras de datos resultantes pueden llegar a tener un cierto grado de complejidad. Sin embargo, a medida que aumenta la información a considerar, y su complejidad, el modelo de datos necesario para representarla puede convertirse, ciertamente, en una construcción complicada.

Querer resolver de repente toda la problemática que puede comportar la modelización de una BD no es, pues, una opción muy realista. Al afrontar una tarea de esta envergadura, es preferible dividirla en subtareas para simplificarla.

Así pues, resulta conveniente descomponer el diseño de BD en diferentes etapas, de modo que en cada una de ellas sólo se examinen ciertos aspectos, o tipos de problemas, para minimizar la posibilidad de error. Entonces, a partir del resultado obtenido en cada fase, se puede continuar trabajando en la siguiente fase, hasta llegar al resultado esperado, al final de la última fase.

Es habitual estructurar el diseño de BD en las tres etapas o fases siguientes:

1. Diseño conceptual.
2. Diseño lógico.
3. Diseño físico.

Fase de diseño conceptual

Lo primero que debe hacerse, durante la fase de diseño conceptual, es recopilar toda la información necesaria de la parte del mundo real que nos proponemos modelizar con una BD.

Esta recopilación de información se realizará por diferentes vías, tales como:

- Entrevistas con los futuros usuarios de la BD que está diseñando.
- Examen de la documentación proporcionada por estos mismos usuarios.
- Observación directa de los procesos a informatizar.

A continuación, deben estructurarse convenientemente los datos necesarios para dar respuesta a todas las necesidades derivadas del conjunto de informaciones compendiadas.

El **objetivo del diseño conceptual** consiste en obtener una especificación sistemática.

Dicha especificación sistemática resultado del diseño conceptual debe cumplir dos tipos de requisitos:

- **De datos.** El modelo resultante debe tener en cuenta la estructura completa de los datos y su integridad.
- **Funcionales.** Un buen esquema conceptual también deberá prever las necesidades básicas en materia de manipulación de datos (es decir, las operaciones de inserción, borrado, consulta y modificación de éstas). Durante las fases posteriores, puede ser conveniente depurar el diseño a fin de optimizar las operaciones a realizar sobre los datos.

Por último, es necesario elegir un modelo de datos de alto nivel y traducir los requisitos anteriores a un esquema conceptual de la futura BD expresado con los conceptos y la notación correspondientes. Uno de los modelos de datos de alto nivel más utilizados es el **modelo entidad-interrelación**.

Expresado en la terminología del modelo ER, el esquema de datos desarrollado durante la fase de diseño conceptual debe especificar todas las entidades necesarias, y las interrelaciones entre ellas, con las cardinalidades adecuadas, así como los atributos que correspondan en cada caso.

El modelo ER

Una **entidad**, en el modelo ER, es una abstracción que nos interesa modelizar, y mediante la cual se agrupan las instancias del mundo real que poseen unas características comunes.

Una **interrelación**, en el modelo ER, es la asociación entre instancias de diferentes entidades tipos. Esta asociación puede darse con diversas cardinalidades.

Un atributo en el modelo ER es una característica de una entidad que nos interesa tener registrada.

El modelo resultante debe revisarse para garantizar la satisfacción de todas las necesidades detectadas, por un lado, y para evitar redundancias de los datos (es decir, repeticiones indeseadas de éstos), por otro.

El resultado de la fase de diseño conceptual pertenece al llamado **mundo de las concepciones**, pero aún no al **mundo de las representaciones**, ya que no se especifica ninguna representación informática concreta.

Como puede verse, durante la fase de diseño conceptual no es necesario tener en cuenta, aún, ni el tipo de BD que se utilizará posteriormente ni, mucho menos, el SGBD o el lenguaje concreto con el que se implementará.

Fase de diseño lógico

En la fase de **diseño lógico**, se trabaja con el modelo abstracto de datos obtenido al final de la etapa de diseño conceptual, para traducirlo al modelo de datos utilizado por el sistema gestor de bases de datos (SGBD) con el que se desea implementar y mantener la BD.

Por tanto, a partir de esta fase de diseño, sí que hay que tener en cuenta la tecnología concreta que debe emplearse en la creación de la BD, ya que la BD resultante puede adecuarse a diferentes modelos lógicos, como por ejemplo los siguientes:

- Jerárquico
- Relacional
- Distribuido
- Orientado a objetos

Pese a la diversidad de posibilidades, lo cierto es que lo más frecuente, a la hora de diseñar una BD, todavía consiste en expresar el esquema conceptual en un modelo ER y, a continuación, traducirlo a un modelo relacional.

Cuando el producto de una fase de diseño lógico es una BD relacional, ésta consiste en un conjunto de relaciones (de lo contrario, llamadas *representaciones tabulares*) compuestas por atributos, algunos de los cuales forman parte de claves primarias o de claves foráneas.

Resulta evidente, pues, que el resultado de la fase de diseño lógico se sitúa ya dentro del llamado **mundo de las representaciones informáticas**.

Las **claves primarias** sirven para distinguir entre sí las distintas tuplas de atributos dentro de una misma relación.

Las **claves foráneas** son unos instrumentos destinados a permitir la interrelación de la respectiva relación con otros.

Fase de diseño físico

El **diseño físico** consiste en realizar ciertos tipos de modificaciones sobre el esquema lógico obtenido en la fase anterior de diseño lógico, con el fin de incrementar la eficiencia.

La eficiencia de un esquema puede comportar la modificación de algunas operaciones que deban realizarse con los datos, aunque comporten un cierto grado de redundancia de éstos, como, por ejemplo:

- Añadir algún atributo calculable a alguna relación.
- Dividir una relación en otras dos o más.
- Incluir en la BD una relación que sea el producto de combinar dos o más relaciones.

Pero la fase de diseño físico también se caracteriza por la posibilidad de adoptar otras decisiones, relacionadas con aspectos de implementación física a menor nivel, y estrechamente vinculadas con el SGBD con el que se trabaja en cada caso, como los siguientes:

- Definición de índices.
- Asignación del espacio inicial para las tablas, y previsión de su ulterior crecimiento.
- Selección del tamaño de las memorias intermedias.
- Parametrización del SGBD según las opciones que éste ofrezca.

Para tomar acertadamente este tipo de decisiones, es necesario tener en cuenta las características de los procesos que operan con los datos, la frecuencia de ejecución de los diferentes tipos de consulta, el grado de volatilidad de los datos, los volúmenes de información a almacenar, etc.

2.1.2 Diseño conceptual de una BD.

Podemos considerar a cualquier empresa, organización, institución, etc. como un sistema con reglas propias de funcionamiento. Este sistema, susceptible de ser informatizado, está compuesto por tres subsistemas:

- **Subsistema de producción.** Se encarga de realizar las actividades propias de la organización de que se trate en cada caso (por ejemplo, fabricar coches, repararlos, venderlos, etc.).

- **Subsistema de decisión.** Se encarga de dirigir, coordinar y planificar las actividades realizadas en el ámbito del subsistema de producción.
- **Subsistema de información.** Se encarga de recoger, almacenar, procesar y distribuir todas las informaciones necesarias para el buen funcionamiento de los otros dos subsistemas.

Las BD sirven para almacenar las representaciones de las informaciones utilizadas en el ámbito de los sistemas de información.

Los elementos que conforman un **sistema de información** son de dos tipos:

- Datos: representaciones de las informaciones.
- Procesos: acciones ejercidas sobre los datos (consultas, modificaciones, cálculos, etc.).

En función de las observaciones anteriores, podemos afirmar que existen dos premisas que todo diseñador de BD debería tener bien asumidas antes de empezar a trabajar en cualquier proyecto:

- No es competencia del diseñador de BD, como tal, tomar decisiones sobre la porción del mundo real que desea modelizar.
- En principio, el diseñador de BD tampoco debe inventarse características de la realidad a modelizar: simplemente debe reflejarlas de la manera más fiel posible en el modelo resultante.

Podemos subdividir esta etapa de diseño conceptual en dos fases sucesivas, que conllevan tareas de diferentes tipos: la **recogida y abstracción de las necesidades** de la organización, por un lado, y la **elaboración de un esquema conceptual** mediante un modelo de datos concreto, por otro.

La **recogida y abstracción de las necesidades** de la organización se lleva a cabo mediante diferentes procedimientos (entrevistas, examen de documentación, etc.), y en esta fase es necesario recoger toda la información necesaria para cubrir todos los requerimientos de datos. Pero, con esta información, debe seguirse un proceso de abstracción que nos permita estructurarla y diferenciar entre las cuestiones esenciales y las accesorias.

La **elaboración de un esquema conceptual** mediante un modelo de datos concreto comporta que toda la información recogida y debidamente estructurada debe expresarse en una notación estandarizada (como los diagramas ER).

2.1.3 Captura y abstracción de los requerimientos de datos

Para realizar la captura y abstracción de los requerimientos, en primer lugar, es necesario averiguar qué necesidades tienen los usuarios de la futura BD. A menudo, estos potenciales usuarios sólo tienen una percepción muy general de lo que necesitan.

Usuarios de las BD: Lo son tanto los operadores que interactúan habitualmente con el sistema, como los destinatarios finales de la información que debe extraerse o incluirse en la BD.

El diseñador debe saber seleccionar las cuestiones esenciales, diferenciarlas de los aspectos accesorios, y descubrir cuáles son los verdaderos intereses de quienes han encargado el diseño de la BD.

Ahora bien, el diseñador de BD, como tal, no debe decidir nada. Su tarea consiste más bien en ayudar a los usuarios potenciales a descubrir qué necesitan exactamente. Es un trabajo pesado y complicado, que sirve para descubrir el verdadero flujo de datos que debe reflejarse en el modelo conceptual resultante. Normalmente, esta función comporta lo siguiente:

- Muchas entrevistas con los futuros usuarios de todos los niveles y secciones de la organización a informatizar. Es necesario anotar todos los detalles surgidos durante las entrevistas, susceptibles de implementación informática.
- Examen exhaustivo de la documentación proporcionada por el cliente, a fin de conocer el funcionamiento interno de la organización.
- Observaciones directas de los distintos procesos de la organización.

Un buen diseñador de BD debe llegar, al menos, a una solución plausible para cada problema que se le haya planteado. En ocasiones, también podrá tomar en consideración soluciones parciales alternativas. Y, en todo caso, finalmente, debe presentar un análisis completo de los requerimientos en el que concrete las especificaciones de la organización que le ha encargado el proyecto de diseño.

Un ejemplo concreto: BD de una Red de Bibliotecas de INS

Examinar los requisitos de diseño que nos permitan establecer un modelo conceptual para una futura BD que dé servicio conjuntamente a las bibliotecas de todos los institutos de Galicia (mediante una aplicación web ulterior, por ejemplo), a fin de poner al alcance de cualquier miembro de la comunidad educativa de secundaria la totalidad de los fondos bibliográficos y de otros recursos audiovisuales, aunque no se encuentren físicamente en el centro docente donde trabaje o estudie el usuario que solicite el préstamo, nos proveerá de un buen escenario para desarrollar un ejemplo de diseño de BD concreto.

No se trata ahora tanto de profundizar en cuestiones de detalle, como de captar globalmente cómo se trabaja durante la fase de diseño conceptual, y aplicar los conocimientos adquiridos previamente -los cuales son necesarios para elaborar los esquemas conceptuales-, en el nuestro caso fundamentalmente con ayuda del modelo de datos ER.

Después de entrevistarnos con el alto cargo del Departamento de Educación, que promueve la coordinación y colaboración de las bibliotecas de todos los institutos (INS), con varios directores y directoras de diferentes bibliotecas interesadas en adherirse a la futura Red de Bibliotecas de INS de Galicia, y también con varios bibliotecarios de éstas, y de recopilar y examinar diferente documentación que nos ha sido proporcionada, los diseñadores de BD encargados del proyecto hemos sintetizado las informaciones recibidas de la siguiente forma:

1. Servicio de préstamo

- Mediante este servicio, se ofrece la posibilidad de sacar los distintos documentos (libros, revistas, vídeos, CD...) de la biblioteca.
- Para utilizar el servicio de préstamo, los usuarios deben tener el carné de biblioteca.
- Este carné sirve para utilizar otros servicios de la biblioteca, como el servicio de acceso a Internet, el servicio Wi-Fi, consulta en BD o participación en algunas actividades que organicen las bibliotecas.
- El mismo carné sirve para toda la Red de Bibliotecas y permite disfrutar de interesantes ventajas en el mundo de la cultura.
- La solicitud del carné de biblioteca debe realizarse desde la misma biblioteca o bien online. En el formulario de solicitud, se piden al usuario los datos personales necesarios (nombre y apellidos, dirección, fecha de nacimiento, etc.).
- Además, los usuarios pueden realizar otras gestiones online relacionadas con el préstamo: reservas y renovaciones de documentos.

2. Normativa de uso del servicio de préstamo de la Red de Bibliotecas

- Tener el carné implica aceptar las normas de funcionamiento de la biblioteca.
- El carné es personal e intransferible.
- Los usuarios menores de dieciséis años necesitan la autorización de los padres para realizarse el carné, así como los mayores de dieciséis años que no tengan DNI.
- Es necesario comunicar, en la biblioteca, cualquier cambio de domicilio o la pérdida del carné.
- Cada lector puede llevarse en préstamo hasta diez documentos entre libros, revistas y audiovisuales, por un plazo de tres semanas, los libros, y una semana, las revistas y audiovisuales.
- Pasado el plazo, cualquier documento en préstamo puede ser renovado, siempre que ningún otro usuario lo haya reservado. La renovación se puede realizar en la biblioteca, por teléfono, por correo electrónico o por Internet.
- Es necesario cuidar los documentos dejados en préstamo. Si un usuario no devuelve los documentos en el plazo fijado, puede ser excluido del servicio de préstamo de las bibliotecas de la Red durante un tiempo equivalente al retrasado en la devolución. Si un usuario pierde o daña un documento, debe notificarlo a la biblioteca y debe comprar el mismo documento o abonar su importe.

- Quedan excluidos de préstamo los siguientes documentos:
 - Algunas obras de referencia: enciclopedias, diccionarios, atlas, etc.
 - Fondo de reserva.
 - Documentos pertenecientes a la colección local.
 - Números corrientes de revistas y diarios.
 - Los documentos que la biblioteca crea conveniente que no salgan de la biblioteca.
- Además del servicio de préstamo al usuario individual, las bibliotecas ofrecen otros tipos de préstamo:
 - Servicio de préstamo interbibliotecario. Mediante el servicio de préstamo interbibliotecario, las bibliotecas de la Red se encargan de localizar y proporcionar los documentos que carece del fondo propio y que están disponibles en otras bibliotecas de la Red. El préstamo entre las bibliotecas de la Red tiene un precio público establecido de 1,20€.
 - Servicio de préstamo a domicilio. Este servicio se dirige a personas con problemas de movilidad temporal o permanente, como enfermos crónicos, personas en período de convalecencia, con discapacidades físicas o personas mayores. No todas las bibliotecas de la Red ofrecen ese servicio. El servicio de préstamo a domicilio tiene un precio público establecido de 1,50€.

3. Consulta del catálogo

- Se debe poder consultar la disponibilidad de los fondos bibliotecarios por los siguientes conceptos:
 - Firma (identificador del ejemplar físico objeto de préstamo)
 - Título
 - Palabra clave contenida en el título
 - Autor (concretamente, por el apellido)
 - Nombre o apellidos incompletos del autor
 - Materia
- Además, las búsquedas deben poder limitarse a los siguientes ámbitos:
 - Una biblioteca concreta
 - Todas las bibliotecas de una sola población
 - Todas las bibliotecas de una sola comarca
- Finalmente, deben poder limitarse los resultados obtenidos en función de los siguientes conceptos:
 - Idioma
 - Tipo de formato
 - Año de publicación

2.1.4 Identificación de entidades

La especificación de los requerimientos de datos sirve como punto de partida para la elaboración del esquema conceptual de la futura BD. A continuación, lo primero que debe hacerse es identificar a las entidades y sus atributos.

La identificación de las entidades y de sus atributos se puede documentar con un listado que siga el siguiente formato:

ENTIDAD1 (Atributo1, Atributo2, ...)

ENTIDAD2 (Atributo1, Atributo2, ...)

....

ENTIDAD (Atributo1, Atributo2, ...)

En el listado, los nombres de las entidades irán con mayúsculas, y los nombres de los atributos empezarán con mayúsculas.

Siguiendo con el ejemplo BD de las bibliotecas, en una primera aproximación, repasando las especificaciones que tenemos, podríamos obtener una solución que contara como mínimo con las tres entidades siguientes, que incorporan los respectivos atributos expresados entre paréntesis:

- IES (Población, Comarca). Posibles atributos adicionales: Nombre, Dirección y Teléfono.
- DOCUMENTO (Formato, Importe, Excluyente Préstamo, Firma, Título, Autor, Materia, Idioma, AñoPublicacion).
- USUARIO (DNI, Nombre, Apellidos, Dirección, FechaNacimiento, Carnet, FechaFinalExclusioPretec).

Pero estamos en condiciones de depurar ese resultado rudimentario. Fijémonos en que habrá muchas poblaciones y comarcas que se repetirán para diferentes INS. Sería mejor, pues, que estos dos atributos constituyeran dos entidades independientes, convenientemente interrelacionadas con la entidad INS: POBLACION y COMARCA.

Otra ventaja de utilizar estas dos entidades en vez de considerarlas atributos de una entidad concreta es que, en caso necesario, las podremos reaprovechar interrelacionándolas con otras entidades.

Lo mismo ocurre con otros atributos de DOCUMENTO: el formato, el autor la materia y el idioma.

Sería mejor tener una entidad para grabar los tipos de formato, que siempre serán los mismos (libro, revista, CD, etc.), nombrada por ejemplo FORMATO, e interrelacionarla con DOCUMENTO.

Estaremos frente al mismo fenómeno con los idiomas (catalán, castellano, inglés, etc.) y las materias (historia, música, física, etc.). Por tanto, será conveniente contar con dos entidades más: IDIOMA y MATERÍA.

Por otra parte, los autores podrán ser autores de una pluralidad de documentos y, a fin de simplificar las búsquedas ulteriores por autor, es preferible destinar una entidad propia para representarlos (podemos llamarla sencillamente AUTOR).

Después de hacer estas consideraciones, nos encontraríamos con unas cuantas entidades más que inicialmente:

- INS (Nombre, Dirección, Teléfono)
- DOCUMENTO (Firma, Título, AñoPublicacion, Importe, ExcluidoPréstamo)
- USUARIO (Carnet, DNI, Nombre, Apellidos, Dirección, FechaNacimiento, FechaFinalExclusioPretec)
- POBLACION(Nombre)
- COMARCA(Nombre)
- FORMATO(Descripción)
- AUTOR (Nombre, Apellidos)
- MATERIA(Descripción)
- IDIOMA(Descripción)

Debemos ser conscientes de que cualquier especificación que no se desprenda directamente de los documentos de trabajo inicial en los que hemos sintetizado los diferentes requerimientos de datos detectados como, por ejemplo, añadir los nuevos atributos a INS debe validarse mediante nuevas entrevistas, examinando nueva documentación, revisando la antigua, o, en su defecto, observando *in situ* el proceso en ejecución que pueda comportar una innovación en la especificación.

En cambio, otros aspectos que pueden parecer más atrevidos (como representar lo que era inicialmente el atributo de una entidad como otra entidad independiente) no deben implicar necesariamente unos procesos de validación como los que acabamos de describir, si se tratan puramente de decisiones técnicas de diseño.

Cabe destacar que, por poco complicada que sea la porción del mundo real a modelizar, normalmente es posible obtener varios modelos conceptuales, en algunos casos alternativos y equivalentes, y en otros orientados a satisfacer en mayor o menor medida diferentes finalidades, pero con resultados igualmente correctos.

2.1.5 Designación de interrelaciones

Tras la captura y abstracción de los requerimientos de datos y de la identificación de entidades, el siguiente paso consiste en establecer las interrelaciones necesarias entre las entidades detectadas.

Las interrelaciones se pueden documentar con el siguiente formato:

Interrelación (Atributo1, Atributo2, ...), entre ENTIDAD y ENTIDAD

En que el nombre de la interrelación empezará con mayúsculas y, en caso de disponer de atributos, éstos irán entre paréntesis y con la inicial del nombre también con mayúsculas. Habrá que especificar el nombre de las entidades que interrelaciona.

En algún caso, al igual que las entidades, las interrelaciones pueden incorporar algún atributo. Inicialmente podemos encontrar, como mínimo, cuatro interrelaciones que sólo afectan a las tres entidades originarias:

- Prestamo (Fecha, Tipo, Precio), entre USUARIO y DOCUMENTO
- RenovacionPrest (Fecha), entre USUARIO y DOCUMENTO
- Reserva (Fecha), entre USUARIO y DOCUMENTO
- Ubicación, entre INS y DOCUMENTO

Pero, después de considerar el establecimiento del resto de entidades, deberemos añadir las interrelaciones correspondientes:

- Está, entre INS y POBLACION
- Pertenece, entre POBLACION y COMARCA
- Tiene1, entre FORMATO y DOCUMENTO
- Tiene2, entre AUTOR y DOCUMENTO
- Versa, entre MATERÍA y DOCUMENTO
- Expresado, entre IDIOMA y DOCUMENTO

Notemos que cuando se repite un mismo nombre en más de una interrelación (como aquí ocurre con Tiene), es necesario numerarlas correlativamente, para evitar confusiones al hacer referencia a cada una.

Y, además, podríamos considerar el establecimiento de una interrelación entre USUARIO y POBLACION para completar su dirección debidamente (como hemos hecho con los INS). Quedaría así:

- Vive, entre USUARIO y POBLACION

2.1.6 Establecimiento de claves

Recordemos que la **clave primaria** de una entidad está constituida por un atributo, o por un conjunto de atributos, cuyos valores son capaces de identificar unívocamente las instancias de aquélla.

En ocasiones, una entidad dispone de más de un atributo, o conjunto de atributos, que están en condiciones de constituir la clave primaria. En este caso, hablamos de **claves candidatas**.

Y una vez que el diseñador elige un atributo o conjunto de atributos concretos para identificar unívocamente las instancias, entre los que cumplen las condiciones (**claves candidatas**), éste o estos atributos pasan a ser la **clave primaria** , y los no seleccionados permanecen como **claves alternativas** .

Un criterio importante a la hora de decidirse para que uno o más atributos formen la clave primaria de una entidad, es que su valor nunca cambie o, al menos, muy raramente.

El atributo o conjunto de atributos que forman la clave primaria deben aparecer subrayados en la documentación.

Por ejemplo, no sería muy prudente decidirse por el número de teléfono móvil como clave primaria de una entidad que almacena a personas, porque, a pesar de ser un número habitualmente de uso personal, puede cambiar a lo largo del tiempo (por ejemplo, una persona puede regalar su móvil a otra). En cambio, el número de DNI puede ser, por lo general, una buena opción, ya que, en principio, este número es personal e intransferible.

Continuando con nuestro ejemplo, estamos en condiciones de encontrar las siguientes claves, formadas por sólo un atributo, las cuales se muestran subrayadas:

- DOCUMENTO (Firma, Título, AñoPublicación, Importe, ExcluidoPrestamo)
- COMARCA (Nombre)
- POBLACION (Nombre)
- FORMATO (Descripcion)
- MATERIA (Descripcion)
- IDIOMA (Descripcion)

La entidad AUTOR no tiene atributo alguno que identifique plenamente sus instancias. Aquí podríamos haber optado por combinar los valores que adopten los atributos Nombre y Apellidos en cada tupla, para no tener que introducir ningún tipo de codificación artificial, pero ésta no deja de ser una opción arriesgada, ya que puede ocurrir que haya dos autores con el mismo nombre y apellidos. Por tanto, añadimos un nuevo atributo, llamado Código, para que no haya problemas con la clave primaria de esta entidad. Por tanto, la entidad nos queda así:

- AUTOR (Codigo, Nombre, Apellidos)

En cierto modo podríamos considerar que USUARIO tiene dos claves alternativas: Carnet y DNI. Pero en realidad no es así, ya que puede haber usuarios menores de edad que todavía no tienen DNI. Y ya sabemos que ninguna clave primaria puede admitir valores nulos, ya que entonces no serviría para distinguir unívocamente las instancias

entre sí. En cambio, todo usuario dispondrá de un número de carné. Por tanto, nos quedará la siguiente estructura:

- USUARIO (Carné, DNI, Nombre, Apellidos, Dirección, FechaNacimiento, FechaFinalExcPrest)

Por último, debemos examinar la entidad INS. El nombre de los institutos se puede repetir, y de hecho se repite. Por tanto, el atributo Nombre no es suficiente para constituir por sí solo la clave primaria de la entidad. Una opción sería establecer algún tipo de codificación. Pero, en este caso, hemos preferido considerar a INS como una entidad débil que depende de POBLACION. Por tanto, la clave primaria de INS estará formada por la clave primaria de la entidad fuerte (el atributo Nombre de POBLACION) más el atributo discriminante de la entidad débil (el atributo Nombre de INS), que también mostramos subrayado:

- INS (Nombre, Dirección, Teléfono)

2.1.7 Establecimiento de cardinalidades

El correcto establecimiento de las cardinalidades adecuadas para cada interrelación depende de las características del mundo real que se quieren modelizar.

Es necesario añadir, pues, la información referente a las cardinalidades de las interrelaciones en la documentación de diseño.

En el modelo que estamos construyendo (siguiendo el ejemplo de la BD de las bibliotecas) podemos establecer las siguientes cardinalidades:

- Un mismo usuario puede tomar en préstamo diferentes documentos, y un mismo documento se puede prestar a diferentes usuarios, a lo largo del tiempo:

Prestamo(Fecha, Tipo, Precio), entre USUARIO y DOCUMENTO: MN
--

- Lo mismo ocurre en caso de renovar un préstamo o de reservar un documento:

RenovacionPrest (Fecha), entre USUARIO y DOCUMENTO: MN
--

Reserva (Fecha), entre USUARIO y DOCUMENTO: MN
--

- Todo documento pertenecerá a la biblioteca de un INS concreto, la cual podrá disponer de muchos documentos:

Ubicación, entre INS y DOCUMENTO: 1-N

- Dentro de una misma población podrán coexistir distintos INS, pero un INS estará en una población concreta. Además, no olvidemos que, al tratarse de una entidad débil, la cardinalidad debe ser 1-N, con la entidad débil junto a la N:

Esta, entre INS y POBLACION: N-1

- Cada población pertenece a una sola comarca, pero cada comarca tendrá más de una población dentro de su territorio:

Pertenece, entre POBLACION y COMARCA: N-1

- Se ha considerado que cada documento sólo tiene un formato, pero evidentemente cada formato será aplicable a muchos documentos distintos:

Tiene1, entre FORMATO y DOCUMENTO: 1-N
--

- Con una cardinalidad MN, hacemos posible registrar a más de un autor para un mismo documento:

Tiene2, entre AUTOR y DOCUMENTO: MN

- En cambio, consideramos que cada documento sólo puede versar sobre una sola materia:

Versa, entre MATERÍA y DOCUMENTO: 1-N

- Un documento (como una película en DVD) podrá estar expresado en diferentes idiomas:

Expresado, entre IDIOMA y DOCUMENTO: MN

- Cada usuario tiene su domicilio en una población concreta. Pero en una misma población pueden vivir distintos usuarios:

Vive, entre USUARIO y POBLACION: N-1

2.1.8 Restricciones de participación y límites de cardinalidad

Se dice que la **participación de una entidad en una interrelación** es **total** si cada una de sus instancias participa una vez, como mínimo, en dicha interrelación, y **parcial** de lo contrario.

- Prestamo(Fecha, Tipo, Precio), entre USUARIO (parcial) y DOCUMENTO (parcial): MN
- RenovacionPrest(Fecha), entre USUARIO (parcial) y DOCUMENTO (parcial): MN
- Reserva (Fecha), entre USUARIO (parcial) y DOCUMENTO (parcial): MN
- Ubicación, entre INS (parcial) y DOCUMENTO (total): 1-N
- Esta, entre INS (total) y POBLACION (parcial): N-1
- Pertenece, entre POBLACION (total) y COMARCA (total): N-1
- Tiene1, entre FORMATO (parcial) y DOCUMENTO (total): 1-N
- Tiene2, entre AUTOR (total) y DOCUMENTO (total): MN
- Versa, entre MATERÍA (parcial) y DOCUMENTO (total): 1-N
- Expresado, entre IDIOMA (total) y DOCUMENTO (total): MN
- Vive, entre USUARIO (total) y POBLACION (parcial): N-1

Además, debemos establecer un límite máximo a la cardinalidad de la interrelación Prestamo porque, en principio, un usuario no puede solicitar en préstamo más de diez documentos simultáneamente.

Las interrelaciones quedarán documentadas, pues, con el siguiente formato:

Interrelacion (Atributo1, Atributo2, ...), entre ENTIDADi(participacion) y ENTITATj (participacion): tipo_de_cardinalidad

En que el nombre de la interrelación empezará con mayúsculas y, en caso de disponer de atributos, éstos irán entre paréntesis, separados por comas, y con la inicial del nombre también con mayúsculas. Habrá que especificar el nombre de las entidades que interrelaciona y el tipo de cardinalidad, que puede ser 1-1, 1-N, N-1 o MN. La participación puede ser total o parcial.

2.1.9 Elaboración de un esquema conceptual.

Una vez tenemos los requerimientos; y hemos identificado las entidades, atributos e interrelaciones; y gracias al conocimiento alcanzado de las necesidades que deben satisfacer los datos, el diseñador está en condiciones de elaborar un esquema conceptual completo de éstas y de sus interrelaciones.

Para confeccionar este modelo conceptual, el diseñador de BD utilizará algún modelo de datos que se adapte a las necesidades del proyecto, y permita continuar trabajando en él durante la fase ulterior de diseño lógico.

El modelo de datos más utilizado para la elaboración del esquema conceptual es el modelo ER, y la documentación del diseño conceptual culmina con el **diagrama ER**.

Aunque el modelo ER sigue siendo el modelo de datos más utilizado, con diferentes variaciones en aspectos de notación, cada vez se utiliza más el modelo UML de datos, muy útil en proyectos que basen la implementación de la programación en el paradigma de la orientación a objetos.

Para empezar a trabajar con los diagramas ER, puede ser cómodo establecer el esqueleto del modelo, sin incluir todavía todos los detalles (atributos, cardinalidades, etc.). En la [figura 2.1](#) existe un esquema conceptual inacabado, expresado con la ayuda del modelo de datos ER, que sólo incluye las entidades e interrelaciones detectadas inicialmente.

Figura 2.1 Diagrama ER inicial

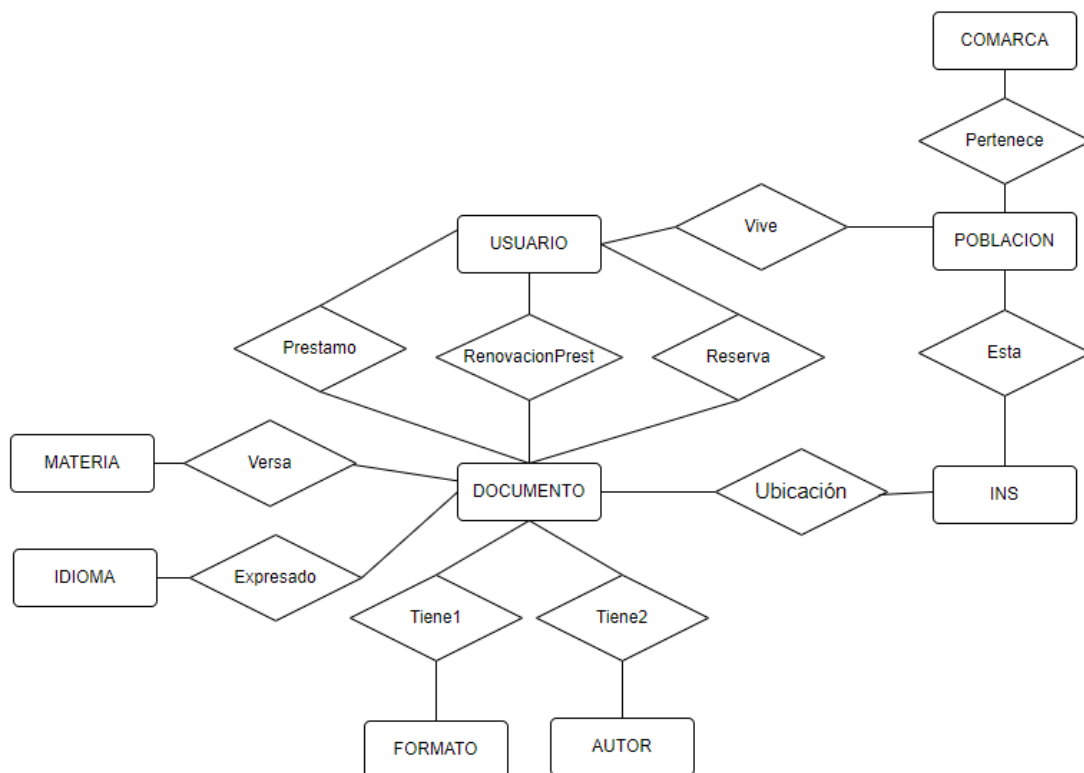
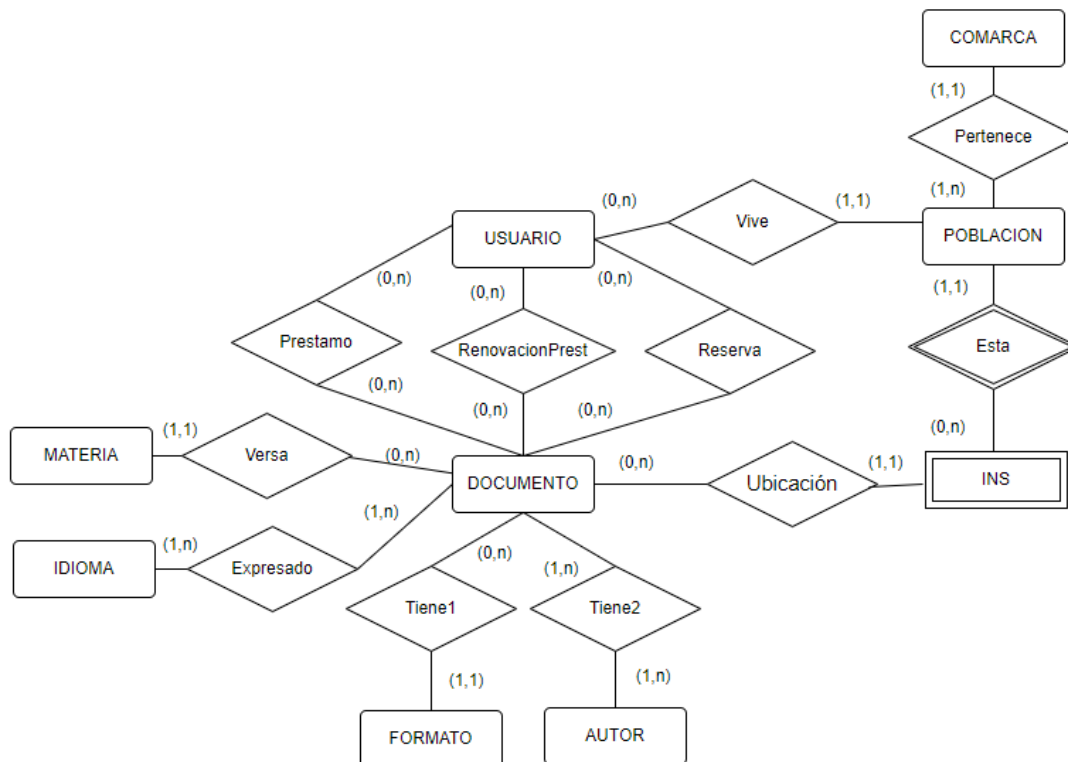


Figura 2.2 Diagrama ER a medias



A la hora de diseñar el diagrama ER, puede ser una buena práctica partir de un primer diagrama inacabado (como el de la [figura 2.1](#)) e incorporar progresivamente más detalles. En la [figura 2.2](#), por ejemplo, ya han quedado establecidas las cardinalidades y las restricciones de participación.

Y el último paso que dará el diseñador para completar el diagrama, de forma que refleje todos los requerimientos previamente detectados, sería introducir todos los atributos y claves especificados.

2.2 Extensiones del modelo Entidad-Relación.

Las estructuras básicas del modelo Entidad-Relación (modelo ER) permiten representar la mayoría de situaciones del mundo real que habitualmente es necesario incorporar en las BD. Pero, en ocasiones, ciertos aspectos de los datos deben describirse mediante unas construcciones más avanzadas del modelo ER, que conllevan una extensión del modelo ER básico. Estas ampliaciones del modelo ER consisten en la especialización, generalización y agregación de entidades.

2.2.1 Especialización.

Nos podemos encontrar con el caso de alguna entidad tipo en la que -además de las características generales, comunes a todas sus instancias- nos interese modelizar, adicionalmente, ciertas características específicas aplicables sólo aparte de sus instancias.

Entonces, podremos considerar que esta entidad tipo contiene otras entidades tipos, de nivel inferior, con características propias.

La **especialización** permite reflejar la existencia de una entidad general, llamada *entidad superclase*, que puede especializarse en diferentes entidades subclase.

La **entidad superclase** permite representar las características comunes de la entidad desde un punto de vista general. Las **entidades subclase**, en cambio, permiten representar las características propias de las especializaciones de la entidad superclase.

Las instancias de las subclases serán, al mismo tiempo, instancias de la respectiva superclase.

El proceso de designación de subclases a partir de una superclase se llama *especialización*.

La especialización en los diagramas ER se representa con un **triángulo**.

Ejemplo de especialización

Hasta ahora contábamos con una única entidad PROFESOR, que nos servía para trabajar con todos los docentes del centro, ya que todavía no habíamos detectado ningún subconjunto de este colectivo que nos hubiera hecho pensar en implementar una especialización.

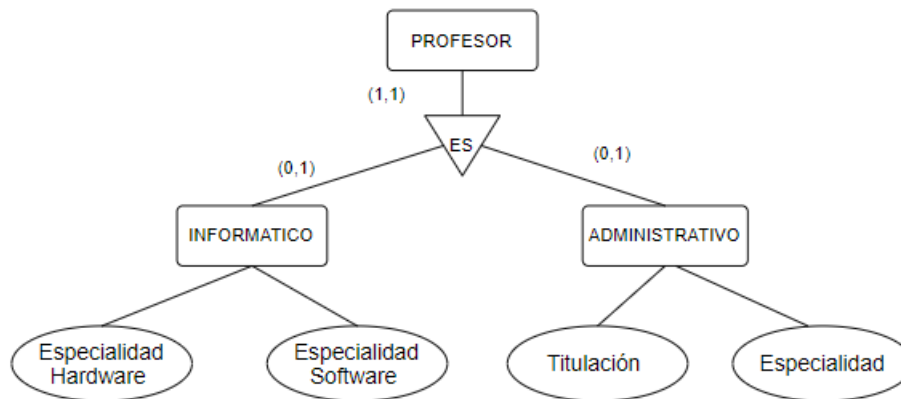
Pero resulta que la dirección del centro quiere implicar, en la gestión de este y en su mantenimiento informático, al profesorado de dos familias profesionales: la administrativa y la informática, respectivamente.

Por tanto, nos interesa tener constancia, por un lado, de la titulación de los profesores de la familia administrativa y de su especialidad, ya que en función de estas características podrán asumir, o no, las responsabilidades que se les quieren encomendar.

También puede resultar útil saber cuál es la especialidad principal, tanto en hardware como en software, del profesorado de la familia de informática, para asignar las tareas de mantenimiento con cierta garantía de éxito.

Como consecuencia de todo ello, implementaremos una especialización de la entidad PROFESOR en dos subclases: ADMINISTRATIVO, que incorporará dos nuevos atributos (Titulación y Especialidad), e INFORMATICO, que incorporará otros dos (Especialidad Hardware y Especialidad Software).

Figura 2.3 Ejemplo de especialización



La especialización permite reflejar las diferencias entre las instancias de una misma entidad, mediante el establecimiento de diferentes entidades de nivel inferior, las cuales agrupan los subconjuntos de instancias con características específicas comunes.

Estas características propias de las subclases pueden consistir tanto en la existencia de atributos como en la participación en interrelaciones, pero en ningún caso pueden ser de aplicación a todas las instancias de la superclase considerada como tal.

La **generalización**, en cambio, es el resultado de observar cómo distintas entidades preexistentes comparten ciertas características comunes (es decir, identidad de atributos o de interrelaciones en las que participan).

En función de las similitudes detectadas entre diferentes entidades, éstas pueden llegar a sintetizarse en una sola entidad, de nivel superior, mediante un proceso de generalización.

La generalización sirve para resaltar las similitudes entre entidades, por encima de las diferencias, así como para simplificar las representaciones de los datos, al evitar la repetición de atributos compartidos por distintas subclases.

Ejemplo de generalización

Hasta ahora, hemos utilizado dos entidades diferentes que nos han servido para modelizar dos categorías, también distintas, existentes en el mundo real: ALUMNO y PROFESOR.

Sin embargo, es evidente que tanto los alumnos como los profesores son personas, aunque con roles diferentes. Por tanto, tendrán una serie de características comunes, que se podrán modelizar de la misma manera.

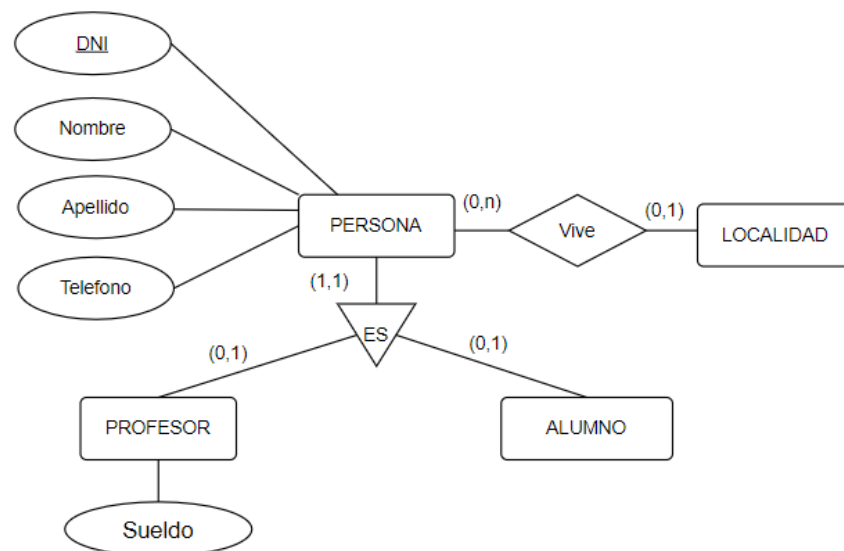
Así, tanto unos como otros tendrán nombre, apellidos, teléfonos de contacto, etc., que se podrán modelizar mediante los mismos atributos.

También es posible que ambas tipologías puedan participar en las mismas interrelaciones. Por ejemplo, para indicar la localidad de residencia, lo habitual es relacionar la entidad que representa a las personas con otra entidad que almacena las diferentes localidades.

En definitiva, partiendo de las entidades ALUMNO y PROFESOR, podríamos crear otra entidad, superclase de las anteriores, y nombrarla, por ejemplo, PERSONA.

De esta forma implementaremos la entidad PERSONA, como generalización de ALUMNO y PROFESOR, la cual contendrá los atributos comunes a sus subclases, y además participará directamente en las interrelaciones que también sean comunes a dichas subclases.

Figura 2.4 Ejemplo de generalización



El producto resultante de la especialización y de la generalización es, pues, idéntico. La diferencia entre ambos recae en el tipo de proceso que conduce a cada una:

- La especialización deriva de un proceso de diseño descendente, durante el cual, a partir de una entidad preexistente, considerada como superclase se detecta la utilidad de establecer ciertas subclases, debido a la existencia de ciertas características (atributos y participaciones en interrelaciones) no aplicables a todas las instancias de la superclase.
- La generalización responde a un proceso considerado de diseño ascendente. Durante este tipo de diseño se valora la utilidad de contemplar unas cuantas entidades preexistentes, llamadas subclases, dependientes de una misma superclase común en todas ellas. La superclase presenta unas características comunes (atributos y participaciones en interrelaciones) en todas las subclases que dependen.

Herencia de propiedades

Tanto en el caso de generalización como en el de especialización, las características de la entidad superclase se extienden hacia las entidades subclase. Como sabemos, estas características pueden consistir o bien en atributos de la entidad superclase, o bien en su participación en diferentes interrelaciones.

Llamamos **herencia de propiedades** a la transmisión de características (atributos e interrelaciones) desde la entidad superclase hacia las entidades subclase.

Puede ocurrir que una misma entidad adopte el rol de subclase en un proceso de generalización o especialización y que, al mismo tiempo, asuma el papel de superclase en otro de estos procesos en los que participe.

Cuando se produce una jerarquía de entidades, las entidades de los niveles inferiores pueden heredar características no sólo de la respectiva superclase, sino también de otras clases de niveles superiores.

Llamamos **herencia múltiple** a la recepción, por parte de una entidad subclase, tanto de las características (atributos e interrelaciones) de su superclase, como de las de otras entidades de niveles superiores, dentro de una estructura jerárquica de entidades con generalizaciones o especializaciones encadenadas.

Ejemplo de jerarquía de entidades y de herencia múltiple

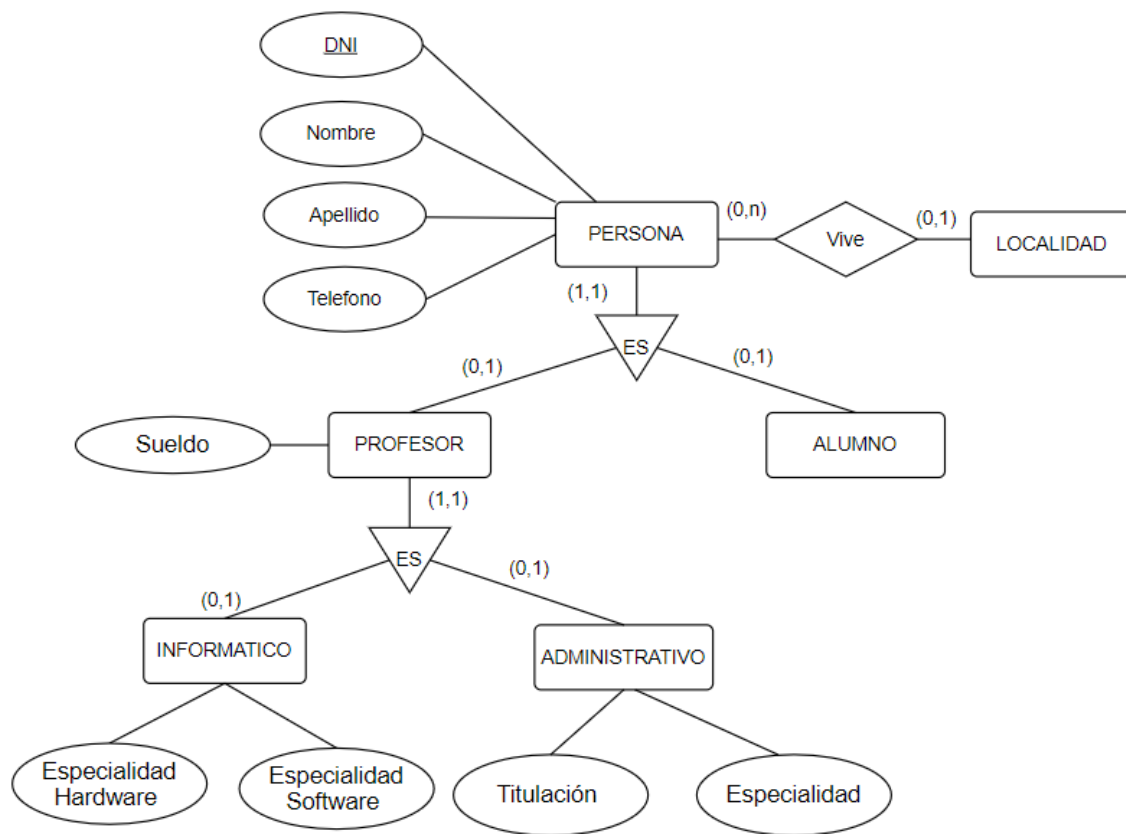
Como resultado de un proceso de generalización hemos conferido, a la entidad PERSONA, la calidad de superclase de las entidades PROFESOR y ALUMNO, consideradas como subclases de aquella.

Al mismo tiempo, pero como resultado de un proceso de especialización, le hemos conferido a la entidad PROFESOR la categoría de superclase de dos nuevas entidades, subclases de la misma: INFORMATICO y ADMINISTRATIVO.

A consecuencia de todo ello, las entidades del nivel inferior (INFORMATICA y ADMINISTRATIVO) no sólo heredarán las características de su superclase (PROFESOR), sino también las de las otras entidades de niveles superiores de las que sean descendientes y, por tanto, herederas.

En este caso, pues, INFORMATICO y ADMINISTRATIVO heredarán las características de su superclase (PROFESOR) y también las propiedades de la superclase de aquella (PERSONA). Pero no heredarán ninguna propiedad de ALUMNO, porque no son descendientes de esa entidad.

Figura 2.5 Ejemplo de herencia múltiple



Restricciones

A fin de modelizar más exactamente la parcela del mundo real que nos interese, se pueden establecer ciertas restricciones sobre las especializaciones o generalizaciones detectadas.

Un primer tipo de restricciones define si las instancias pueden pertenecer simultáneamente o no a más de una subclase de una estructura simple (es decir, que cuente con una sola superclase y un solo nivel de subclases) de generalización o especialización. En estos casos, las entidades de tipo subclase pueden ser de dos tipos:

- **Disyuntiva (o exclusiva):** Una misma entidad instancia no puede aparecer en dos entidades subclase distintas. Se representa en el diagrama añadiendo una etiqueta con la letra D.
- **No disyuntiva (o inclusiva):** Una misma entidad instancia puede aparecer en dos (o incluso en más de dos) entidades subclase distintas. Se representa en el diagrama añadiendo una etiqueta con la letra E.

Un segundo tipo de restricciones especifica si toda instancia de la superclase debe pertenecer simultáneamente a una o más de las subclases o no. Aquí las entidades de tipo subclase también pueden ser de dos tipos:

- **Totales.** Toda instancia de la entidad superclase debe pertenecer simultáneamente, como mínimo, a una de sus entidades subclase. Se denota con la etiqueta T.
- **Parciales.** Algunas instancias de la entidad superclase pueden no pertenecer simultáneamente a ninguna de sus entidades subclase. Se denota con la etiqueta P.

Combinando estas restricciones obtenemos, pues, cuatro posibilidades aplicables a las subclases de una generalización o especificación. Hay que separar las letras que se incluyen en la etiqueta con una coma:

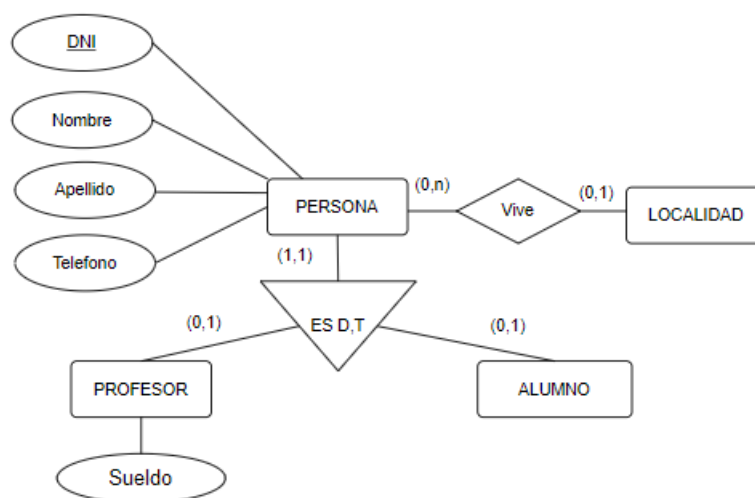
- D, T (exclusiva y totales)
- D, P (exclusiva y parciales)
- E, T (inclusiva y totales)
- E, P (inclusiva y parciales)

Ejemplo de subclases D, T

Deberemos considerar disjuntas las subclases de PERSONA si los reglamentos de funcionamiento del centro no permiten que ningún profesor se matricule como alumno, simultáneamente con el ejercicio de su labor docente.

Al mismo tiempo, las consideraremos totales si nuestra BD registra exclusivamente los datos de profesores y alumnos, sin ocuparse de otras categorías de personas (como podría ser el personal administrativo, de mantenimiento, de limpieza, etc.).

Figura 2.6 Ejemplo de subclases D,T

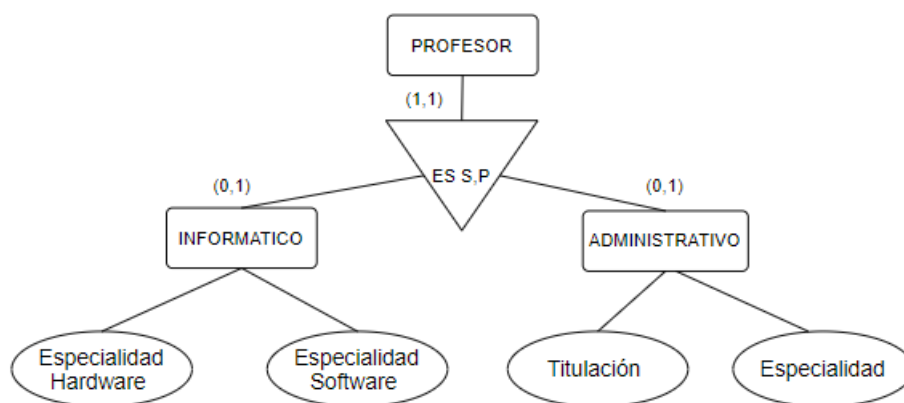


Ejemplo de subclases E, P

Deberemos considerar solapadas las subclases de PROFESOR si queremos reflejar el hecho de que algunos profesores, a pesar de ejercer como tales con una especialidad concreta en un curso académico, pueden tener otras especialidades. Por tanto, un profesor podrá ser simultáneamente INFORMATICO y ADMINISTRATIVO.

Por otra parte, las consideraremos parciales porque en nuestro instituto podrán existir, con toda seguridad, profesores de otras especialidades (como electrónicas, comerciales, etc.), que no serán ni informáticos ni administrativos.

Figura 2.7 Ejemplo de subclases E,P



Para reflejar una combinación de características aún más complicada, debe recurrirse a una especificación textual que acompañe al diagrama.

Ejemplo de subclases con distintas restricciones

Imaginemos que queremos añadir una nueva subclase de PERSONA, con el fin de incluir al personal de administración y servicios del centro. Nombraremos a esta nueva entidad ADM_SERV.

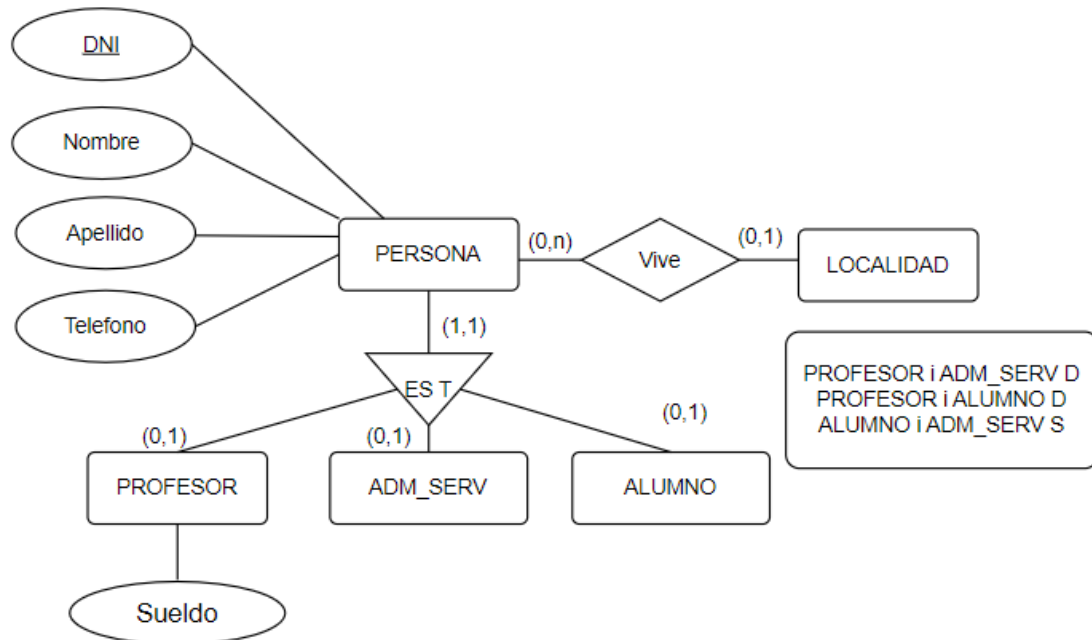
Ya hemos expuesto más arriba que los reglamentos internos de nuestro instituto no permiten que ningún profesor sea, simultáneamente, alumno del centro. Pero ahora nos encontramos con que, al personal de administración y servicios, sí se le permite matricularse como alumno en alguno de los estudios impartidos en el centro al tiempo que desempeñan su labor profesional.

Las tres subclases serán totales, como antes, porque todo el mundo tendrá que pertenecer a alguna de las tres categorías reflejadas.

PROFESOR y ALUMNO serán disjuntas entre ellas, al igual que PROFESOR y ADM_SERV, al no poder compatibilizarse las condiciones mencionadas. Pero, al mismo tiempo, ALUMNO y ADM_SERV serán solapadas, porque una persona podrá estar incluida en estas dos categorías simultáneamente, según hemos visto.

Para reflejar una realidad como ésta, no existe otro remedio que utilizar una especificación textual que, acompañando el diagrama, aclare debidamente las características específicas de cada subclase o agrupación de éstas.

Figura 2.8 Ejemplo de subclases con distintas restricciones



Notación

Tanto la especialización como la generalización se representan mediante un triángulo que incluye, en su interior, la etiqueta **ES**. Esta etiqueta indica que toda instancia de cualquiera de las subclases es, al mismo tiempo, una instancia de la superclase correspondiente (por ejemplo, tanto un informático como un administrativo serán, al mismo tiempo, un profesor).

Para distinguir claramente la superclase en los casos en que existe un gran número de entidades subclase implicadas en la estructura, o bien cuando resulta difícil, por las características del diagrama, alinear claramente todas las subclases, es conveniente indicar los límites de cardinalidad de la generalización o especialización. Para ello, basta con añadir una etiqueta del tipo mín..máx, con el fin de expresar los límites respectivos, junto a la línea que une cada entidad con el triángulo que representa la generalización o especialización, en la que mín y máx podrán tener los siguientes valores:

- **1..1** en la línea que enlaza la superclase, pues toda instancia de cualquier subclase siempre constituirá, simultáneamente, una y sólo una instancia de la superclase.
- **0..1** en la línea que enlaza cada subclase, porque no necesariamente toda instancia de la superclase deberá ser, simultáneamente, instancia de la subclase en cuestión (lo podrá ser de otra subclase, o podrá no serlo de ninguna), si estamos en presencia de una restricción de parcialidad).

Las entidades que forman parte de una estructura de generalización o especialización se representan como el resto de las entidades: cada una con un rectángulo que incorpora el nombre respectivo, y los atributos respectivos rodeados dentro de elipses ligadas a su entidad con una línea. Si los atributos forman una clave primaria, su nombre deberá subrayarse.

En términos de notación diagramática, no se establece diferencia alguna entre una generalización y una especialización. Las diferencias entre ambos fenómenos se reducen al proceso que se ha seguido para derivar en cada uno de ellos, pero no en el resultado, que es siempre el mismo: el establecimiento de una superclase y de unas subclases con unas restricciones concretas, que se representan añadiendo, en la etiqueta ES, las iniciales de las dos restricciones aplicables separadas por una coma:

- D, T (exclusiva y totales)
- D, P (exclusiva y parciales)
- E, T (inclusiva y totales)
- E, P (inclusiva y parciales)

2.2.2 Agregaciones de entidades

Con las reglas básicas del modelo ER, sólo se pueden modelizar interrelaciones en las que participan exclusivamente entidades, pero no puede expresarse la posibilidad de que una interrelación participe directamente en otra interrelación. Pero existe un mecanismo, llamado *agregación*, que permite superar la limitación descrita anteriormente, considerando una interrelación entre entidades como si fuera una entidad, y utilizándola como tal.

La **agregación de entidades** es una abstracción, mediante la cual una interrelación se trata como una entidad de mayor nivel, que agrupa a las entidades interrelacionadas gracias a ella. La agregación debe tener el mismo nombre que la interrelación sobre la que se define.

La utilidad de una agregación de entidades, pues, consiste en que la interrelación en la que se basa se puede interrelacionar con otras entidades. Una agregación de entidades se denota recuadrando a todas las entidades que participan en una interrelación determinada, a fin de construir una nueva entidad que puede establecer las propias interrelaciones.

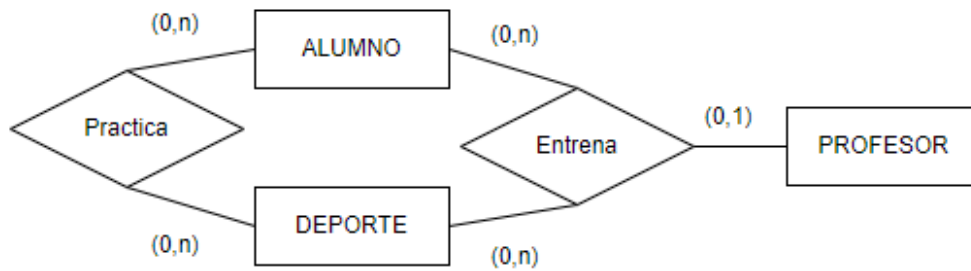
Ejemplo de agregación de entidades

Consideramos la interrelación Practica, binaria y de cardinalidad NM, que tiene lugar entre las entidades ALUMNO y DEPORTE, que ya conocemos.

Imaginemos ahora que se quiere tener constancia del profesor que, en su caso, se dedica a entrenar a un alumno que practica un deporte determinado. Y recordemos que ya existe en nuestro modelo una entidad llamada PROFESOR.

Una alternativa para representar esta realidad consistiría en crear una interrelación ternaria, llamada, por ejemplo, Entrena, entre las entidades ALUMNO, DEPORTE y PROFESOR ([figura 2.9](#)).

Figura 2.9 Ejemplo de interrelaciones redundantes



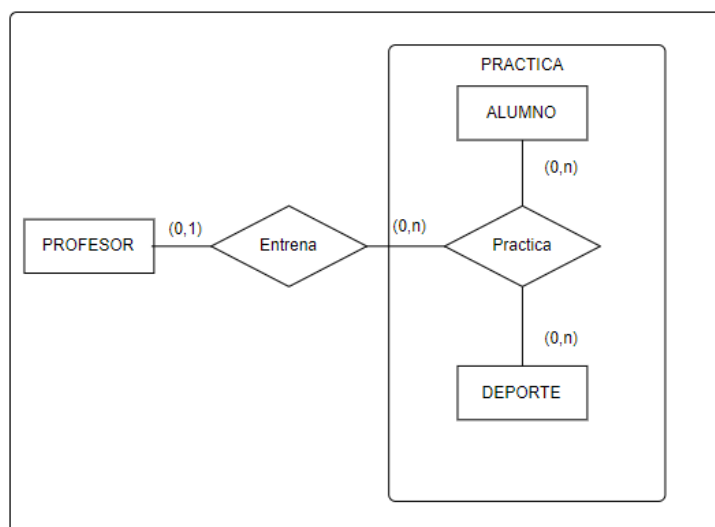
De esta forma, puede parecer que las interrelaciones Practica y Entrena se pueden combinar en una única interrelación. Pero esto no es del todo cierto, puesto que habrá interrelaciones entre ALUMNO y DEPORTE que no dispondrán necesariamente de un profesor que actúe como entrenador.

Ahora bien, existe información redundante en el esquema propuesto hasta ahora, ya que toda combinación entre instancias de las entidades ALUMNO y DEPORTE que hay en Entrena también está en Practica.

Si el entrenador sólo fuera un valor, podríamos plantearnos simplemente añadir un atributo a la interrelación Practica, que se llamara, por ejemplo, Entrenador. Pero al existir una entidad (PROFESOR) que contiene la instancia aplicable en cada caso, cuando es necesario, debemos descartar esta posibilidad.

Así pues, la mejor forma de reflejar todas estas circunstancias es hacer uso de una agregación de entidades. En este caso, es necesario considerar la interrelación Practica, entre ALUMNO y DEPORTE, como otra entidad de nivel más alto, llamada PRACTICA. Y, seguidamente, se puede establecer una interrelación binaria con cardinalidad 1-N entre PROFESOR y la agregación PRACTICA, y nombrarla Entrena, y que incluya las combinaciones necesarias entre ambas, para modelizar quien entrena la práctica de los deportes por parte de los alumnos, cuando se produce esta circunstancia ([figura 2.10](#)).

Figura 2.10 Ejemplo de agregación de entidades sin redundancia



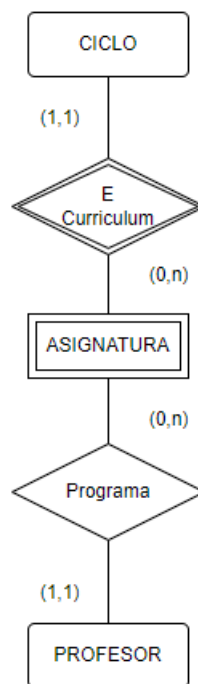
Las agregaciones de entidades en los diagramas ER se representan como **agrupación rectangular** de las entidades y relaciones que integran.

La técnica de las agregaciones engloba a la de las entidades débiles, pero aún resulta más potente: siempre que utilicemos una entidad débil, la podremos sustituir por una agregación, pero no al revés. Ahora bien, es necesario mantener a las entidades débiles en el modelo ER porque, aunque resultan menos complejas que las agregaciones, normalmente son suficientes para modelizar la mayoría de las situaciones que se producen en el mundo real.

Ejemplo de sustitución de una entidad débil por una agregación

Recuperamos la entidad débil ASIGNATURA. Ahora imaginamos que, para establecer cierto control en materia de coordinación pedagógica, se necesita saber quién es el profesor responsable de realizar la programación didáctica de cada asignatura.

Entre PROFESOR y ASIGNATURA puede establecerse una interrelación binaria de cardinalidad 1-N para representar este hecho, que se llama, por ejemplo, Programa.



Pero, alternativamente, podríamos modelizar este dato convirtiendo ASIGNATURA en una agregación.

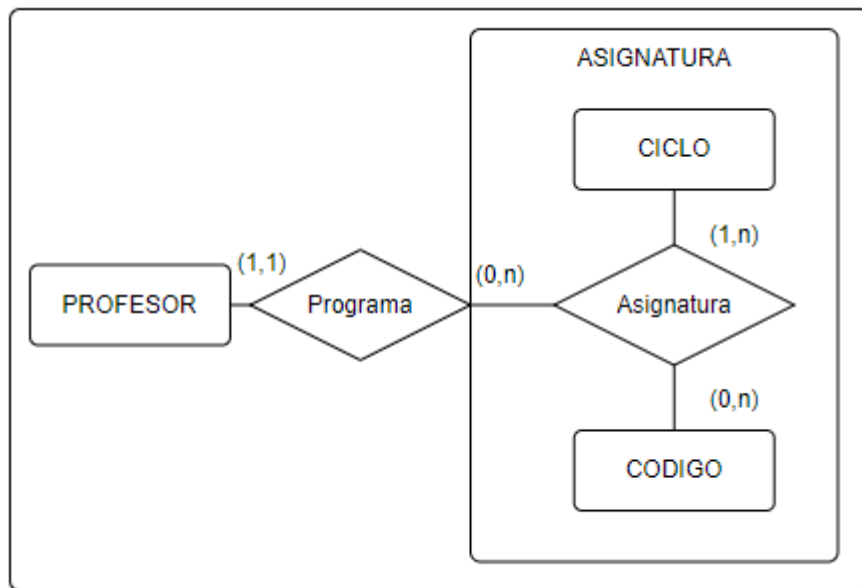
Para conseguirlo, en primer lugar, deberíamos considerar una nueva entidad, y nombrarla por ejemplo CODI, que almacenara simplemente códigos de asignatura (como C1, C2, C3, etc.).

A continuación, deberíamos establecer, por un lado, una interrelación binaria de cardinalidad NM entre CICLO y CÓDIGO, y nombrarla Asignatura. Y, por otra parte,

también deberíamos obtener una agregación de la interrelación entre CICLO y CÓDIGO (es decir, Asignatura).

Por último, deberíamos interrelacionar la agregación resultante con la entidad PROFESOR, con una sencilla interrelación binaria con cardinalidad 1-N, llamada Programa ([figura 2.12](#)).

Figura 2.12 Ejemplo de sustitución de una entidad débil por una agregación (resultado)



Las agregaciones también se representan frecuentemente **incluyendo, dentro de un recuadro, sólo el rombo de cuya interrelación provienen las entidades implicadas**.

Notación

Las agregaciones de entidades se representan incluyendo dentro de un recuadro a todas las entidades que participan en una interrelación determinada.

Desde una interrelación, se puede enviar una flecha (de punta sencilla o doble, para expresar la cardinalidad 1 o N, respectivamente) hasta el rombo incluido dentro del recuadro que indica la existencia de una agregación (o hasta el mismo recuadro, exactamente igual que si se tratara de una simple entidad).

Toda agregación debe tener el mismo nombre que la interrelación sobre la que se define.

2.2.3 Ejemplo: BD de un instituto de formación profesional

Ahora desarrollamos un ejemplo de diseño conceptual de BD, correspondiente a un instituto de formación profesional, para ilustrar por separado los distintos conceptos y su respectiva modelización. Se trata de diseñar una BD para gestionar al personal del instituto (compuesto por los profesores, y por los trabajadores de administración y servicios) y su alumnado, además de los estudios impartidos.

Las siguientes descripciones resumen los requisitos de los usuarios de la futura BD:

- Las personas que forman parte de nuestra comunidad educativa se identifican mediante el DNI (o documento equivalente, como la tarjeta de residencia).
- También queremos conocer, de estas personas, el nombre y apellidos, la dirección, un (y sólo uno, por el momento) teléfono de contacto, y la fecha de nacimiento.
- Además, debemos tener registrada la localidad de residencia, teniendo en cuenta que la BD debe poder almacenar, para otros usos, localidades donde nadie viva.
- Toda persona de la comunidad educativa pertenece, como mínimo, a uno de los tres subtipos siguientes:
 - Profesores
 - Alumnos
 - Personal de administración y servicios
-
- Las personas sólo pueden mantener un tipo de relación laboral con el centro educativo. Por tanto, los profesores no pueden pertenecer simultáneamente al colectivo del personal de administración y servicios.
- Tampoco está permitido que los profesores se matriculen en el centro de trabajo en ninguno de los estudios impartidos. Por tanto, los profesores no se pueden considerar simultáneamente alumnos.
- Sin embargo, sí está permitido a los integrantes del colectivo de administración y servicios que se matriculen, fuera de su horario laboral, en alguno de los estudios impartidos. Por tanto, el personal de administración y servicios puede pertenecer simultáneamente a la categoría de alumno.
- Debemos tener constancia del sueldo de los profesores y del personal de administración y servicios.
- Organizativamente todo profesor está asignado a un solo departamento. Y cada departamento tiene asignado a uno de sus profesores como coordinador.
- Todo profesor tiene reconocida una determinada especialidad (o más de una). Pero internamente, la BD del instituto sólo necesita registrar cuáles de los profesores que tiene asignados el centro pertenecen a las especialidades de informática o de administración, a fin de asignarles tareas específicas adicionales a las docentes que les son propias.
 - De cada informático, queremos saber las especialidades profesionales, cuando las haya, tanto en el ámbito del hardware como también en el del software.

- De cada administrativo, queremos conocer la titulación académica y la especialidad profesional, si la tiene.
- Los alumnos pueden practicar algunos deportes en las instalaciones del centro. E incluso pueden disponer de algunos profesores como entrenadores personales, que se han ofrecido voluntariamente para realizar esta tarea.
- Como es lógico, al tratarse de un centro de formación profesional, el instituto de nuestro ejemplo ofrece diferentes estudios estructurados en ciclos formativos, y cada ciclo formativo tiene sus propias asignaturas. Nos interesa, pues, codificar las asignaturas de la misma forma que se hace en el currículo oficial del ciclo formativo al que pertenecen. El problema es que estos códigos se repiten para todos los ciclos formativos, ya que la codificación siempre consiste en una C (por ser la inicial de la palabra *crédito*) seguida de un número entero (C1, C2, C3, y así sucesivamente).
- Dentro de un mismo ciclo formativo, se puede exigir a los alumnos que, por matricularse en algunas asignaturas, hayan superado alguna asignatura (o más de una).
- Por otra parte, siempre existe un profesor encargado de realizar la programación didáctica de cada asignatura. Sin embargo, un mismo profesor puede encargarse de la programación didáctica de más de una asignatura.
- Todos los alumnos del centro tienen un compañero que actúa como delegado en el ámbito de una asignatura y se encarga, por ejemplo, de distribuir los materiales o baterías de ejercicios. Un mismo alumno puede actuar como delegado en el ámbito de más de una asignatura. Pero cada alumno sólo tendrá un delegado en cada asignatura en el que esté matriculado.
- Finalmente, la BD debe recoger en qué asignaturas está matriculado cada alumno en cada curso académico, y la nota final obtenida.

La [figura 2.13](#) muestra un diagrama ER que cumple con los requisitos mencionados anteriormente.

A continuación, se muestra una lista de todas las entidades que aparecen en el diagrama, acompañadas de sus respectivos atributos (subrayados si forman parte de una clave primaria).

- DEPARTAMENTO
 - *CódigoDepartamento* , NombreDepartamento
- LOCALIDAD
 - *CódigoLocalidad* , NombreLocalidad
- PERSONA
 - *DNI* , Nombre, Apellidos, Dirección, Teléfono, FechaNacimiento
- PROFESOR (subclase de PERSONA)
 - *DNI* , Sueldo
- ADMO_SERVICIO (subclase de PERSONA)
 - *DNI* , Sueldo

- ALUMNO (subclase de PERSONA)
 - *DNI*
- INFORMATICO (subclase de PROFESOR)
 - *DNI* , EspecialidadHardware, EspecialidadSoftware
- ADMINISTRATIVO (subclase de PROFESOR)
 - *DNI* , Titulación, Especialidad
- DEPORTE
 - *CódigoDeporte*
- CURSO
 - *CodiCurso*
- CICLO
 - *CódigoCiclo* , NombreCiclo
- ASIGNATURA (entidad débil: CodiAsignatura la identifica parcialmente, y necesita el código del ciclo correspondiente para identificarse completamente)
 - *CódigoAsignatura* , NombreAsignatura

Por último, hay que comentar algunos de los aspectos más complejos de este modelo, proporcionado a modo de ejemplo:

- Las subclases en las que se especializa PROFESOR (INFORMATICO y ADMINISTRATIVO) son solapadas (E) entre ellas, y además parciales (P):
 - Son solapadas porque las instancias de la superclase pueden pertenecer simultáneamente a ambas categorías.
 - Son parciales, porque no todas las instancias de la superclase deben pertenecer necesariamente a alguna de las dos categorías.
- Las subclases que dan lugar a la generalización de PERSONA (PROFESOR, ADM_SERV y ALUMNO) son totales, pues toda instancia de la superclase debe pertenecer simultáneamente a alguna de las tres subclases citadas. Sin embargo, respecto al hecho de si pueden pertenecer simultáneamente a diferentes subclases o no, tienen restricciones específicas, y las combinan de dos en dos. Esta particularidad está especificada textualmente dentro del diagrama:
 - PROFESOR y ADM_SERV: las entidades son disjuntas entre sí, porque las instancias respectivas no pueden pertenecer al mismo tiempo a ambas.
 - PROFESOR y ALUMNO: se da la misma circunstancia que en el caso anterior.
 - ALUMNO y ADM_SERV: las entidades son solapadas entre sí, porque las instancias respectivas sí pueden pertenecer al mismo tiempo a ambas.
- Entre DEPARTAMENTO y PROFESOR existen dos interrelaciones porque sirven para modelizar dos realidades diferentes: la coordinación del

departamento por parte de un profesor (con cardinalidad 1:1), y el hecho de que una pluralidad de profesores estén adscritos al mismo (con cardinalidad 1 :N).

- La localidad de residencia de las personas se ha implementado mediante entidad independiente, y no como un simple atributo de la entidad PERSONA. De esta forma, se evitarán redundancias, porque cada localidad sólo se registrará una vez dentro de la BD, aunque después se interrelacionará con todas las instancias de PERSONA que sea necesario.
- Se ha optado por establecer una agregación a partir de la interrelación Practica, a fin de permitir establecer otra interrelación (Entrena) entre ésta y PROFESOR, que evita la redundancia de datos que habría si se hubiera utilizado una interrelación ternaria entre ALUMNO, DEPORTE y PROFESOR, ya que contendría todas las combinaciones de la interrelación entre ALUMNO y DEPORTE. Y no se podría implementar simplemente una ternaria entre ALUMNO, DEPORTE y PROFESOR, y suprimir dicha binaria, porque los alumnos también pueden practicar los deportes por libre, sin ningún profesor que los entrene.
- Para representar la figura del alumno delegado de asignatura, ha sido necesario recurrir a una interrelación recursiva ternaria, ya que es necesario interrelacionar a cada alumno que actúa como delegado con sus alumnos representados y, además, con la asignatura de qué se trate en cada caso.
- Para representar los prerrequisitos de matriculación, hemos añadido otra interrelación recursiva, en este caso binaria, que sirve para asociar las asignaturas entre sí cuando es necesario.
- Fijémonos en que la interrelación ternaria Matricula, entre CURSO, ALUMNO y ASIGNATURA, con cardinalidad M:N:P, tiene un atributo propio para almacenar la nota final de cada alumno.

Figura 2.13 Ejemplo: BD de un instituto de formación profesional

