



**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación

GUÍA MODELOS ENTIDAD RELACIÓN

**Departamento Nacional de Planeación
Bogotá, 2019**

HISTORIAL DE VERSIONES

Versión	Fecha	Descripción	Autor
1.0	2019-08-30	Versión Inicial	Jorge Valenzuela
1.5	2020-07-12	Versión ajustada con descripción detallada de contenidos	Jorge Eduardo Henao

CUADRO DE ELABORACIÓN Y APROBACIÓN.

Elaborado Por	Revisado Por	Aprobado Por
Nombre: Jorge Valenzuela	Nombre: Jorge Eduardo Henao Garzón	Nombre: Diego Alonso Espinoza Jiménez
Cargo: Ingeniero de Requerimientos	Cargo: Arquitecto Empresarial	Cargo: Coordinador GGSI
Fecha: 30-08-2019	Fecha: 12-07-2020	Fecha: 12-07-2020

CONTENIDO

1	OBJETIVO	6
2	CONCEPTOS	7
3	ALCANCE	7
4	IMPORTANCIA DEL MODELADO EN INGENIERÍA DE SOFTWARE	8
5	VOCABULARIO, SINTAXIS Y NOTACIONES PARA MODELOS E/R	8
5.1	VOCABULARIO	8
5.2	SINTAXIS	9
5.3	NOTACIÓN	10
6	EJEMPLO	14
7	GLOSARIO DE TÉRMINOS Y DEFINICIONES	6
8	BIBLIOGRAFÍA	16

Índice de ilustraciones

Ilustración 1. Ejemplo de un modelo entidad relación con notación de Chen	11
Ilustración 2. Ejemplos de representación de una relación con diferentes notaciones	13
Ilustración 3. Representación de una relación entre dos entidades con UML.....	14
Ilustración 4. Diagrama o representación gráfica del modelo	16

Índice de tablas

Tabla 1. Cardinalidad en las notaciones más comunes	12
Tabla 2. Ejemplo lectura interpretativa en las notaciones más utilizadas	13
Tabla 3. Punto de vista basado en lenguaje UML.....	14

1 OBJETIVO

El objetivo de este documento es proporcionar una introducción a los modelos entidad relación como artefacto con doble utilidad: por una parte, como herramienta para entender, discutir y comunicar un contexto o espacio de la realidad y por otra parte como mecanismo para definir, organizar, discutir, estructurar y presentar los datos de un sistema de información (S.I.) que representará dicha realidad. El documento también provee un ejemplo práctico para entender mejor este artefacto, su elaboración, sus objetivos, ventajas y el impacto en la calidad final de un modelo físico de base de datos relacional, que es el componente fundamental de datos para sistemas de información transaccional en el DNP.

Se presentan unos conceptos generales que ilustran los propósitos fundamentales de un modelo entidad relación. Posteriormente introduce los elementos principales del modelo considerados dentro de un contexto o ámbito particular denominado el dominio del discurso. En seguida se hace énfasis en la importancia del modelo como herramienta semántica y de gestión de conocimiento. Luego se describe y explica los lenguajes de modelado desde su vocabulario, sintaxis y notación. Por último, se ofrece un ejemplo breve pero representativo de un modelo entidad relación, diagramado con notación Barker.

2 TÉRMINOS Y DEFINICIONES

Notación: Sistema de signos convencionales que se adopta para expresar conceptos matemáticos, físicos, químicos, de ingeniería, etc.

Diagrama: Representación gráfica, generalmente esquemática, para describir o comunicar algo.

Base de datos: Información en movimiento, desde la fuente en donde se genera el dato o la información, se transforma a través de procesos y llega al usuario o dependencia para apoyar la gestión.

Instancia: Caso o ejemplar particular y diferenciado de los elementos de un conjunto o de los individuos de una clase o especie.

Modelo: Esquema teórico y formal de un sistema o de una realidad compleja que se elabora para facilitar su comprensión y el estudio de su comportamiento

Modelo Relacional: organización y gestión de bases de datos consistente en el almacenamiento de datos en tablas compuestas por filas, o tuplas, y columnas o campos.

Relación: Es el elemento central del modelo relacional. Son los objetos principales sobre los que debe recogerse información y generalmente denotan personas, lugares, cosas o eventos de interés. Una relación tiene un nombre, un conjunto de atributos que representan sus propiedades

Tupla: Representa cada una de las ocurrencias de la relación (equivale a lo que conocemos como ocurrencia de un registro, en gestión de archivos)

3 CONCEPTOS

Un Modelo Entidad Relación (MER) es una herramienta de la ingeniería de software con dos propósitos: uno – que podemos llamar secundario – que nos facilita definir la organización estructural de los datos de un Sistema de Información y otro – que podemos llamar principal – que nos permite elaborar, compartir y comunicar con usuarios no técnicos un entendimiento consensuado de los conceptos, sus características y sus relaciones asociados a la realidad que debe representar el Sistema de Información.

Existe una sobrevaloración subjetiva del propósito secundario de un MER desde los ingenieros de software, principalmente los programadores, para estos, lo preponderante es la estructura final de los datos que van a ser persistidos y manejados en un sistema específico de gestión de bases de datos. Dado que es posible determinar de cualquier manera una organización de datos para un Sistema de Información, se suele obviar o desestimar la elaboración de un modelo entidad relación o, peor aún, se suele confundir con un diagrama que representa el diseño físico de una base de datos.

Para la OTSI y el DNP el valor fundamental del MER es el de transferencia de conocimiento institucional derivada de su propósito principal, este enfoque guarda conformidad con lo expresado por Peter Chen, el autor original del modelo, quien dejaba claro la importancia de incorporar información semántica importante sobre el mundo real y con el autor Richard Barker, quien extendió el modelo y aportó una nueva notación, Barker también reiteraba el valor del modelo en la comunicación con otros interesados e involucrados en el desarrollo y uso del Sistema, interesados que son diferentes a los ingenieros de software.

Un tercer autor, John Cartmell, ha continuado la evolución del modelo y aunque ha eliminado la palabra relación dejando solo modelado de entidades (Entity Modelling) mantiene su propósito usual de servir como estructuración de los datos de un sistema de información hace especial relevancia en el propósito del modelo en ofrecer un marco de trabajo para desarrollar y representar conceptos que se usan en dominios de conocimiento con los cuales no se está familiarizado.

Por otra parte, elaborar un MER considerando su propósito principal, redundante en la calidad final del diseño físico de la base de datos, pues casi que garantiza que el diseño resultante estará en tercera forma normal y libre de anomalías como redundancia o falta de integridad referencial.

4 ALCANCE

En esencia, un MER establece o propone un dominio (o universo) del discurso en forma de narración o descripción de un contexto particular de la realidad a través de un conjunto no vacío de cosas acerca de las cuales se habla en ese determinado contexto.

Esas cosas de las que se habla en el dominio del discurso son las que se denominan **entidades** dentro del MER, las entidades se consideran en abstracto como un tipo de noción o concepto que representa posibles instancias particulares de ese tipo de entidad; de las entidades se predicen hechos o propiedades, algunas propiedades son importantes dentro del dominio del discurso porque caracterizan y describen a la entidad, estas propiedades o características se denominan **atributos** de la entidad. Algunos de los atributos de la entidad la describen, otros, por adjetivación, la clasifican, la cuantifican, la cualifican, la identifican o expresan un estado de esta.

El tercer tipo de elementos de un MER son las **relaciones**, que establecen conexiones entre las entidades consideradas en el modelo, con contenido semántico, - es decir, significativo y relevante dentro del dominio del discurso -. Las relaciones entre entidades en un MER pueden significar una composición entre instancias de entidades (todo / parte) o de referencia o de correspondencia etc.

5 IMPORTANCIA DEL MODELADO EN INGENIERÍA DE SOFTWARE

Para la ingeniería y la ciencia práctica, modelar o hacer modelos es una tarea de conceptualización y comunicación; los modelos constituyen un primer lenguaje de discusión y entendimiento, modelar es recorrer el camino entre comprender y hacer. La utilización sistemática de modelos en actividades de ingeniería no es concepto moderno, es difícil imaginar, por ejemplo, la construcción de grandes obras de ingeniería civil sin apelar a su uso. La ingeniería de software también da tal preponderancia a los modelos que se convierte en utilidad efectiva para la resolución de problemas y están siempre presentes en todo el ciclo de vida del software.

Dentro del ámbito de la comunicación, los modelos son constructos lingüísticos con sintaxis y semántica, es decir, tienen notación y significado que permiten representar una realidad (describen el dominio del discurso, una problemática, un sistema de gestión bajo estudio, etc.) o ilustran una propuesta posible (prescriben el diseño o especificación de una solución o sistema de información, un artefacto, un componente, etc.).

En realidad, y como podría deducirse del párrafo anterior, antes que hablar de modelos es conveniente hablar de lenguajes de modelado y precisar que, dependiendo del conjunto de símbolos que conforman su notación o vocabulario y de las reglas de las expresiones válidas que definen su sintaxis, un lenguaje de modelado puede ser textual – como XML – o visual – como UML – o utilizar simultáneamente símbolos y palabras.

Los lenguajes de modelado más útiles combinan representación icónica y de texto en su notación y sintaxis lo que permite una mejor interpretación semántica de los modelos producidos. El lenguaje de modelado entidad relación pertenece a este tipo de lenguajes.

De cualquier manera, sea textual, gráfico o combinado, un modelo como producto lingüístico debe ser, de acuerdo con su notación y sintaxis, consistente y bien formado frente a la realidad que representa (sea descriptiva o prescriptiva) y, de acuerdo con sus reglas semánticas, lógico y consecuente frente al significado de las construcciones que ofrece.

6 VOCABULARIO, SINTAXIS Y NOTACIONES PARA MODELOS E/R

6.1 VOCABULARIO

Los antecedentes del MER se encuentran en un documento publicado por Charles Bachman en 1969 [1], trabajo que es anterior inclusive al artículo de Edgar Codd que dio lugar al modelo relacional de base de datos [2]. El artículo de Bachman introduce el concepto de Entidad asociado al de clases de entidad (entity class) y al de conjuntos de entidades (entity set) que conforman relaciones entre entidades. Además de introducir los conceptos, Bachman propone lo que sería la primera notación gráfica sobre un modelo de datos, en donde las entidades se representan con cajas y las relaciones con flechas dirigidas que conectan las entidades.

Peter Pin-Shan Chen es el creador del modelo entidad relación y se basó en las ideas de Bachman, a quien cita y agradece por la motivación en su artículo seminal de 1976 [3], el modelo original propuesto en el documento es uno de los trabajos más citados en el campo de las ciencias de la ingeniería del software y goza de prestigio y galardones. Su léxico o vocabulario de conceptos básicos se ha mantenido invariable hasta ahora y comprende los tres elementos ya mencionados en el capítulo 3:

- Entidad
- Atributo
- Relación

Entidad es una cosa, objeto o evento con existencia física o conceptual (real o imaginario), acerca del cual existe la necesidad de almacenar, gestionar o mantener información de conformidad con el comportamiento y funcionalidad dentro del dominio del discurso o realidad considerada para el modelo.

Atributo es algo que se predica o se dice de una entidad como característica o propiedad de esta. Usualmente una entidad tiene un número plural de atributos y estos sirven a una entidad para, entre otras cosas:

- Identificarla.
- Describirla.
- Cualificarla.
- Clasificarla.
- Expresar un estado.

Relación es una conexión o asociación con contenido semántico - es decir, significativa y nombrable – entre entidades consideradas dentro del dominio del discurso.

6.2 SINTAXIS

La sintaxis define las reglas o restricciones aplicables a los elementos de los modelos entidad relación bien formados.

Cada tipo de entidad debe identificarse de forma única con un nombre que represente un conjunto o clase de cosas, no una instancia particular y se recomienda que ese nombre sea en mayúsculas fijas y en singular. Por ejemplo, un tipo de entidad que guarde información de individuos humanos debería llamarse PERSONA, de ninguna manera Carolina o Mauricio, que serían posibles instancias de la entidad PERSONA.

Las entidades son mutuamente exclusivas, un objeto o cosa considerada en el modelo solo puede ser representado por un tipo de entidad.

Cada instancia de entidad debe encontrarse separada e identificada de forma única y diferenciada de las demás instancias. Si existieran dos instancias de PERSONA con valor “Carolina” para un atributo nombre, deben existir otro(s) atributo(s) cuyo valor(es) diferencie(n) cada instancia.

Las relaciones son asociaciones binarias en el sentido de que corresponde siempre a la interacción entre dos entidades o de una entidad consigo misma.

Toda relación tiene dos extremos, para cada uno de los cuales debe proveerse:

- Una leyenda textual.
- Un grado o cardinalidad (cero, uno o muchos).

La leyenda define y explica textualmente la semántica de la relación, el grado (o cardinalidad de mapeo) indica para cada instancia de una entidad con cuantas instancias de la otra entidad puede asociarse, la cardinalidad no se define

con cantidades exactas pero si precisa determinar un número mínimo (cero o uno) y un número máximo (cero, uno o muchos) en cada lado de la relación.

Para algunos autores el número mínimo en la cardinalidad implica una semántica diferenciada que define la opcionalidad u obligatoriedad de la relación, es decir, si el número mínimo es cero la relación es opcional (puede o no existir la relación), de lo contrario es obligatoria (debe existir la relación), esa diferenciación se expresa en la notación.

Considerando dos entidades, A y B, las cardinalidades válidas para representar en el modelo son las siguientes:

- Uno a Uno (1:1). Una instancia de una entidad A se relaciona con solo una instancia de la entidad B.
- Uno a Muchos (1:M). Una instancia de una entidad A se relaciona con muchas instancias de la entidad B.
- Muchos a Muchos (M:N). Una instancia de una entidad A se relaciona con muchas instancias de la entidad B y a la vez una instancia de la entidad B se relaciona con muchas instancias de la entidad A.

La obligatoriedad indica si esa relación se fuerza o no, es decir, si debe existir una instancia de la entidad para exigir la relación o si la relación es opcional, es decir, puede no existir una instancia de entidad.

Algunos autores tratan la cardinalidad y la opcionalidad bajo un mismo concepto, en cuyo caso agregan a la cardinalidad un valor mínimo de cero (0). Otros autores, como Richard Barker, las tratan como conceptos distintos, que se muestran esquemáticamente diferenciados.

Los atributos se representan textualmente con un nombre en singular y en minúsculas para diferenciarlos de las entidades.

6.3 NOTACIÓN

La notación original de Peter Chen considera iconos o figuras para cada elemento del modelo, así, las entidades se representan con rectángulos, las relaciones con un rombo y los atributos con óvalos. Las líneas unen las relaciones (los rombos) con las entidades (rectángulos) y las entidades con los atributos (óvalos), tal como se muestra en la figura x.

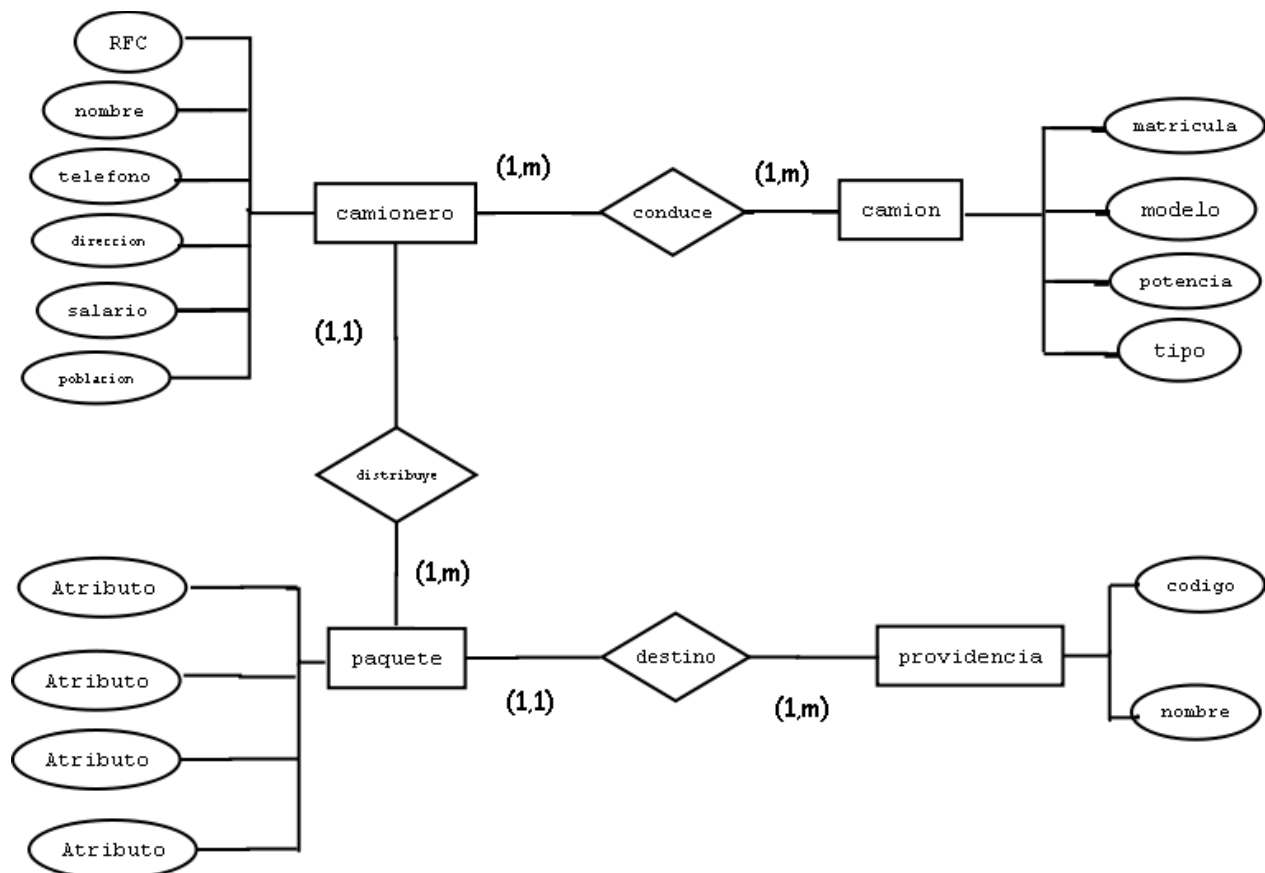


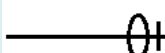
Ilustración 1. Ejemplo de un modelo entidad relación con notación de Chen

Si bien la notación de Chen sigue siendo muy utilizada, otras notaciones posteriores han logrado aceptación y popularidad, particularmente tres se consideran válidas en el DNP para dibujar el modelo:

- Notación Barker
- Notación IE (Information Engineering)
- Notación IDEF1X

Estas tres notaciones obvian el elemento gráfico para los atributos y los relegan a simple descripción textual dentro de las entidades, las entidades mantienen la forma de rectángulo, pero en la notación de Barker los bordes son redondeados. Las relaciones se representan igualmente con líneas que conectan las entidades, pero ciertas variaciones en el dibujo de las líneas aportan semántica para la cardinalidad y la opcionalidad

Las tres notaciones manejan cardinalidad y opcionalidad. La siguiente gráfica muestra las diferencias:

Grado	IE	IDEF1X	Barker
Cero uno	Círculo vacío y barra 	Línea con círculo relleno y Z	Línea punteada

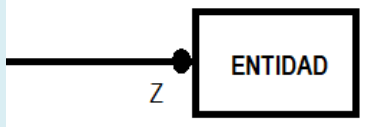
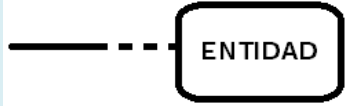

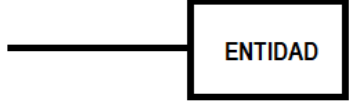
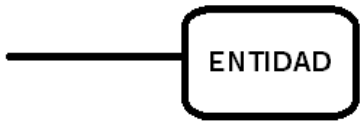

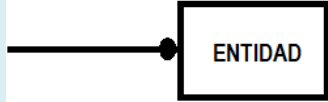
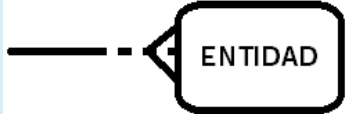

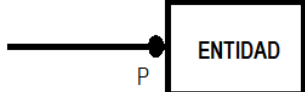

			
Uno y solo uno	Dos barras al final de la línea 	Línea continua 	Línea continua 
Cero o muchos	Círculo hueco y pata de gallo 	Línea con círculo relleno 	Línea punteada con pata de gallo 
Uno o muchos	Una barra y pata de gallo 	Línea con círculo relleno y P 	Línea continua con pata de gallo 

Tabla 1. Cardinalidad en las notaciones más comunes

Un ejemplo más elaborado para representar la relación entre dos entidades dadas (Empleado y dependencia), establece que es opcional que una **Dependencia** tenga empleados (es decir, la empresa permite, como regla de negocio, que pueda existir una dependencia aunque no tenga empleados a cargo). Como se ha mencionado, la opcionalidad puede manejarse dentro de la cardinalidad, considerando el cero (0) como opcional. A continuación, el ejemplo gráfico en las 3 notaciones:

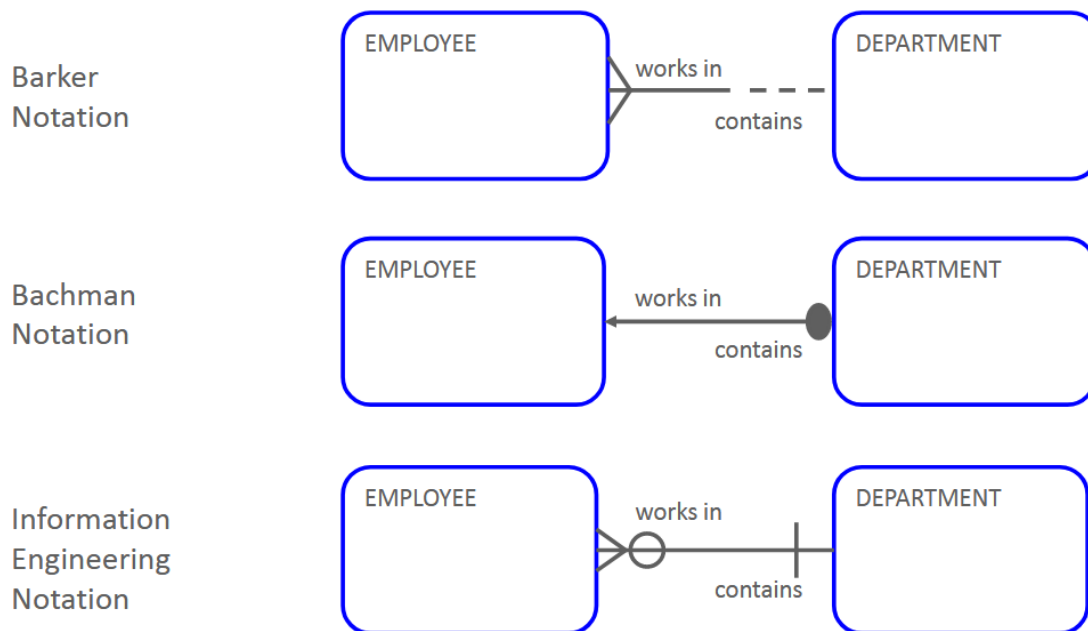


Ilustración 2. Ejemplos de representación de una relación con diferentes notaciones

La interpretación del modelo demanda enunciar la relación desde una entidad a la otra y viceversa, entonces las lecturas serían:

Notación	Lectura interpretativa
Barker	Cada Empleado debe trabajar para una y solo una Dependencia Cada Dependencia puede estar compuesta de uno o más Empleados
Idef1x	Cada Empleado trabaja en una y solo una Dependencia Cada Dependencia está compuesta de cero, uno o más empleados
Information Engineering	Cada empleado trabaja en una y solo una dependencia Cada dependencia está compuesta de cero, uno o más Empleados

Tabla 2. Ejemplo lectura interpretativa en las notaciones más utilizadas

Como alternativa a las notaciones citadas, se puede emplear un lenguaje de dominio genérico para elaborar modelos de datos, tal como el lenguaje UML. Pero esto se hace adaptando los diagramas de clase de UML como diagramas de datos, por ejemplo, debe añadir estereotipos (nuevos símbolos, por ejemplo <<entity>> o <<id>>), evitar algunos símbolos e inclusive cambiar la semántica de ciertos elementos (considerar las relaciones como predicados, no como roles); éstas y otras razones evidencian el esfuerzo que conlleva forzar una notación genérica en vez de usar un lenguaje específico de dominio, como las notaciones que producen modelos entidad relación, pues ofrecen símbolos más apropiados para describir y especificar datos. Además, no se debe olvidar que un modelo de clases no es lo mismo que un modelo entidad relación. Por ello, se privilegia y aconseja usar las notaciones entidad relación enumeradas anteriormente. Por lo anterior, debe quedar claridad en los puntos de vista utilizados, es decir, en la descripción de los componentes gráficos que sirven para la construcción del modelo. Como ejemplo, se plasma a continuación como adaptar los diagramas de clase UML para representación de modelos entidad relación.

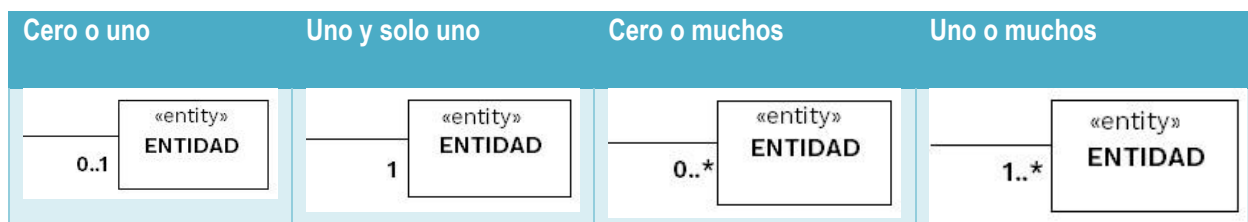


Tabla 3. Punto de vista basado en lenguaje UML

Notación UML para representar la relación ejemplo entre la entidad Dependencia y la entidad Empleado.

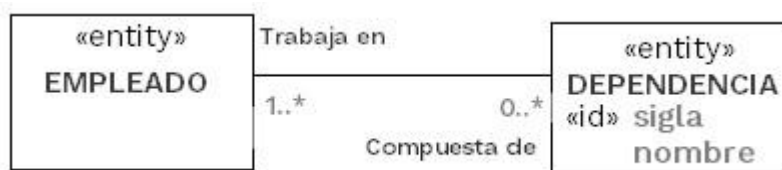


Ilustración 3. Representación de una relación entre dos entidades con UML

7 EJEMPLO

El ejemplo que se va a exponer en este documento corresponde a un sistema de información para gestionar las evaluaciones y notas en un colegio colombiano que ofrece cursos para primaria, secundaria y educación media.

Las entidades más relevantes que conforman el dominio del discurso son las siguientes:

NIVEL EDUCATIVO La educación formal colombiana establece tres niveles institucionales de: Básica primaria, Básica secundaria y Media.

GRADO Cada uno de los grados de estudio, la educación formal colombiana define once grados enumerados del 1 al 11, repartidos entre los tres niveles, así:

- Básica Primaria: Ciclo que consta de cinco grados, cada uno de un año de duración, nombrados de primero a quinto grado, ingresando hacia los 6 años.
- Básica Secundaria: Ciclo que consta de cuatro grados, cada uno de un año de duración, nombrados de sexto a noveno, ingresando usualmente a los 11 o 12 años.
- Educación media: Ciclo que comprende los dos últimos grados, décimo y undécimo.

CURSO Cada uno de los grupos correspondientes a cada grado, identificados generalmente con las letras A, B, C, etc. por ejemplo 9A y 9B para representar dos cursos del grado 9.

AULA Espacio físico donde transcurre cada curso.

CARGA Relación de las materias que se encargan a cada profesor en un año lectivo

PROFESOR Instructor encargado de enseñar las materias correspondientes a cada grado en cada curso.

ÁREA Agrupación de materias por temas fundamentales y obligatorios definidas por la norma colombiana, son:

- a) Ciencias naturales y educación ambiental.
- b) Ciencias sociales, historia, geografía, constitución política y democracia.
- c) Educación artística.
- d) Educación ética y en valores humanos.
- e) Educación física, recreación y deportes.
- f) Educación religiosa.
- g) Humanidades, lengua castellana e idiomas extranjeros.
- h) Matemáticas.
- i) Tecnología e informática.

MATERIA Asignatura objeto de estudio en los distintos grados, que está contenida en un área específica.

ALUMNO Estudiante asignado a cada curso de un grado.

EVALUACIÓN Prueba o examen en el que se revisa el rendimiento de cada alumno en cada materia dentro de un curso.

NOTA Calificación final que permite determinar si el alumno será promovido o no al siguiente grado, se calcula con base en las evaluaciones.

7.1 Diagrama o esquema gráfico del MER

El Modelo Entidad Relación se representa esquemáticamente en un diagrama como un grafo dirigido en donde los vértices o nodos son las entidades y las aristas son las relaciones.

Las entidades se dibujan mediante cajas de bordes redondeados, dentro de las cuales se coloca el nombre de la entidad. El tamaño y disposición de la caja que representa la entidad dentro del diagrama son arbitrarios, requiriéndose que posean suficiente espacio para colocar su nombre.

Es posible que haya sinónimos válidos de uso equivalentes en el dominio de aplicación analizado. Se debe escoger un nombre primario; los sinónimos se pueden mostrar en mayúsculas, precedidos de una barra inclinada / (diagonal), por ejemplo, alumno o estudiante son términos válidos para la misma entidad.

La figura 3 presenta el esquema gráfico del modelo, hecho con notación Barker.

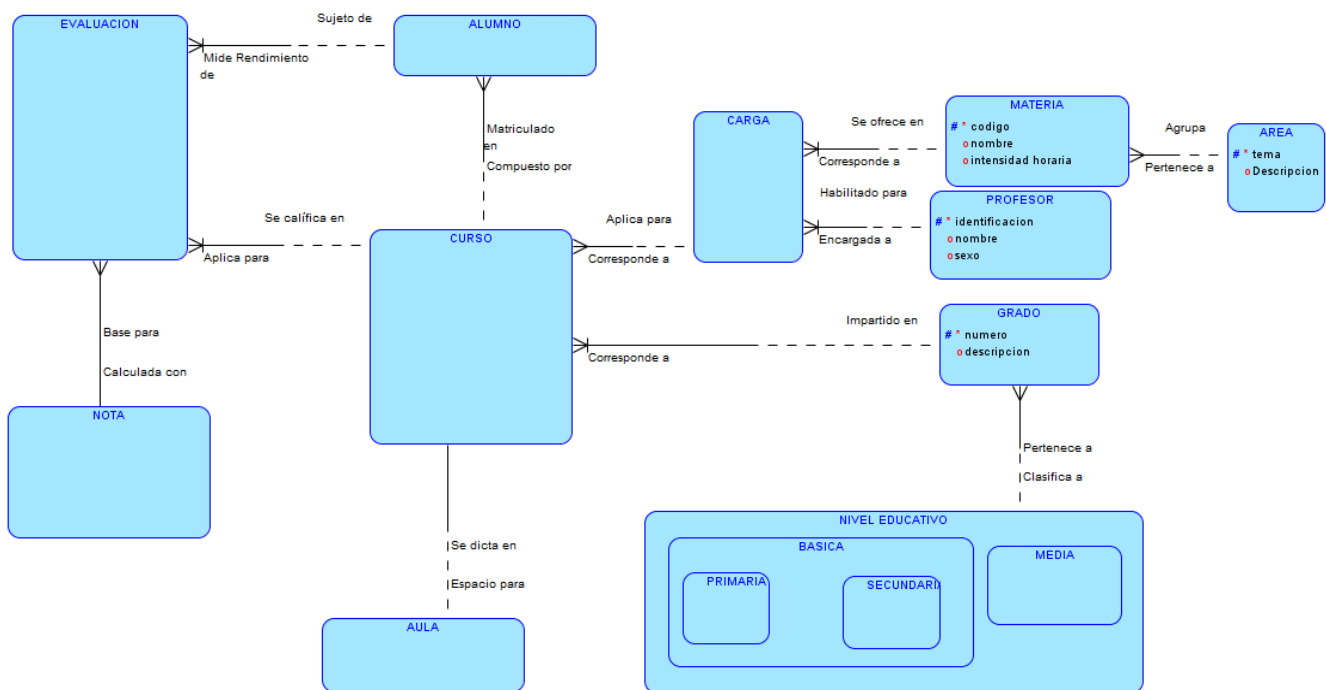


Ilustración 4. Diagrama o representación gráfica del modelo

8 BIBLIOGRAFÍA

- [1] C. W. Bachman, «Data structure diagrams,» de *Data Base*, 1969.
- [2] E. F. Codd, «A relational model of data for large shared data bank,» *Communications of the ACM*, vol. 13, nº 6, pp. 377 - 387, 1970.
- [3] P. P.-S. Chen, «The entity-relationship model—toward a unified view of data,» *ACM Transactions on Database Systems (TODS)*, vol. 1, nº 1, pp. 9 -- 36, 1976.
- [4] M. Cohn, «User Stories Applied: For Agile Software Development,» 2004.
- [5] R. Singh, «International Standard ISO/IEC 12207 software life cycle processes,» *Software Process Improvement and Practice*, pp. 35-50, 1996.
- [6] Team, CMMI Product, «Capability Maturity Model Integration (CMMI), Version 1.1--Continuous Representation,» 2002.
- [7] R. V. a. L. C. Y. O'Connor, «Software project management in very small entities with ISO/IEC 29110,» de *European Conference on Software Process Improvement*, 2012.