



<https://eva.fing.edu.uy/course/view.php?id=536#section-2>

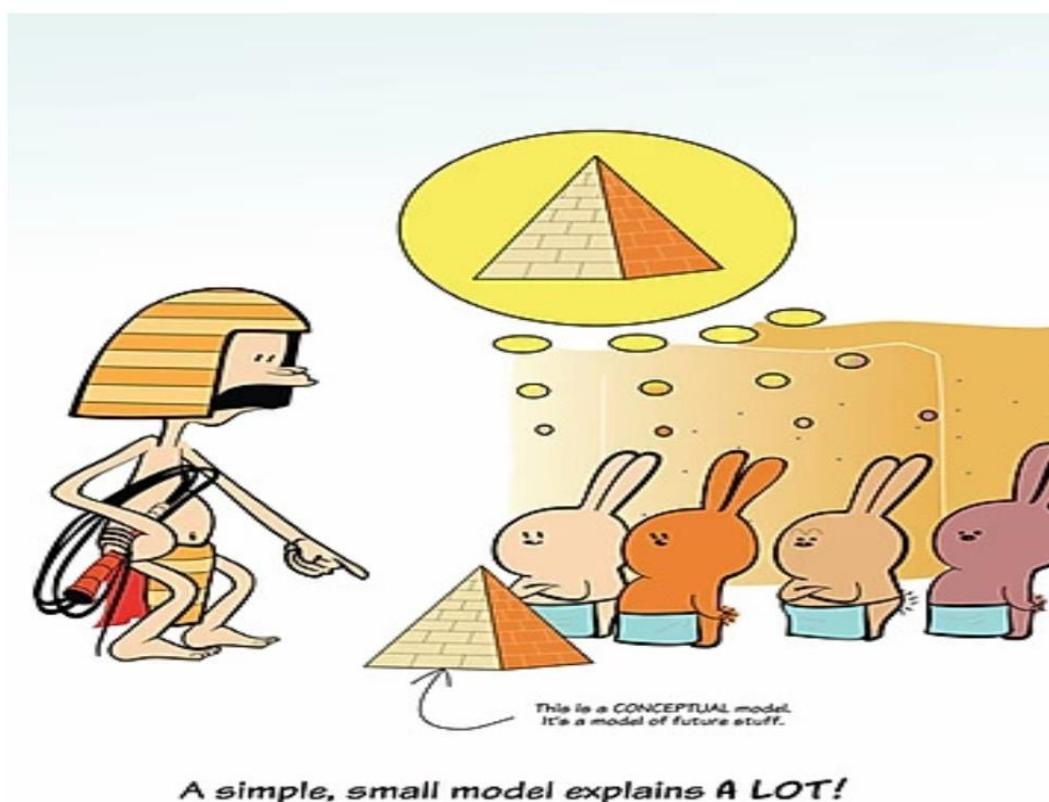
Regina Motz

InCo- Fing - Universidad de la República

Objetivos de la Unidad

Introducir los diferentes modelos de datos que permiten representar y proveer de semántica escenarios del mundo real.

Entender que estos modelos tienen diferente grado de expresividad y analizar los que forman parte de la Arquitectura de la Web Semántica, con especial atención a XML-Schema.



I. Modelos de Datos

(Recorrer la lección del eva para responder los cuestionarios de actividad individual)

El conocimiento que tenemos de una realidad lo podemos representar utilizando un MODELO. Contar con un modelo nos permite describir una situación de forma tal que podemos estudiar sus propiedades o sus reglas de comportamiento y explicarlas para que otras personas puedan comprenderlas. Algunos modelos pueden servir para simular el comportamiento de una situación y hacer predicciones, como los modelos matemáticos de propagación del COVID-19 por ejemplo.

Los modelos también son usados para representar una situación con un grado específico de detalle. Las diferencias en los grados de detalle representados por el modelo dependen del punto de vista

desde el cual se quiera estudiar la situación. Según las diferencias en el grado de detalle que representan se habla de modelos con distinto nivel de abstracción. El nivel de abstracción no está relacionado con la capacidad de información que un modelo puede representar sino en qué características de la realidad se desea focalizar.

Algunos modelos con los que trabajamos en sistemas de información son:

- modelos conceptuales
- modelos de flujo de datos
- modelos lógicos
- modelos físicos

Cada uno de estos modelos tiene por objetivo explicar la realidad con foco en resolver un aspecto específico de la misma.

El objetivo principal de un modelo conceptual es transmitir los principios fundamentales y la funcionalidad básica del sistema que representa, tratando de responder a la pregunta de

¿Qué es la realidad?

Mientras que los modelos de flujo de datos tratan de responder a la pregunta de

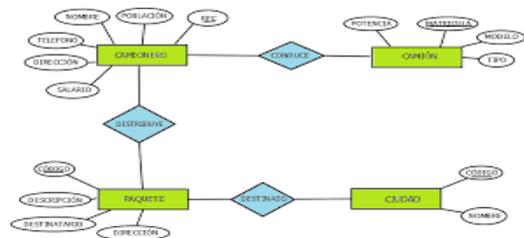
¿Cómo funciona la realidad?

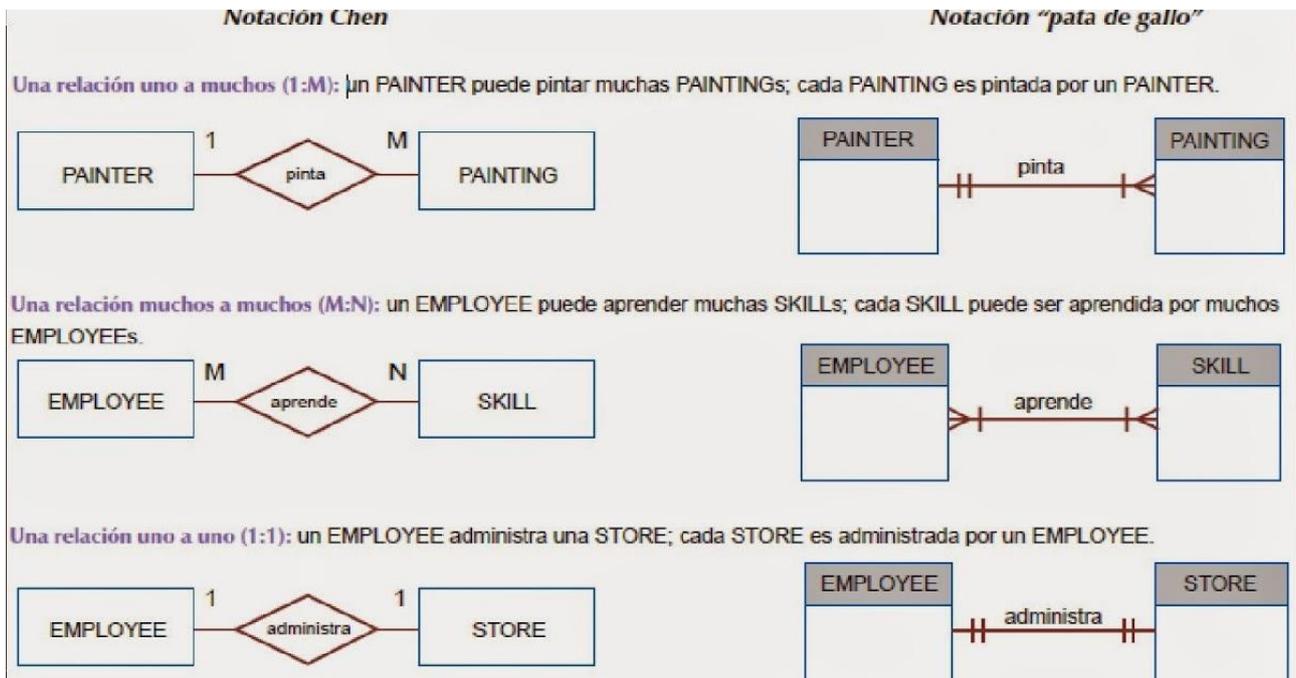
En la siguiente sección tomemos como ejemplo las Bases de Datos:

CASO BASES DE DATOS

Si nos fijamos en los modelos que usamos para diseñar, desarrollar e implementar una base de datos encontramos los siguientes tres niveles [1]:

1. Para la fase de diseño: un Modelo Conceptual, como el Modelo Entidad-Relación (MER) que nos ayuda a representar los elementos que entendemos relevantes de la realidad.





Referencias:

[1] "Fundamentals of Database Systems 5th Ed.", Elmasri & Navathe, Addison Wesley, 2007. (Capítulo 3) .

-----Actividad Individual en el EVA-----

Realiza las preguntas que encuentras en la Lección: Modelos de Datos en el EVA.

Continuando con modelos en general

En bases de datos los modelos que se utilizan se clasifican según su nivel de abstracción pero cuando queremos capturar la semántica de una realidad nos interesa clasificar a los modelos por su capacidad de expresión. Es decir, un modelo con mayor nivel de abstracción evita tener que pensar en detalles de cómo representar aspectos tales como un cierto tipo de relación, por ejemplo la relación jerárquica, ya que provee un constructor que lo resuelve, por ejemplo los modelos orientados a objetos lo proveen. Sin embargo, la capacidad de expresión tiene que ver con poder o no representar ese aspecto, por ejemplo poder representar jerarquías.

Encontramos modelos simples para entender una realidad, que permiten describir de forma ordenada los elementos que componen el fragmento de la realidad en estudio.

Por ejemplo, si consideramos una farmacia es preferible tener un catálogo o un vocabulario controlado de los productos que comercializa que sólo nombrar las marcas que vende. La Figura 1 nos muestra un ejemplo de algunos de los productos que comercializa una farmacia, usando un modelo de catálogo con códigos numéricos y usando una taxonomía de términos de un vocabulario controlado.



CÓDIGOS

productos de limpieza	01
Alcohol líquido	01011
Alcohol en gel	01012
Cloro	0102
Hipoclorito	0103

Vocabulario controlado

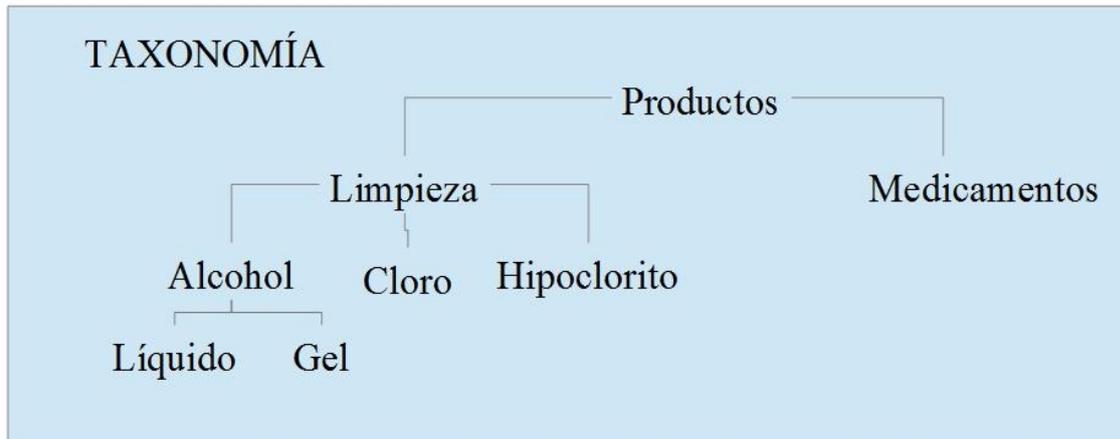


Figura 1: Ejemplo de modelado con catálogos y taxonomías.

Si se usara un tesauro, tendríamos una taxonomía expandida donde se representan también términos sinónimos y se distingue el término más utilizado. También se agregan términos referentes o los más utilizados. Por ejemplo se acostumbra a veces llamar al producto por el nombre de la marca más conocida y ese nombre se agrega entonces como nombre referente al término en que se asocia. También se pueden incluir las traducciones del término a varios idiomas.

Pero a pesar que los tesauros representan un mayor número de relaciones que las taxonomías, observamos que los tesauros no permiten agregar relaciones entre los términos que no sean los ya definidos, ver por ejemplo WordNet.

Para expresar la complejidad de una situación de la realidad es necesario poder definir relaciones entre los conceptos de forma flexible. Esto se puede hacer mejor con los modelos de datos como el MER que permiten definir cualquier tipo de relaciones binarias entre conceptos. Se clasifican entonces el MER con mayor nivel de expresividad que el tesauro.

Los modelos de datos también nos permiten definir que una relación puede tener la propiedad de ser función. Si estamos trabajando con un modelo orientado a objetos tenemos además la propiedad de herencia en las relaciones de jerarquía.

Si además de la propiedad de ser función o de herencia queremos definir una relación con otras propiedades, como tener la propiedad de ser transitiva por ejemplo, no lo podemos hacer ni con el MER ni con el Modelo Relacional ni el orientado a objetos. Pero sí lo podemos hacer con las Ontologías pues presentan un espectro de propiedades de relaciones entre conceptos más amplio que los modelos de datos. Esta es una de las razones por las que se ubica a las ontologías como más expresivas que los modelos de datos.

La Figura 2 muestra una definición coloquial para cada uno de estos modelos junto a su clasificación según el grado de expresividad.

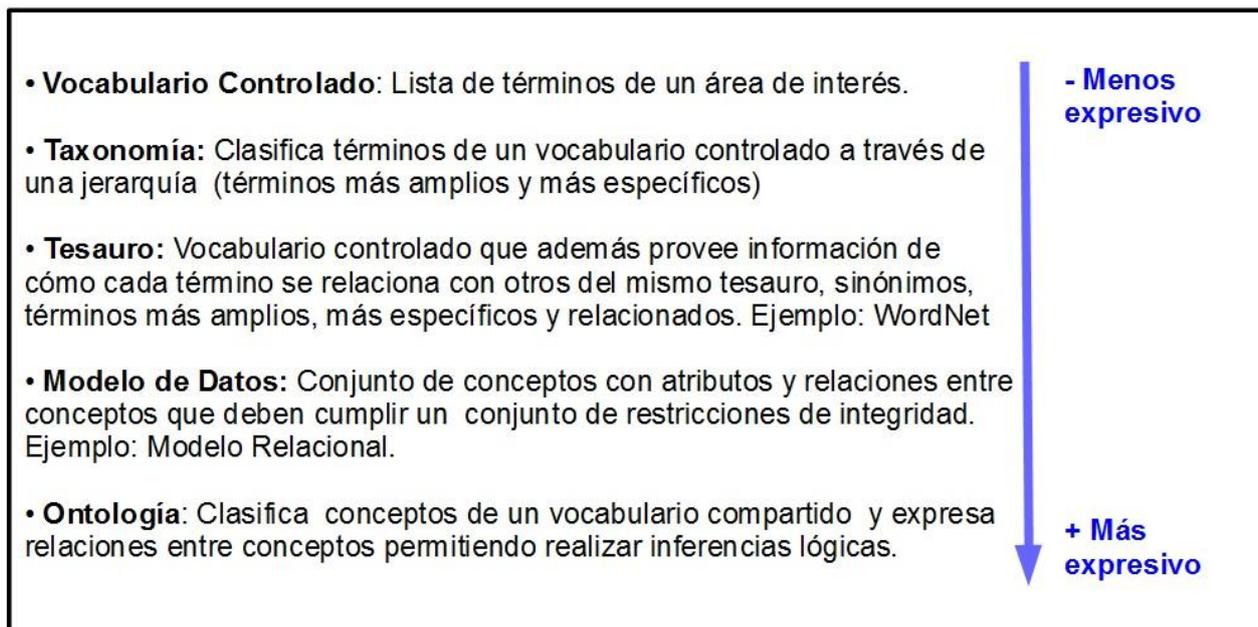


Figura 2: Niveles de expresividad en distintos modelos.

¿Qué significa que las ontologías nos permitan hacer inferencias lógicas?

Para contestar esta pregunta mira el Ejemplo-Animales en la Lección del EVA. <Ejemplo-Animales>

-----**Actividad Individual en el EVA**-----

Realiza las preguntas que encuentras en la Lección Modelo de Datos en el EVA.

¿Qué relación tienen los modelos con los metadatos?

---> a los metadatos también necesitamos representarlos y para eso usaremos distintos modelos.

Directamente asociado al nivel de expresividad de los modelos se los clasifica también respecto a las posibilidades que brindan para realizar integraciones de datos, desde sintácticas, estructurales o semánticas. Esto se puede observar en la Figura 3 con el cuadro que terminábamos la Unidad 2.

Espectro de Metadata

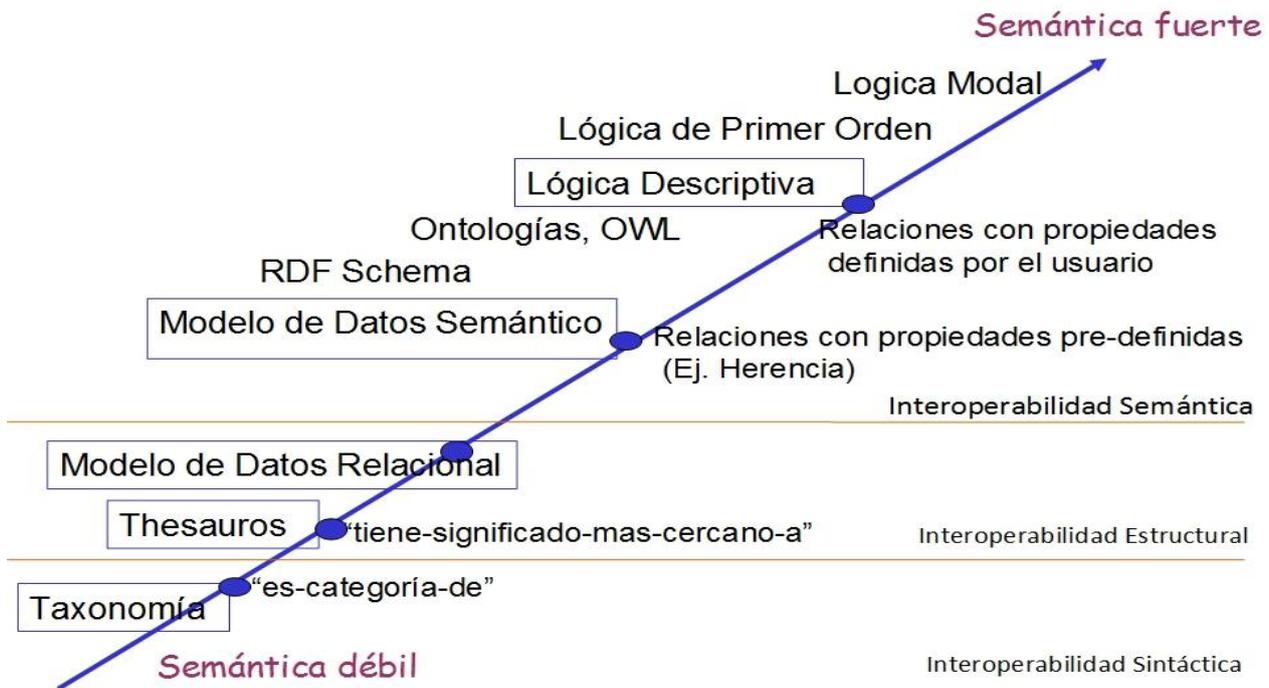


Figura 3. Espectro de metadatos según su capacidad para apoyar la integración de datos

La bibliotecología ha sido pionera en el uso de metadatos, mira el video: "De las bibliotecas a la web semántica" que encuentras al final de la lección, el video muestra a través de un ejemplo la relación existente entre los modelos usados para organizar, recuperar e integrar información de las bibliotecas con relación a lo que implementa la web semántica.

En la siguiente parte de la Unidad analizaremos el papel que cumple XML en la arquitectura de la Web Semántica. Lo encuentras en archivo Material-XML.pdf en el EVA de la Unidad 3.

Antes de continuar mira la animación "De las bibliotecas a la web semántica", la encuentras en el EVA.