

Correspondencias entre taxonomías XBRL y ontologías en OWL

Unai Aguilera, Joseba Abaitua

Universidad de Deusto, EmergiaTech

Resumen

Todo documento XBRL contiene cierta información semántica que se representa mediante XML Schema y Xlink. Sin embargo estos mecanismos no fueron creados con el objetivo de dar cuenta de información de este tipo. Por ello pueden plantearse correspondencias entre taxonomías XBRL y ontologías de conceptos en OWL. De esta forma, la información contenida en un documento XBRL podrá relacionarse con conceptos de otras ontologías favoreciendo el proceso de razonamiento. Se han detectado, sin embargo, algunos impedimentos de OWL, como el de incluir datos XML Schema creados por el usuario. Este y otros problemas, así como sus posibles soluciones son objeto de análisis en este documento.

Introducción

La adopción de XBRL [1] facilita el intercambio de información financiera entre entidades así como su tratamiento y almacenamiento posterior. Esta tecnología, basada en XML [2], se fundamenta en la utilización de taxonomías comunes y compartidas de conceptos financieros. De esta forma, las entidades que utilizan XBRL pueden generar informes cuyos datos se anotan y estructuran en un sistema conceptual común. Gracias al uso de estas taxonomías consensuadas, los datos pueden ser etiquetados de manera que se explicita su significado exacto. Esto permite que posteriormente, cuando el documento sea interpretado, pueda obtenerse el significado de cada uno de los datos de una forma precisa, a partir de los conceptos establecidos anteriormente en la taxonomía. Este tipo de anotaciones es de facto una anotación semántica. Es posible, por ello, la creación de aplicaciones informáticas que extraigan la información contenida en documentos XBRL y razonen sobre ella, teniendo en cuenta el significado e interrelaciones de cada uno de los datos.

En XBRL las taxonomías se crean utilizando XML Schema [3] y XLink [4]. La tecnología XML Schema fue creado para resolver algunas carencias de los documentos DTD a la hora de definir la estructura y tipos de datos para los documentos XML. La tecnología XLink por su parte permite crear referencias entre recursos. La suma de estas dos tecnologías permite establecer las relaciones entre conceptos que son necesarias para construir las taxonomías.

Sin embargo, aunque en un documento XBRL existe una cierta carga de información semántica, sus posibilidades se encuentran muy limitadas por la tecnología utilizada para representarla. Hay que tener en cuenta que ni XML Schema ni XLink fueron creados con el objetivo de representar este tipo de información. Pese a que, como es el caso de XBRL, se pueden utilizar con esta finalidad, la expresividad que proporcionan es insuficiente para capturar toda la carga semántica de los datos representados. Esta falta de expresividad repercute en la capacidad de razonamiento de las aplicaciones que utilicen los documentos XBRL. Efectivamente, pueden existir aplicaciones que procesen documentos XBRL como datos de entrada que, además de razonar con conceptos relativos al ámbito financiero, necesiten relacionar estos conceptos con otros procedentes de dominios diferentes.

En la actualidad, con la aparición de la web semántica, se está tendiendo a la utilización de lenguajes creados explícitamente para la representación de dominios conceptuales, en forma de ontologías (el ejemplo más destacado es OWL [5]). Con ellos se obtiene la expresividad y facilidad para la representación de los conceptos de las que carecen las taxonomías. Por otro lado, estos lenguajes favorecen la reusabilidad de las ontologías creadas. Cualquier ontología creada anteriormente puede ser referenciada y reutilizada desde otra ontología. De esta manera, además de permitir reducir el tiempo de desarrollo de las nuevas ontologías, se consigue que los conceptos más comunes mantengan una interpretación uniforme en distintas ontologías.

Ontologías complementarias con XBRL

En este documento se analiza una posible correspondencia entre taxonomías XBRL y otras ontologías en OWL. Con este fin se extraerá la información conceptual de las taxonomías XBRL y se creará una ontología que represente la misma información semántica. Con ello no se pretende modificar el estándar XBRL, sino añadir una capa semántica que permita que la taxonomía (y su derivable ontología) de conceptos financieros pueda ser relacionada y utilizada junto a otras ontologías de otros dominios de conocimiento, ampliando la cobertura conceptual de la misma y mejorando significativamente las posibilidades de razonamiento semántico de las aplicaciones que la utilicen.

Este procedimiento no sólo beneficiará a las aplicaciones financieras que recurran a las ontologías XBRL ampliadas, sino también a otras aplicaciones semánticas, como agentes inteligentes de un uso más general, diseñadas para el tratamiento de la información en el lenguaje que manejan habitualmente.

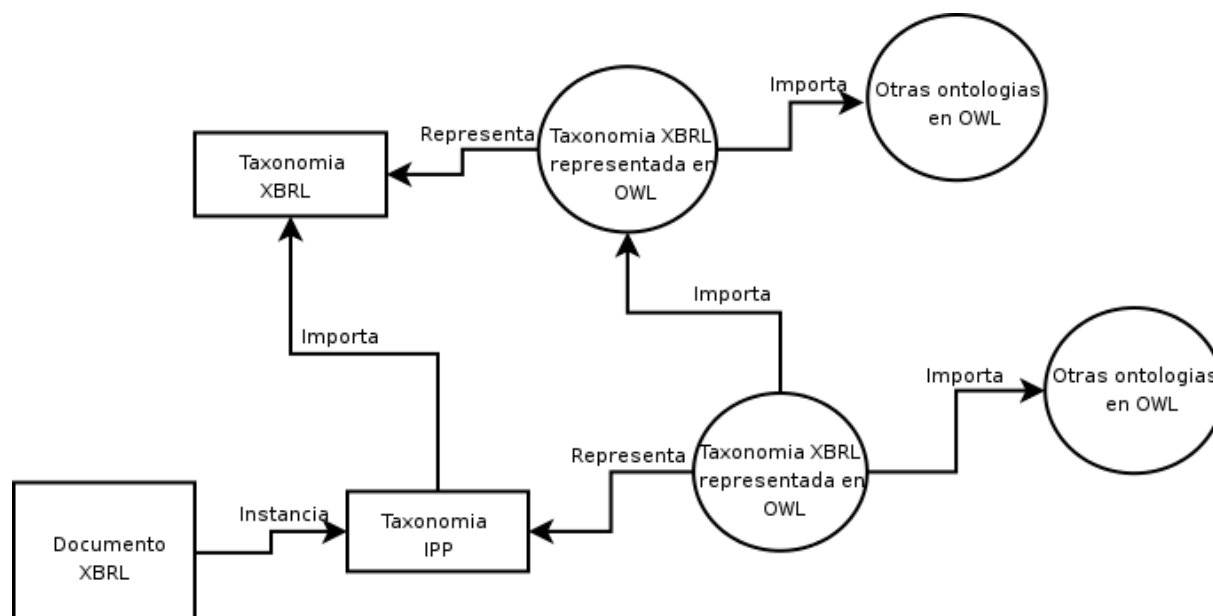


Figura 1: Diagrama Extensión Semántica XBRL

En la Figura 1 puede observarse el diagrama de la solución propuesta. Un documento XBRL se crea mediante una instancia de los conceptos representados en una o varias taxonomías. En el caso del ejemplo presentado, el documento XBRL es una instancia de la taxonomía IPP que a su vez utiliza

los conceptos genéricos (tipos y relaciones) definidos en la taxonomía común de XBRL. Esta última es una taxonomía genérica que define conceptos comunes utilizados por todas las taxonomías específicas definidas dentro de XBRL. Estas taxonomías son por lo tanto nuevamente representadas utilizando en esta ocasión un lenguaje específico para la creación de ontologías, como puede ser OWL, que presenta las ventajas comentadas anteriormente. Así, estas ontologías sirven de puente de unión entre los conceptos del dominio financiero, capturados en las taxonomías XBRL, y otros conceptos contenidos en otras ontologías procedentes de otros dominios. De esta forma será posible aprovechar todo el potencial que puede proporcionar la web semántica sin necesidad de modificar el estándar XBRL.

Correspondencia entre XBRL y OWL

Se han estudiado las características de XBRL con el propósito de realizar la correspondencia entre la taxonomía XBRL y una ontología descrita en lenguaje OWL. Se detalla a continuación la forma ideada para representar en OWL las relaciones de la taxonomía XBRL:

- **Linkbase de Definición:** Establece relaciones entre los elementos de la taxonomía (generalización/especialización, equivalencias, etc.). La información semántica que proporcionan estas referencias entre conceptos pueden ser representadas mediante el uso de construcciones existentes para tal fin en OWL. OWL permite especificar relaciones de concepto/subconcepto, equivalencias, uniones de conceptos y otras propias de las ontologías del lenguaje OWL, como pueden ser las propiedades y las restricciones.
- **Linkbase de Etiquetas:** Con ellas se especifican las etiquetas que se utilizarán en un determinado idioma para representar cada elemento (concepto) de la taxonomía. En OWL existe un conjunto de etiquetas que soportan el multilingüismo.
- **Linkbase de Cálculo:** Permite especificar relaciones algebraicas entre elementos de la taxonomía (el elemento A es la suma de B y C). Actualmente sólo permite realizar sumas y restas, pero se espera que incluya operaciones más complejas en revisiones posteriores del estándar.
OWL no proporciona directamente un lenguaje para representar relaciones matemáticas entre los conceptos. Pero debido a que esta información es de tipo semántica, podrá ser representada mediante la creación de una ontología en OWL (o la utilización de una ontología ya existente que represente las operaciones matemáticas posibles). De esa forma, relacionando las operaciones matemáticas con los conceptos financieros se podrá indicar que un concepto se obtiene como resultado de una operación matemática a partir de otros.
- **Linkbase de Referencia:** Son referencias a documentos externos. Esta información puede ser representada en OWL a través de una propiedad cuyos valores posibles sean URIs o identificadores de un archivo de documentos.
- **Linkbase de Presentación:** Es información relativa a cómo debe mostrarse el informe en las aplicaciones de visualización.
El lenguaje OWL no introduce ninguna construcción para representar la visualización de los conceptos de la ontología. Si esta información, que sí puede estar presente en un documento XBRL, quiere representarse en OWL, deberá construirse una estructura que permita su representación en OWL mediante el uso de clases y relaciones entre ellas.

Tipos de datos

XBRL define nuevos tipos para los datos que pueden aparecer en un documento. Así se introducen nuevos tipos de datos más complejos creados a partir de los tipos básicos incluidos en el estándar XML Schema. Por ejemplo, se introduce el tipo *nonZeroDecimal* a partir del tipo *Decimal*, definido en XML Schema, y la introducción de varias restricciones que impiden que cualquier dato de este tipo pueda tomar el valor “0”.

Por lo tanto, es necesario que estas restricciones en los tipos de datos sean representados guardando su significado en la ontología creada. La forma más sencilla y directa de realizarlo será referenciar desde la ontología en OWL los tipos de datos definidos en el fichero *XBRL-instance.xsd*. Sin embargo, aunque OWL permite trabajar con los tipos de datos predefinidos en XML Schema, con el fin de especificar el tipo de las propiedades de las clases, sólo permite trabajar con estos tipos predefinidos y no con tipos complejos definidos por el usuario [6]. Hasta la fecha, en el estándar de OWL no existe una forma clara de introducir referencias a los tipos de datos creados por el usuario. Existen diferentes propuestas para incluir esta posibilidad en el estándar pero ninguna de ellas ha sido elegida aún. Para solucionar este problema, y poder utilizar estos tipos de datos dentro de las ontologías en OWL, se propone crear una ontología que los represente. Aunque esta ontología recoge gran parte de la información contenida en el XML Schema, por la restricción comentada anteriormente, no se ha llegado todavía a representar toda la información referente a los tipos de datos.

A modo de ejemplo se ha creado una ontología de tipos de datos. Una representación visual de esta ontología puede verse en la Figura 2.



Figura 2: Ontología de Tipos de Datos

Los tipos de datos de XML Schema han sido representados de la siguiente forma:

- Las relaciones de especialización se han representado directamente mediante la estructura que proporciona OWL para representar la especialización/generalización de conceptos. Así, en la ontología existe una jerarquía de tipos de datos que representa la jerarquía existente en

conceptos heredan los atributos correspondientes a un elemento XBRL y los atributos correspondientes a su tipo específico. Por ejemplo, el concepto financiero “Inmovilizado” además de ser del tipo “XBRLElement” es del tipo “MonetaryItemType” heredando por lo tanto los atributos correspondientes.

Conclusión

En este trabajo se ha pretendido mostrar como se puede interconectar la taxonomía XBRL con ontologías creadas en OWL. Aunque, como se ha visto, parte de la información contenida en la taxonomía puede transformarse en una ontología OWL de una forma más o menos directa, existen determinados aspectos que deben ser estudiados en profundidad para elegir la mejor forma de representarlos. Gran parte de estos problemas se resolverán cuando se introduzca en OWL la posibilidad de utilizar tipos de datos XML Schema creados por el usuario. Hasta ese momento, los tipos de datos complejos deben representarse mediante una ontología OWL que no es la más adecuada para representar todas las restricciones de datos que se pueden representar en XML Schema.

A pesar de estos problemas, que se espera que sean solucionados en un futuro, la representación de la taxonomía XBRL en un lenguaje de ontologías proporcionará varias ventajas con respecto a la situación actual, ampliando el dominio de razonamiento de las aplicaciones a dominios fuera del ámbito financiero y permitiendo aprovechar todas las posibilidades de la web semántica.

Bibliografía

<http://www.xbrl.org>
<http://www.w3.org/XML>
<http://www.w3.org/XML/Schema>
<http://www.w3.org/TR/xlink/>
<http://www.w3.org/2004/OWL/>
<http://www.w3.org/TR/swbp-xsch-datatypes/>