

La representación de relaciones conceptuales en una ontología

Amparo Alcina

Universitat Jaume I

alcina@uji.es

Resumen

Las bases de datos terminológicas no cuentan con una estructura eficaz para representar los conceptos y sus relaciones. Presentamos cómo hemos adaptado las ontologías para desarrollar análisis detallado de los conceptos y sus relaciones, que está en la base de la elaboración de definiciones terminológicas. Mostramos el análisis conceptual, formalización y su implementación. Además de las ventajas de la gestión, estos sistemas facilitan también la consistencia y el razonamiento con los datos. El uso de ontologías en terminología abre el camino hacia los datos enlazados e intercambio de datos terminológicos en el nuevo paradigma de la Web Semántica.

Palabras clave

Terminología, tecnología, concepto, ontología, definición, Web semántica.

Abstract

The terminology databases do not have an effective structure to represent the concepts and their relations. We present how we have adapted ontologies to develop detailed analysis of the concepts and their relationships, which is at the basis of the elaboration of terminological definitions. We show the conceptual analysis, formalization and its implementation. In addition to the advantages of management, these systems also facilitate consistency and reasoning with data. The use of ontologies in terminology opens the way to linked data and exchange of terminological data in the new paradigm of the Semantic Web.

Keywords

Terminology, Technology, Concept, Ontology, Definition, Semantic Web.

0. Introducción¹

La terminología es sin duda uno de los grandes quebraderos de cabeza en la actividad traductora. Conocer el significado de los términos que aparecen en los textos de especialidad, encontrar su equivalencia en la lengua meta entre los expertos en un dominio, aprender cómo se combinan en el texto y cuáles son las colocaciones que suenan más naturales son algunas de las inquietudes que se manifiestan en el proceso de traducción.

¹ Esta investigación forma parte del proyecto: "PRO-ONTODIC: Protocolos para la creación de diccionarios terminológicos basados en ontologías (Modelo ONTODIC)" (Ref. UJI-B2018-65), financiado por el Plan de investigación 2018 de la Universitat Jaume I.

La actividad terminológica aplicada a la traducción comprende la recopilación de los términos, la descripción de sus aspectos lingüísticos (fonológicos, morfológicos, morfosintácticos y de uso), la redacción de su definición y la ubicación del concepto al que representa en una estructura conceptual del dominio de especialidad. Los aspectos relacionados con el significado, la elaboración de la definición y la ubicación del término en una estructura conceptual, requieren a su vez explicitar con detalle las características que definen el concepto y las relaciones que lo vinculan con otros conceptos.

En terminología, se emplea un gran esfuerzo para extraer las características y relaciones conceptuales, recopilarlas y asignarlas a los conceptos, elaborar árboles de campo, organizar grupos de conceptos y trabajar sistemáticamente la información para redactar definiciones de términos correctas desde el punto de vista de su contenido y de las normas lexicográficas. Pocas veces quedan esos datos de trabajo (relaciones conceptuales, árbol de campo, especificación de características) plasmados en los diccionarios, los recursos lexicográficos o terminológicos impresos o electrónicos. Las definiciones, redactadas en lenguaje natural y destinadas a su lectura solo por humanos, no muestran los elementos del proceso de trabajo. Como afirmaba Sager:

Dentro de la teoría terminológica se acepta que los conceptos deberían ordenarse según ciertos esquemas de clasificación conceptuales y presentarse dentro de una estructura sistemática. Para lograr esto, se caracteriza a los conceptos mediante las relaciones que forman con sus conceptos colindantes. [...] Sin embargo, pese a que percibimos enseguida un gran número de relaciones entre los términos, muy pocas se reflejan en la estructura de los glosarios. (Sager, 1990:54)

El cambio del formato impreso al uso de las bases terminológicas mejoró la flexibilidad en el acceso a los datos que incluyen los diccionarios, pero la información sobre el significado sigue sin aparecer desglosada. Los sistemas de bases de datos terminológicos, basados en estructuras de campos y registros, resultan insuficientes. Se requieren formas más flexibles de estructurar la información.

Se comenzaron a aprovechar los sistemas de bases de conocimiento desarrollados en Ingeniería del conocimiento, luego llamadas *ontologías*, para realizar el análisis conceptual detallado de los términos.

En nuestra línea de investigación ONTODIC, sobre Metodología y técnicas para la elaboración de diccionarios terminológicos, buscamos desarrollar sistemas más flexibles y potentes, que hagan compatible tanto las necesidades del terminólogo y el traductor, como las necesidades actuales del procesamiento automático del lenguaje natural. El trabajo consistió en la estructuración del árbol de campo (Alcina, 2009), la búsqueda de relaciones conceptuales y características en el dominio (Maroto, 2007) y la elaboración de plantillas de definición para algunos grupos de términos (Alcina y Valero, 2008; Valero y Alcina, 2009; 2010; 2015). En las primeras fases del proyecto, las relaciones conceptuales se implementaron en el programa Protégé-Frames (Maroto, 2007); más adelante, con la evolución de la herramienta, se usó Protégé-OWL para la implementación de características (Estellés

Palanca, 2014) y para la implementación de definiciones (Alcina y Valero Doménech, 2017).

En este trabajo presentamos, en primer lugar, por qué las bases de datos terminológicas no pueden ayudar en la gestión del análisis de conceptos y la necesidad de usar bases de conocimiento y ontologías. En segundo lugar, explicamos en cómo desarrollamos el análisis de conceptos y su formalización orientada a su gestión en una ontología. Por último, presentamos cómo hemos implementado ese análisis y formalización de conceptos en el editor de ontologías Protégé. Concluiremos con una reflexión sobre las ventajas que nos aporta el uso de este programa en la gestión conceptual.

1. De las bases de datos terminológicas a las ontologías

Desde los inicios de la informática de usuario, los traductores encontraron en las bases de datos terminológicas un aliado para realizar las tareas de almacenamiento, consulta e intercambio de datos terminológicos, frente a los sistemas en papel que se habían estado usando (Alcina, 1997).

1.1. Las limitaciones de las bases de datos terminológicas

Los programas de gestión de bases de datos terminológicas permiten, a menudo, configurar la estructura de campos de manera flexible, de modo que el usuario diseña la configuración que le resulta más adecuada a sus fines. Algunos de estos datos son la denominación del término, su categoría gramatical, su definición, uno o varios contextos de uso, el dominio de especialidad al que pertenece, equivalencias en una o varias lenguas, sinónimos, imagen. También se pueden añadir campos que indiquen el dominio de especialidad al que pertenece el término, e incluso la rama y/o subrama del árbol de campo al que se ha asignado. Cada uno de esos datos se recoge en campos independientes. Las bases de datos terminológicas orientadas al concepto presentan, en cada ficha terminológica, la información relacionada con los términos que designan un mismo concepto de una o varias lenguas. En la Figura 1 se muestra de forma esquemática la información de una ficha en la base de datos.

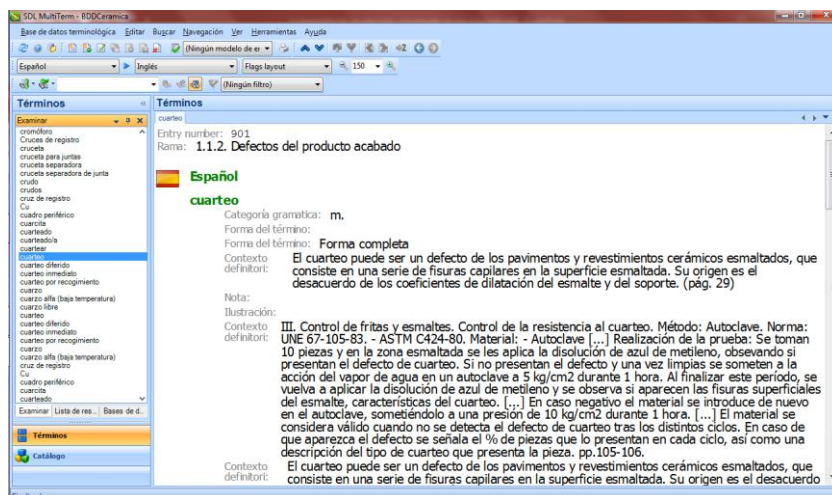


Figura 1. Ejemplo de ficha en una base de datos terminológica

El sistema de navegación en las bases de datos suele consistir en: un panel mostrando los términos ordenados alfabéticamente disponible en uno o más idiomas, sistemas de búsqueda mediante la forma plena, formas truncadas, uso de comodines, uso de filtros (que permiten seleccionar un término o un conjunto de términos a partir de la delimitación de criterios en alguno de los campos) y referencias cruzadas (que permiten saltar de una ficha a otra pinchando en alguno de los enlaces —términos, fuente u otros elementos— que aparecen en la ficha). Para una clasificación completa de las formas de búsqueda en diccionarios electrónicos puede consultarse (Pastor y Alcina, 2010).

Por otra parte, las bases de datos terminológicas apenas ofrecen funciones relacionadas con los conceptos, ni la suficiente flexibilidad para gestionar de forma detallada y cómodamente esa clase de información. En relación con la información conceptual, una ficha terminológica incluirá campos como rama y subrama del dominio, definición del término o términos relacionados. Sin embargo, el sistema no ofrece una gestión especial de estos campos como sería deseable.

En primer lugar, por ejemplo, no existe una función que permita listar los términos organizados por ramas y subramas. La información sobre la rama y subrama del dominio aparecen en cada ficha de forma independiente, sin posibilidad de acceder al mapa conceptual completo del dominio o de mostrar de forma conjunta los términos clasificados en cada rama, o de navegar por el sistema de conceptos así organizados. En segundo lugar, las características y relaciones conceptuales del término están contenidas implícitamente en la definición, pero no existe la posibilidad de que sean consultadas o manejadas de forma individual o independiente de cada ficha. No hay opciones de visualización que permitan comparar las características conceptuales y las definiciones de los distintos términos, o de listar los términos vinculados por un tipo de relación conceptual en particular.

Ante las limitaciones de las bases de datos terminológicas para procesar la información conceptual de los términos, por una parte, y la toma de conciencia en el ámbito de la Traducción de la necesidad de los traductores de acceder a las estructuras de conocimiento de los dominios, por otra parte, se comenzó a considerar la necesidad de usar otro tipo de recursos que facilitaran el acceso y la gestión de la información conceptual.

En la década de los 90, Meyer introduce las bases de conocimiento apoyándose en la incipiente disciplina de la Ingeniería del conocimiento. Su investigación da lugar a herramientas como CODE, IKARUS o Dockman, que permitían navegar a través de sistemas de conceptos y sus relaciones (Meyer y otros, 1997). Desde entonces, diversos grupos de investigación que desarrollan aspectos conceptuales del término han desarrollado herramientas para crear bases de conocimiento terminológico, como Caos (Madsen y Thomsen, 2009a; b) y Ontoterm (Faber, 1999; 2002; Moreno Ortiz, 2002).

Recientemente, estas bases de conocimiento empiezan a difundirse en la Web, como el innovador proyecto EcoLexicon. EcoLexicon es una base de conocimiento terminológica sobre medioambiente, fundamentada en la Terminología basada en marcos y desarrollada por el grupo Lexicon de la Universidad de Granada (León-Araúz y otros, 2019).

Más allá de los desarrollos para la terminología, la Ingeniería del conocimiento ha evolucionado en sus fundamentos, las herramientas y los estándares de datos. Aunque se retroalimentan, la comunicación entre una y otra disciplina no termina de ser fluida y, en general, van por caminos muy distintos. Pero resulta necesario, para poder beneficiarse de los nuevos avances, conocer cuál es la evolución de esta disciplina y por qué se desarrolla de esta forma.

1.2. Las bases de conocimiento y ontologías en Ingeniería del conocimiento

Una aplicación informática ejecuta instrucciones utilizando unos datos. Actualmente, las instrucciones y procesos se guardan de forma separada de los datos mismos. Esto lleva a que esos datos se crean, se modifican y se gestionan de forma independiente en bases de datos. De forma paralela, se desarrollan lenguajes y formatos estándar para facilitar el intercambio y combinación de los datos. Estas bases de datos evolucionaron a las bases de conocimiento que, con el desarrollo de la Web semántica, han pasado a llamarse *ontologías*. Estas bases de conocimiento no se basan en estructuras de campos y registros, sino en estructuras más complejas y eficientes. De la gestión de los datos se había pasado a la gestión del conocimiento, origen de la Ingeniería del conocimiento.

Las ontologías almacenan información que describe los objetos de un dominio –ya sea de la salud, social o comercial– y organizan estos objetos en clases. Los elementos básicos de una ontología son objetos, propiedades y clases.

Las descripciones de objetos mediante propiedades y su clasificación deben atender a reglas lógicas y ser consistentes entre sí. De este modo, a partir de los datos aportados y de la aplicación de reglas de lógica descriptiva de las que está dotada la ontología, puedan inferirse datos nuevos. El uso de reglas permite también controlar la consistencia de los datos entre sí, es decir, que no incurran en contradicciones, y avisar cuando se introducen datos que no cumplen con las condiciones o requisitos previstos para un objeto.

Para gestionar el conocimiento en la Web Semántica, se han desarrollado lenguajes estándar como RDF y OWL, que se han extendido a otras aplicaciones. La necesidad de disponer de ontologías de dominios en distintos ámbitos, ha llevado a crear programas específicos con la función de crear y modificar ontologías, los *editores de ontologías* (Allemang y Hendler, 2011; Baader y otros, 2017).

Los editores muestran las ontologías mediante interfaces gráficas que facilitan al usuario, incluso no experto, la gestión de los componentes de las ontologías. Incorporan funciones de insertar y gestionar clases en taxonomías, describir objetos describiendo sus propiedades y clasificarlos basándose en ellas. Los sistemas incorporan razonadores y reglas que controlan la consistencia de los datos y realizan inferencias.

Entre los editores de ontologías, ha cobrado especial relevancia el editor Protégé, de la Universidad de Standford (Horridge, 2011), con una comunidad muy amplia de usuarios. Es de código abierto, por lo que toda una comunidad de usuarios desarrolla ampliaciones del programa, funciona en varias plataformas (Windows, MacOs, Web), se actualiza constantemente. Permite el uso de razonadores que se

aplican a la ontología elaborada para comprobar su coherencia y para hacer inferencias, como Hermit y Pellet. Las últimas versiones de Protégé ya no se basan en frames, sino en el lenguaje RDF y OWL2, y permite guardar las ontologías en diversos formatos.

Desde la página web del proyecto Protégé es posible descargar el programa y documentación (Musen, 2015). Es recomendable descargar también el manual de uso del programa, que cuenta con ejemplos para desarrollar paso a paso una ontología sencilla (Horridge, 2011). Al editor se pueden añadir plugins con diversas funciones, que pueden descargarse y activarse, o no, según las necesidades de cada proyecto.

En el marco de la Ingeniería del conocimiento, también se han desarrollado instrumentos que permiten vincular léxico a los conceptos de una ontología. El sistema SKOS permite vincular sinónimos (Miles y Bechhofer, 2009). El modelo Lemon vincula la ontología con información lingüística de distintos niveles (fonología, sintaxis, variantes léxicas, entre otras) (McCrae y otros, 2010).

2. El análisis de conceptos en ONTODIC

Algunos proyectos terminológicos orientados a dar cuenta del significado y las relaciones conceptuales y crear bases de conocimiento, han diseñado las herramientas y sus funciones de acuerdo con las necesidades de gestión terminológica y conceptual.

En el proyecto ONTODIC, en cambio, optamos por usar las herramientas y lenguajes estándar propias de la Ingeniería del conocimiento. Esta no era tarea fácil. Al reto de familiarizarse con los fundamentos, métodos y técnicas de una disciplina diferente, se unía el reto de adaptar las peculiaridades de la representación del léxico y la terminología en forma de ontologías. De ahí que nuestro primer objetivo fue desarrollar la metodología y las técnicas para *encajar* las necesidades de análisis conceptual en terminología con lo que nos ofrecían las ontologías (Alcina, 2009).

Para conseguir nuestro objetivo, estudiamos los fundamentos de las ontologías, incluyendo la lógica descriptiva que las guía, y la metodologías y técnicas establecidas para su generación (por ejemplo, Methontology), los lenguajes de representación en que se basan (RDF, OWL) (Gómez Pérez y otros, 2004). A partir de aquí, se realizó un análisis y contraste de los elementos de que consta una ontología (clases, propiedades de objetos, individuos, propiedades de datos, metadatos), por una parte y de los elementos que resultan del análisis conceptual de los términos de un dominio y cómo se relacionan entre ellos (conceptos, relaciones conceptuales, rasgos definitorios, definición, plantillas de definición), por otra.

A través de este análisis y contraste, decidíamos qué informaciones necesitábamos explicitar en el análisis conceptual y cómo debíamos prepararlas. Con estos análisis de conceptos preparados, iniciamos las tareas de implementación en la ontología.

En los apartados siguientes, mostramos cómo preparamos el análisis conceptual para ajustarlo al objetivo de su implementación en el editor de ontologías. En primer lugar, revisamos brevemente los principales aspectos de las relaciones

conceptuales y sus tipos. Seguidamente, presentamos la aplicación del análisis de conceptos a un grupo de conceptos de la cerámica industrial.

2.1. Las relaciones conceptuales

Los conceptos de un dominio, representados por los términos, mantienen vínculos unos con otros. No nos cabe duda de la existencia de vínculos entre los conceptos representados por *tela* y *tejer*, entre *baldosa* y *mosaico*, entre *uva* y *vino*, entre *quimioterapia* y *cáncer*². Podemos afirmar que la tela es el resultado de la actividad de tejer, que el mosaico es un tipo de baldosa más pequeña y decorativa, que la uva es la materia prima de la que se elabora el vino y que la quimioterapia es el tratamiento que se usa para la curación del cáncer. Estos vínculos entre conceptos muestran lo que llamamos *relaciones conceptuales* (Meyer y otros, 1997; Sager, 1990).

En la formalización, representamos esquemáticamente estas descripciones verbales que hemos utilizado para expresar las relaciones entre dos conceptos. Así, etiquetamos esas relaciones como ‘es resultado de’, ‘es un tipo de’, ‘es materia prima de’ o ‘es tratamiento de’. De modo que obtenemos tres partes: 1) el concepto origen, 2) la relación conceptual y 3) el concepto meta. En la Tabla 1 se muestran esquemáticamente los conceptos (en las columnas laterales, en versalitas) y las relaciones conceptuales que los vinculan (en la columna central, en cursiva).

Concepto	Relación	Concepto
TELA	...es resultado de...	TEJER
MOSAICO	...es un tipo de...	BALDOSA
UVA	...es materia prima de...	VINO
QUIMIOTERAPIA	..es tratamiento de...	CÁNCER

Tabla 1. Ejemplos de conceptos y sus relaciones conceptuales

Las relaciones no son exclusivas de una determinada pareja de términos, sino que se reutilizan en el marco de cada dominio para unir numerosas parejas de términos.

Así, en el marco de un determinado dominio, la aparición reiterada de una relación conceptual para vincular parejas de conceptos permite agrupar los términos con relaciones idénticas. La agrupación de esos conjuntos de conceptos contribuye también a estructurar el conocimiento del dominio. Por ejemplo, permitirá agrupar los términos *mosaico*, *losa*, *losanga*, *mosaiquete*, *zanquín*, *zócalo* por el hecho de que todos ellos mantienen una relación ‘es un tipo de’ con el término *baldosa*.

² Usamos letras versalitas o inicial en mayúsculas cuando hacemos alusión a conceptos, diferenciándolos así de los términos, para los que usaremos cursivas.

Por ejemplo, en el ámbito de la Medicina, la relación ‘es tratamiento de’ que hemos utilizado para vincular la pareja Quimioterapia – Cáncer, también vinculará otras parejas, como Antibiótico – Infección bacteriana, o Analgésico – Dolor. Y el hecho de que los conceptos Quimioterapia, Antibiótico y Analgésico figuren como concepto origen de la relación ‘es tratamiento de’ nos lleva a entender que entre ellos puede existir también algún tipo de vínculo. Lo mismo ocurre con los conceptos meta de esta relación: Cáncer, Infección bacteriana y Dolor.

Algunas relaciones conceptuales se repiten y mantienen la base de su significado estable en el marco de diversos dominios. Así, una determinada relación conceptual vincula parejas de conceptos cuyo contenido informativo tiene poco o nada que ver con el contenido de otras parejas de conceptos. Por ejemplo, el significado de los términos *mosaico* y *aldosa* nada tiene que ver con el significado de los términos *mariposa* e *insecto*; sin embargo, la relación conceptual ‘es un tipo de’, que vincula cada pareja, se mantiene estable e idéntica. Al igual que ocurre en el dominio de la cerámica, en el dominio de la biología la relación conceptual permite estructurar el conocimiento y en este caso agrupar el conjunto de términos *mariposa*, *hormiga*, *libélula*, *grillo*, que mantienen la relación ‘es un tipo de’ con el término *insecto*.

Por otra parte, también vemos que no todas las relaciones se manifiestan en todos los dominios o con la misma frecuencia. Existen relaciones que se manifiestan en determinados dominios y escasamente en otros. Así, por ejemplo, la relación ‘es tratamiento de’, que vincula un término con significado de ‘tratamiento médico’, como *quimioterapia*, con un concepto con significado de ‘enfermedad’, como *cáncer*, es propia del ámbito de la salud, medicina, farmacología. Lo mismo ocurre con otras relaciones que serán más o menos específicas de algún dominio. En cambio, las relaciones ‘es resultado de’ o ‘es materia prima de’ aparecerán en dominios relacionados la industria la textil y la cerámica.

El análisis del comportamiento de las relaciones conceptuales permite distinguir las propiedades que las caracterizan y hacen posible su aplicación en distintos casos y, en su caso, en distintos dominios. Se trata de encontrar las propiedades lógicas de las que, a su vez, se podrá deducir determinadas conclusiones o inducir determinadas generalizaciones. De entre los distintos tipos de relaciones que conoceremos en un dominio, en este caso, nos centraremos en explicar las relaciones genérico-específico para distinguirlas del resto de relaciones, con el objetivo de señalar más adelante la diferente implementación que de una y otras se hará en la ontología y sus consecuencias.

Las relaciones genérico-específico, también llamadas *lógicas*, se basan en que el concepto específico comparte un conjunto de características con el concepto genérico, y añade otras características que el genérico no posee. Por ejemplo, el concepto Mosaico (concepto específico de Aldosa) comparte todas las características del concepto Aldosa (por ejemplo, ‘es un revestimiento cerámico’, ‘fabricado por moldeado y cocción de la pasta cerámica’, etc.) y añade otras que Aldosa no tiene (como ‘tamaño inferior a 7x7cm’, ‘se combina con otras piezas para formar superficies lisas o composiciones artísticas’).

La propiedad de semejanza de características en que se basa la relación genérico-específico se formaliza a modo de reglas lógicas en una ontología, y lleva consigo la posibilidad de ser computadas con el mecanismo de herencia de características. En el apartado de implementación, veremos cómo las ontologías razonan con los conceptos que mantienen esta relación.

2.2. Formalización de conceptos y relaciones

En la primera fase de ONTODIC, nos centramos en la incorporación del análisis de conceptos y flexibilización de las consultas, de modo que permitieran la consulta onomasiológica. Llevamos a cabo el análisis de algunos términos de la cerámica industrial, con el propósito de comprobar cuáles eran los aspectos referentes a características y relaciones conceptuales de los términos susceptibles de ser formalizados y, posteriormente implementados en forma de ontología.

Tomaremos como ejemplo el análisis conceptual del término *corazón negro*. Este término se refiere a un defecto que en ocasiones presentan las baldosas cerámicas como resultado de la presencia de ciertas partículas y que se manifiesta durante la cocción de una pieza cerámica, dando lugar a una mancha oscura en la superficie de la baldosa. Una vez aparece el defecto, la baldosa será desechada porque no hay solución para este tipo de defecto. Esta información la encontramos en diversos contextos extraídos de textos de especialidad del ámbito de la cerámica (Ejemplo 1). También encontramos esta información en el diccionario como definición (Ejemplo 2).

Ejemplo 1

El cuarteo puede ser uno de los defectos de los pavimentos y revestimientos cerámicos esmaltados, que consiste en una serie de fisuras capilares en la superficie esmaltada. Su origen es el desacuerdo de los coeficientes de dilatación del esmalte y del soporte. En la actualidad es un problema poco común debido a los modernos procesos de fabricación y a los controles de calidad que se establecen en la fabricación de estos productos cerámicos.

Ejemplo 2

Corazón negro: Defecto que se presenta como una mancha oscura en la superficie de la pieza cerámica, cuya causa es la presencia de carbono y los óxidos de hierro reducidos en la fase de cocción, y que tiene como consecuencias la disminución de la calidad y las propiedades del producto final e hinchamientos y deformaciones pirolásticas.

Tanto en los contextos como las definiciones detectamos fragmentos de información más pequeños que muestran características que se refieren a una clase de información. Por ejemplo, una clase de característica se refiere a la descripción del 'aspecto físico' que presenta el defecto, mancha oscura, otra clase se refiere al 'lugar o parte de la baldosa donde se presenta el defecto', superficie de la baldosa; otra información se refiere al 'origen o causa del defecto' como la presencia de partículas; y así, podemos etiquetar cada una de las características consideradas relevantes en la descripción del término *corazón negro*. Hemos marcado cada tipo de información con una etiqueta que identifica el tipo de significado al que se asocia. En los ejemplos siguientes (Ejemplos 3 y 4), podemos ver que se han usado etiquetas

como GENUS, TIPO DE PRODUCTO, NATURALEZA FÍSICA, ZONA, CAUSA, FRECUENCIA, CONSECUENCIA.

Ejemplo 3

El cuarteo puede ser un [defecto] GENUS [de los pavimentos y revestimientos cerámicos esmaltados] TIPO DE PRODUCTO, que consiste en [una serie de fisuras capilares] NATURALEZA FÍSICA [en la superficie esmaltada] ZONA. Su origen es el [desacuerdo de los coeficientes de dilatación del esmalte y del soporte] CAUSA. En la actualidad es un [problema poco común] FRECUENCIA debido a los modernos procesos de fabricación y a los controles de calidad que se establecen en la fabricación de estos productos cerámicos.

Ejemplo 4

Corazón negro: [defecto] GENUS [que se presenta como una mancha oscura] NATURALEZA FÍSICA [en la superficie de la pieza cerámica] ZONA, [cuya causa es la presencia de carbono y los óxidos de hierro reducidos] CAUSA [en la fase de cocción] FASE, y [que tiene como consecuencias la disminución de la calidad y las propiedades del producto final e hinchamientos y deformaciones pirolásticas] CONSECUENCIA.

Las etiquetas (y la clase de información que les acompaña) que hemos usado para el término *corazón negro* no es exclusiva de este defecto, sino que aparece de forma reiterada y consistente en los distintos tipos de defecto que los expertos en cerámica industrial señalan para este tipo de producto. Por ello, una vez analizado un conjunto de estos defectos, se obtuvo el factor común a todos ellos representado estas clasificaciones en forma de tablas. En cada tabla, dedicada a un grupo de conceptos, cada columna representa un concepto y cada fila representa la etiqueta referida a una característica (Alcina y Valero, 2008; Valero y Alcina, 2010; 2015). En el cruce de la columna (un concepto) con una fila (etiqueta de característica) se rellena el valor que para cada concepto. En la Tabla 2 podemos ver los valores asociados para los términos *corazón negro* y *cuarteo*.

CONCEPTO	CORAZÓN NEGRO	CUARTEO
GENUS	Defecto	Defecto
NATURALEZA FÍSICA	Mancha oscura	Fisuras capilares
ZONA AFECTADA	Superficie de la baldosa	Superficie esmaltada
CAUSA	Presencia de carbono y los óxidos de hierro reducidos	Desacuerdo de los coeficientes de dilatación del esmalte y del soporte
FASE EN QUE SE PRODUCE	Cocción	
MÉTODO DE FABRICACIÓN	Cualquiera	
TIPO DE PRODUCTO	Cualquier producto cerámico	Pavimentos y revestimientos cerámicos esmaltados
CONSECUENCIA	Disminución de la calidad del producto final	Disminución de la calidad del producto final

	Hinchamientos y deformaciones pirolásticas	
FRECUENCIA	Frecuente	Poco común
GRAVEDAD	Grave	
SOLUCIÓN	Ninguna	

Tabla 2. Tabla de formalización de conceptos y características

Como resultado de esta formalización, tenemos que: en la primera fila (concepto) aparecen los nombres de los conceptos analizados, y en la primera columna aparecen los nombres de las características analizadas para ese grupo de conceptos. En cada celda interior, se leen los valores que corresponden al cruce de un concepto (vertical) con un tipo de característica (horizontal). Pues bien, podemos establecer que entre un concepto que se describe y el valor que recibe respecto a una característica, existe una relación conceptual.

Por tanto, si tomamos el concepto a describir, la relación que los une y el concepto valor tenemos ya un triplete. Para nombrar la relación, y distinguirla de la característica, seguimos la convención en Protégé para formación de estos elementos de vínculo: tomamos el nombre de la característica y añadimos una partícula introductoria. Así, de la característica ‘Naturaleza física’, formamos la relación ‘tieneNaturalezaFísica’. En la Tabla 3 se muestra el ejemplo de este triplete con tres elementos: ‘concepto1relaciónconcepto2’.

Concepto1	Relación	Concepto2
Corazón negro	tieneNaturalezaFísica	Mancha oscura

Tabla 3. Ejemplo de triplete

3. La implementación en el editor de ontologías Protégé

En esta propuesta de implementación ontológica, usamos la versión 5.2. del editor Protégé, que se basa en los lenguajes OWL y RDF, que organizan los datos en tripletes. De ahí que el objetivo de las fases de análisis conceptual y formalización se dirigía a obtener enunciados de tres elementos, o dos elementos vinculados por una relación.

3.1. La creación de la taxonomía de conceptos (‘Class’)

La implementación que presentamos utiliza los elementos ‘Classes’ y ‘Object Properties’ del editor de ontologías, como veremos a continuación. Explicamos en primer lugar en qué consisten estos componentes.

En una ontología, cada concepto del dominio se implementa como clase ‘Class’ en la jerarquía de clases (‘Class hierarchy’), también llamada *taxonomía*. La jerarquía comienza siempre con la clase genérica ‘Thing’ a la que se subordina cualquier nueva clase. Seleccionada una clase es posible introducir clases subordinadas (‘Add subclass’) y coordinadas (‘Add sibling class’). Una vez creada una clase, también es posible moverla para subordinarla a cualquier otra clase en la jerarquía. Una forma alternativa de incorporar clases es escribiendo la taxonomía (o

copiándola desde un editor de texto) como lo haríamos en un editor de texto usando la herramienta específica ('Create class hierarchy'). Por convención, las clases (conceptos) se escriben con la primera letra en mayúsculas.

En la taxonomía de clases, el editor aplica la lógica de la relación jerárquica IS-A ('es un', 'es un tipo de'), equivalente a la relación conceptual genérico-específico en Terminología. Por ejemplo, la relación conceptual entre los términos *mosaico* y *baldosa* o entre *corazón negro* y *defecto* corresponden a este tipo de relación genérico-específico (*Mosaico es-un-tipo-de Baldosa*; *Corazón_negro es-un-tipo-de Defecto*).

El tipo de elemento 'Class' lo hemos usado para representar tanto los conceptos a describir como los conceptos que sirven para describir otros conceptos.

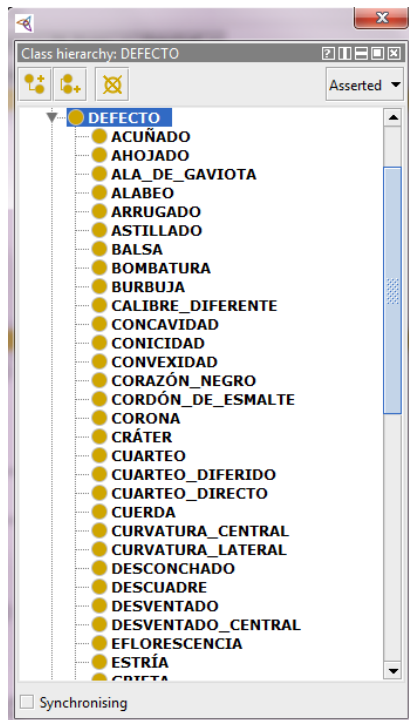


Figura 2. Jerarquía de clases en Protégé

En la Figura 2, se muestra la taxonomía de clases con el concepto genérico Defectos en la parte superior y sus conceptos específicos, como Corazón negro y Cuarteo, subordinados a Defectos.

También los conceptos que hemos asociado con un determinado nombre de característica se han subordinado a un mismo concepto genérico, con el nombre de la característica. Por ejemplo, los conceptos Mancha oscura y Fisuras capilares aparecen subordinados al concepto Naturaleza Física en la taxonomía (ver Figura 3).

Texto original enviado para su publicación en:

Alcina, Amparo. 2020. "La representación de relaciones conceptuales en una ontología." In *Enotradulengua. Vino, lengua y traducción*, edited by Miguel Ibáñez Rodríguez. Berlín: Peter Lang.

Enlace de la editorial: <https://www.peterlang.com/view/title/69064>

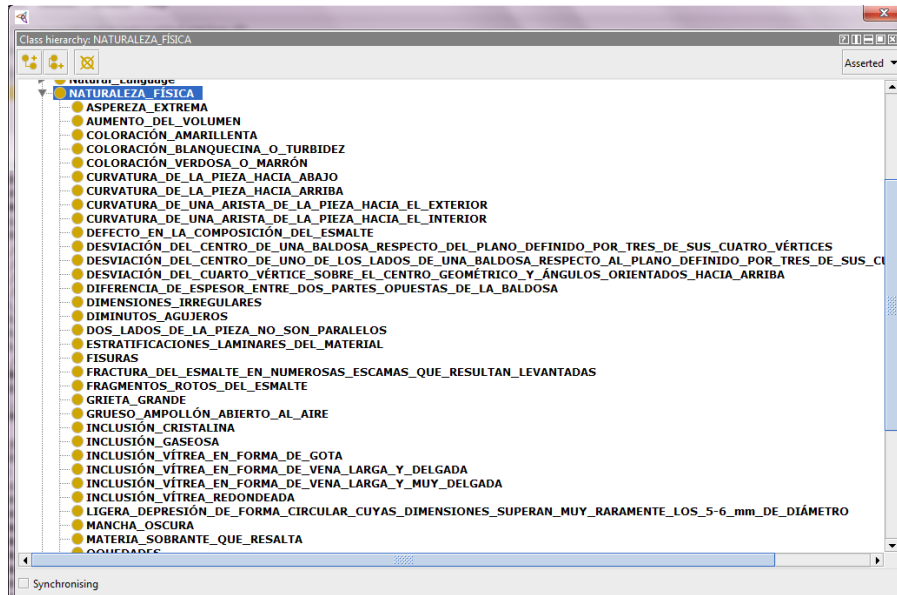


Figura 3. Jerarquía de clases mostrando la clase 'Naturaleza física' y sus subclases

3.2. La descripción de conceptos ('Object Property')

Los elementos del tipo propiedad de objeto ('Object Property') sirven para describir los objetos de una ontología. Por convención, se escriben con su letra inicial en minúscula.

En el elemento propiedad de objeto hemos implementado cada una de las relaciones que nos permiten vincular dos conceptos. Por ejemplo, para vincular el concepto CORAZÓN NEGRO con el concepto Mancha oscura, hemos creado la relación tieneNaturalezaFísica, que a su vez se reutilizará para vincular los conceptos CUARTEO y Fisuras capilares, formando triplete, como mostramos en la Tabla 4.

Clase	Propiedad de objeto	Clase
Corazón negro	hasNaturalezaFísica	Mancha oscura
Cuarteo	hasNaturalezaFísica	Fisuras capilares

Tabla 4. Tripletes formados por clase - propiedad de objeto - clase

Otras relaciones que hemos creado son: *afectaaProducto*, *afectaaZonadelProducto*, *apareceenFasedeFabricación*, *debidoaCausa*, *isPartOf*, *tieneConsecuencia*, *tieneGravedad*, *tieneSolución*, como puede verse en la Figura 4.

Texto original enviado para su publicación en:

Alcina, Amparo. 2020. "La representación de relaciones conceptuales en una ontología." In *Enotradulengua. Vino, lengua y traducción*, edited by Miguel Ibáñez Rodríguez. Berlín: Peter Lang.

Enlace de la editorial: <https://www.peterlang.com/view/title/69064>

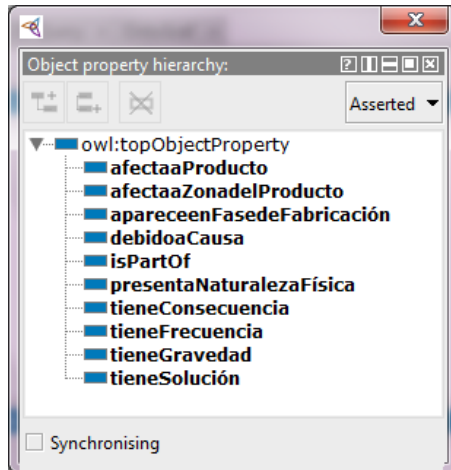


Figura 4. Jerarquía de propiedades de objeto

La descripción de un concepto (ver Figura 5) muestra el conjunto de parejas Relación-Valor que haya sido introducido para describir ese concepto. En la Figura 5, observamos la descripción del concepto CORAZÓN NEGRO. Podemos ver que en la sección superclases ("Superclasses") han quedado descritas las distintas características, formadas por el nombre de una propiedad de objeto, un cuantificador y el valor para esa propiedad.



Figura 5. Descripción del concepto Corazón negro

El plugin Ontograph, que se instala como un plugin del editor, crea grafos a partir del concepto o conceptos seleccionados, y permite escoger las relaciones que queremos ver en la imagen. En la Figura 6 vemos una captura de pantalla de Ontograph que muestra el concepto Calibre diferente conectado mediante flechas (arcos) de distinto color que representan las distintas relaciones. En el panel derecho se enumeran las relaciones indicando el color usado para representarlas.

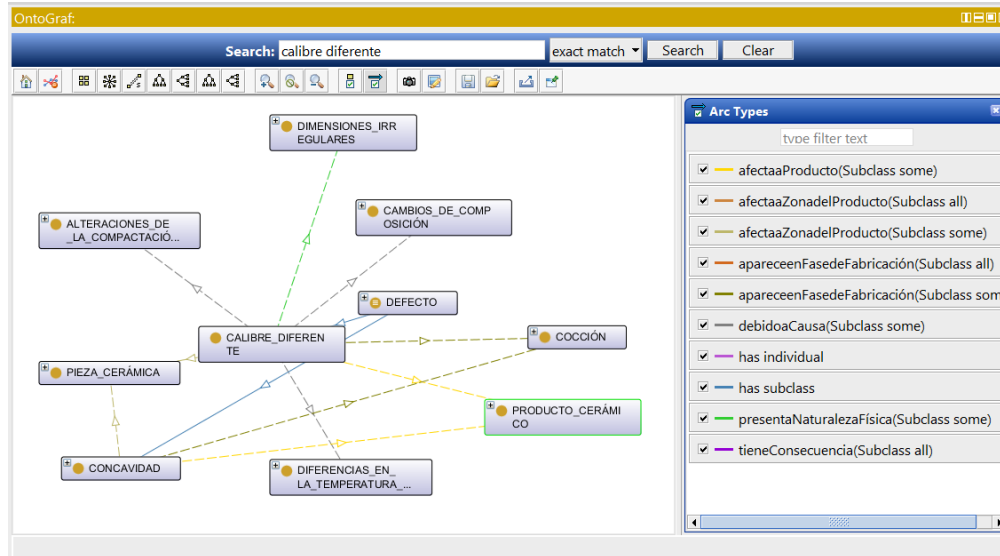


Figura 6. Grafo obtenido con Ontograph

En resumen, el editor de ontologías ofrece un entorno que nos permite gestionar formalmente el análisis de los conceptos, su descripción y sus relaciones, de forma más ágil y eficaz que la que nos proporcionaba el simple uso de tablas en un documento de texto o base de datos genérica o relacional. Además, nos permite ver la organización jerárquica de la taxonomía de conceptos (Figura 2), nos muestra la relación de expresiones usadas para describir las características de un mismo tipo (Figura 3) y así poder homogeneizar su redacción o comparar la descripción de distintos conceptos (Figura 5) para obtener los patrones de definición y mostrar visualmente las relaciones entre conceptos (Figura 6).

4. Razonamiento e inferencias en la ontología

La ventaja definitiva que nos ofrecen las ontologías es que incorporan herramientas que razonan a partir de los datos que se aportan. Estas herramientas pueden razonar gracias a que en ellas se ha implementado reglas de la lógica descriptiva que manejan de forma consistente los elementos con los que puede contar una ontología, es decir, las clases, los objetos y sus propiedades. Obviamente, para que las reglas funcionen y los razonamientos den resultados correctos, es necesario que los datos que se proporcionan estén organizados, formalizados e implementados conforme a dicha lógica.

En la comunicación especializada, y en la cotidiana, hacemos deducciones a menudo. Si nos dicen que ‘tempranillo es una variedad de uva que se cultiva en La Rioja’, asumimos que las características que conocemos de la uva (es decir, que es una fruta, que forma racimos o que con ella se puede elaborar vino, entre otras) se aplican también a tempranillo. Hemos aprendido a hacer este tipo de razonamiento y lo hacemos a diario, sin apenas darnos cuenta de las operaciones de procesamiento que conlleva, de forma automática.

El editor de ontologías puede realizar ese tipo de razonamiento, pues contiene *reglas* que, basadas en la lógica, *operan* con los significados. Es necesario, como decíamos, que estén debidamente formalizados.

Pondremos como ejemplo la implementación de una regla con la relación ISA, equivalente a la relación genérico-específico. Cuando usamos la relación ‘es un tipo de’, entre dos conceptos A y B, como en ‘A es un tipo de B’, lo que esto *significa* es que A posee todas las características que posee B, más otras características que B no posee.

Veamos varias formas de razonamiento que se pueden obtener de esta regla planteando unos casos. Partimos de que hay datos que se afirman acerca de unos hechos, a los que llamamos *explícitos*; y hay otros datos que no se afirman pero que se pueden *inferir* aplicando una regla que conocemos y a partir de unos datos que sí se afirman, a los que denominamos *implícitos*.

CASO A. *Datos explícitos*: 1) Se afirma que el concepto Defecto se describe con la característica ‘disminución de la calidad’. 2) Se afirma que Corazón negro es un concepto específico de Defecto. *Datos implícitos*: 3) En aplicación de la regla de la relación genérico-específico, se infiere que Corazón negro también cumple la característica ‘disminución de la calidad’.

En la descripción del concepto Corazón negro (ver Figura 7), vemos que la característica ‘disminución de la calidad’ no aparece en el apartado ‘SubClass Of’ (donde se afirman los datos) sino en el apartado ‘Subclass of (Anonymous ancestor)’ pues se trata de una característica que ha heredado del concepto genérico. de la descripción del concepto específico.

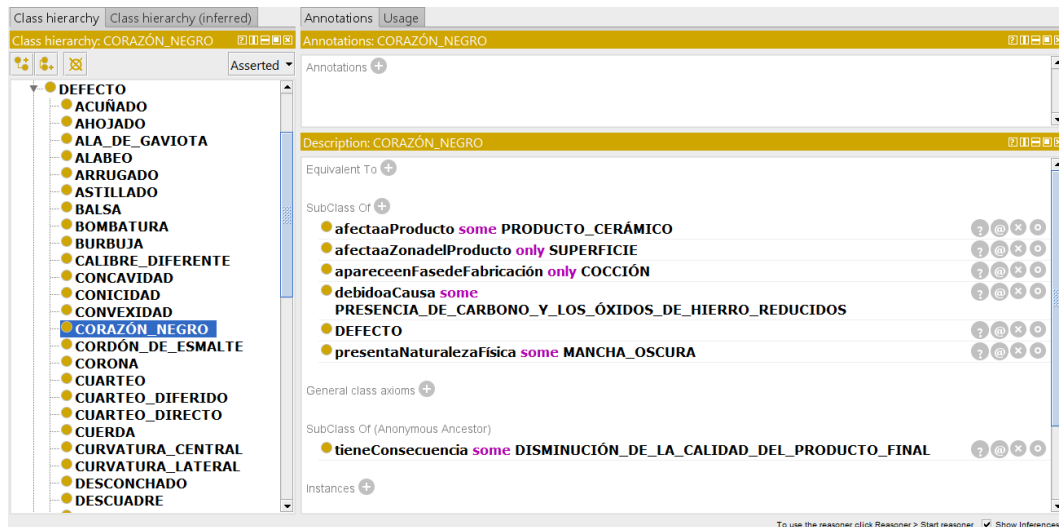


Figura 7. Descripción del concepto Corazón negro

CASO B. *Datos explícitos*: 1) Se afirma que la descripción de Defecto cumple la característica ‘disminución de la calidad’ y además se afirma que tener esta cualidad es condición necesaria y suficiente para que un concepto se clasifique como tipo de Defecto. 2) Se afirma que el concepto Cráter, (situado como un concepto dependiente de ‘Thing’) contiene la característica ‘disminución de la calidad’ (ver Figura 9). *Datos implícitos*: 3) En aplicación de la regla 1), se infiere que Cráter es un concepto específico de Defecto.

En Protégé, podemos expresar la premisa 1), el hecho de que una característica es necesaria y suficiente, mediante la función ‘Defined Class’. Para ello, se

Texto original enviado para su publicación en:

Alcina, Amparo. 2020. "La representación de relaciones conceptuales en una ontología." In *Entradulengua. Vino, lengua y traducción*, edited by Miguel Ibáñez Rodríguez. Berlín: Peter Lang.

Enlace de la editorial: <https://www.peterlang.com/view/title/69064>

selecciona la característica y se elige la opción ‘Convert to Defined Class’ en el menú Edit. La clase Defecto aparece ahora con el símbolo ‘☰’ en la taxonomía; y la característica que hemos convertido en ‘Defined Class’ aparece ahora en el apartado ‘Equivalent To’ (en lugar del apartado ‘SubClass Of’) como podemos ver en la Figura 8.

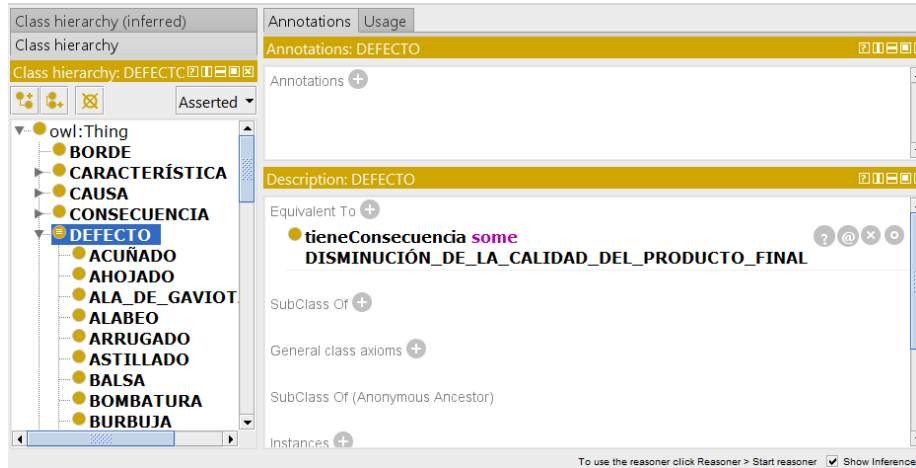


Figura 8. Descripción del concepto Defecto

En la Figura 9, se afirma que el concepto Cráter está descrito con la característica ‘tieneConsecuencia Disminución de la calidad’ (ver panel derecho ‘Description’), y no se ha afirmado su dependencia de ningún concepto genérico (ver panel de jerarquía de clases).

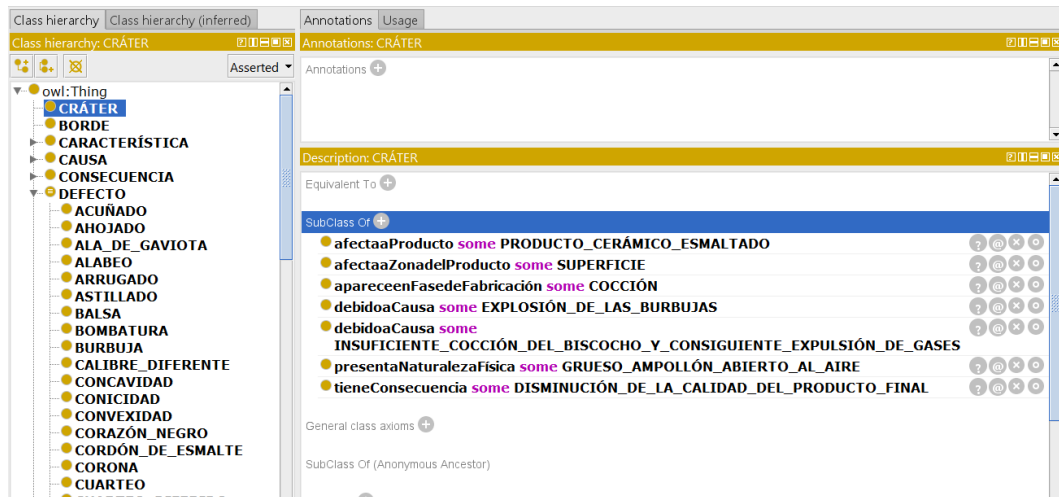


Figura 9. Descripción del concepto Cráter

Para obtener la ejecución de la inferencia que hacemos en 3), aplicamos el razonador Pellet (menú ‘Reasoner’, marcamos el razonador Pellet y a continuación ‘Start Reasoner’). Una vez el razonador se ha ejecutado, vemos que Cráter se muestra como concepto específico de Defecto tanto en el panel de jerarquía de clases inferida (‘Class hierarchy (inferred)’), como en la descripción del concepto Cráter. Los datos inferidos se muestran destacados sobre fondo coloreado, como vemos en la Figura 10.

Texto original enviado para su publicación en:

Alcina, Amparo. 2020. "La representación de relaciones conceptuales en una ontología." In *Enotradulengua. Vino, lengua y traducción*, edited by Miguel Ibáñez Rodríguez. Berlín: Peter Lang.

Enlace de la editorial: <https://www.peterlang.com/view/title/69064>

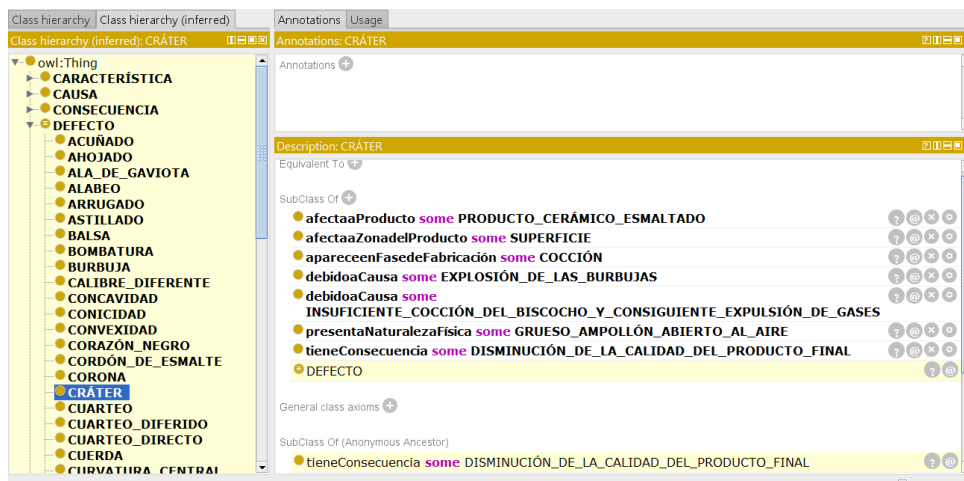


Figura 10. El concepto Cráter desde el panel jerarquía de clases inferida

5. Conclusiones

La terminología puede contar con las ontologías como un instrumento eficaz para desarrollar el análisis, formalización e implementación de conceptos, como hemos mostrado. Además, puede beneficiarse de ventajas como el razonamiento, que contribuye a potenciar la consistencia y la agilidad en la ampliación de datos. En nuestros proyectos, seguimos trabajando para complementar este sistema de análisis conceptual y añadir conocimiento lingüístico de otros niveles (como morfológico, gramatical, sintáctico).

Entendemos que la evolución natural e inminente de los recursos terminológicos es su adaptación al paradigma tecnológico de la Web Semántica y Linked Data, que cuenta con nuevas tecnologías, las ontologías, que se basan en nuevos lenguajes y formatos (RDF, OWL). Estas tecnologías proporcionan mayor flexibilidad en la representación y estructuración de datos y mejoran la consistencia en su gestión. Del uso y adaptación eficaz de estas tecnologías a la terminología dependerá en buena medida el rendimiento que pueda obtenerse de esos datos en la Web semántica.

REFERENCIAS BIBLIOGRÁFICAS

- ALCINA, Amparo (1997): «Soportes de almacenamiento y formas de difusión de datos terminológicos. Las fuentes de información en terminología», *Revista Española de Lingüística Aplicada*, 12, 221-233.
- ALCINA, Amparo (2009): «Metodología y técnicas para la elaboración de diccionarios onomasiológicos» en ALCINA, AMPARO; ESPERANZA VALERO y ELENA RAMBLA (eds.): *Terminología y Sociedad del conocimiento*. Berne: Peter Lang, 33-58.
- ALCINA, Amparo y Esperanza VALERO DOMÉNECH (2017): Description of the terminological concept in an ontology. *Terminology & Ontology: Theories and applications 2017*. Chambéry (Francia).
- ALCINA, Amparo y Esperanza VALERO (2008): Análisis de las definiciones del diccionario cerámico científico-práctico. Sugerencias para la elaboración de patrones de definición. vol. 4. Debate terminológico. Disponible en <<http://seer.ufrgs.br/index.php/riterm/article/view/23841>>.

- Alcina, Amparo. 2020. "La representación de relaciones conceptuales en una ontología." In *Enotradulengua. Vino, lengua y traducción*, edited by Miguel Ibáñez Rodríguez. Berlín: Peter Lang.
Enlace de la editorial: <https://www.peterlang.com/view/title/69064>
- ALLEMANG, Dean y Jim HENDLER (2011): *Semantic Web for the Working Ontologist. Effective Modeling in RDFS and OWL*, Waltham: Elsevier.
- BAADER, Franz; Dresden Ian HORROCKS; Carsten LUTZ y Uli SATTLER (2017): *An Introduction to Description Logic*: Cambridge University Press.
- ESTELLÉS PALANCA, Anna (2014): *Ontología de características de la baldosa cerámica desde la Terminología*, Traducción y comunicación, Universitat Jaume I.
- FABER, Pamela (1999): «Conceptual analysis and knowledge acquisition in scientific translation», *Terminología y Traducción*, 2, 97-123.
- FABER, Pamela (2002): «Oncoterm: Sistema bilingüe de información y recursos oncológicos» en ALCINA, AMPARO y SILVIA GAMERO PÉREZ (eds.): *La traducción científico-técnica y la terminología en la sociedad de la información*. Castelló: Publicacions de la Universitat Jaume I, 177-188.
- GÓMEZ PÉREZ, Asunción; Mariano FERNÁNDEZ LÓPEZ y Oscar CORCHO (2004): *Ontological Engineering*, New York: Springer.
- HORRIDGE, Matthew (2011): *A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools*, Manchester: University of Manchester. Disponible en <<http://goo.gl/m2ChLn>>.
- LEÓN-ARAÚZ, Pilar; Arianne REIMERINK y Pamela FABER (2019): «EcoLexicon and by-products: integrating and reusing terminological resources», *Terminology, Sp. Issue Terminology and e-dictionaries*, 25: 2, in press.
- LEXICON RESEARCH GROUP EcoLexicon terminological knowledge database. Disponible en <<http://ecolexicon.ugr.es/>>.
- MADSEN, Bodil Nistrup y Hanne Erdman THOMSEN (2009a): CAOS – A tool for the Construction of Terminological Ontologies vol. NODALIDA 2009 Conference Proceedings, ed. por JOKINEN, KRISTIINA y ECKHARD BICK. 279-282.
- MADSEN, Bodil Nistrup y Hanne Erdman THOMSEN (2009b): «Terminological concept modelling and conceptual data modelling», *Int. J. Metadata, Semantics and Ontologies*, 4: 4, 239-249.
- MAROTO, Nava (2007): *Las relaciones conceptuales en la terminología de los productos cerámicos y su formalización mediante un editor de ontologías*, Traducción y Comunicación, Universitat Jaume I.
- MCCRAE, John; Guadalupe AGUADO DE CEA; Paul BUITELAAR; Philipp CIMIANO; Thierry DECLERCK; Asunción GÓMEZ PÉREZ; Jorge GRACIA; Laura HOLLINK; Elena MONTIEL-PONSODA; Dennis SPOHR y Tobias WUNNER (2010): *The lemon cookbook*. Disponible en <<https://www.lemon-model.net/lemon-cookbook/>>.
- MEYER, Ingrid; Karen ECK y Douglas SKUCE (1997): «Systematic Concept Analysis within a Knowledge-Based Approach to Terminology» en WRIGHT, SUE ELLEN y GERHARD BUDIN (eds.): *Handbook of Terminology Management*. Philadelphia: John Benjamins, 98-118.
- MILES, Alistair y Sean BECHHOFFER (eds.) (2009): *SKOS Reference: W3C Recommendation 18 August 2009*. Disponible en <<http://www.w3.org/TR/skos-reference>>.

Alcina, Amparo. 2020. "La representación de relaciones conceptuales en una ontología." In *Entradulengua. Vino, lengua y traducción*, edited by Miguel Ibáñez Rodríguez. Berlín: Peter Lang.

Enlace de la editorial: <https://www.peterlang.com/view/title/69064>

- MORENO ORTIZ, Antonio (2002): «Representación de la información terminológica en OntoTerm: un sistema gestor de bases de datos terminológicas basado en el conocimiento» en FABER, PAMELA y CATALINA JIMÉNEZ (eds.): *Investigar en Terminología*. Granada: Comares, 25-70.
- MUSEN, M.A. (2015): «The Protégé project: A look back and a look forward», *AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence*, 1: 4, [June 2015]. También disponible en <<http://protege.stanford.edu/index.html>>.
- PASTOR, Verónica y Amparo ALCINA (2010): «Search Techniques in Electronic Dictionaries: A Classification for Translators», *International Journal of Lexicography*, 23: 3, 307-307. También disponible en <<http://ejournals.ebsco.com/direct.asp?ArticleID=4AF28EA21BCC85219A44>>.
- SAGER, Juan Carlos (1990): *A practical course in terminology processing*, Amsterdam/Philadelphia: John Benjamins.
- VALERO, Esperanza y Amparo ALCINA (2009): Linguistic realization of conceptual features in terminographic dictionary definitions. International Workshop on Definition Extraction Proceedings, ed. por SIERRA, GERARDO; MARA POZZI y JUAN-MANUEL TORRES. 54-60. Borovets, Bulgaria.
- VALERO, Esperanza y Amparo ALCINA (2010): «Exploración de características conceptuales en contextos ricos en conocimiento mediante un programa de análisis cualitativo», *Revista de Lingüística y Lenguas Aplicadas*, 5, 241-254.
- VALERO, Esperanza y Amparo ALCINA (2015): «Aspectos críticos de la formalización de características conceptuales en la definición terminográfica», *Terminalia*, 11, 30-44. Disponible en <http://revistes.iec.cat/index.php/Terminalia/article/view/119727/pdf_627>.