



# Tecnologías para el Modelado, Procesamiento y Gestión de Conocimiento

Jesualdo Tomás Fernández Breis

Universidad de Murcia, IMIB-Arrixaca

Murcia, España

jfernand@um.es

Marcos Menárguez Tortosa

Universidad de Murcia, IMIB-Arrixaca

Murcia, España

marcos@um.es

Catalina Martínez Costa

Medical Graz University

Graz, Austria

catalina.martinez@medunigraz.at

José Antonio Miñarro Giménez

Medical Graz University

Graz, Austria

jose.minarro-gimenez@medunigraz.at

María del Carmen Legaz García

FFIS, IMIB-Arrixaca

Murcia, España

mcarmen.legaz@ffis.es

Manuel Quesada Martínez

Universidad Miguel Hernández

Elche, España

mquesada@umh.es

Astrid Duque Ramos

Universidad de Antioquia

Medellín, Colombia

astrid.duquer@udea.edu.co

Ángel Esteban Gil

FFIS, IMIB-Arrixaca

Murcia, España

angel.esteban@ffis.es

Dagoberto Castellanos Nieves

Universidad de La Laguna

Tenerife, España

dcastell@ull.es

**Abstract**—El grupo de investigación *Tecnologías para el Modelado, Procesamiento y Gestión de Conocimiento* lleva quince años investigando, desarrollando y aplicando tecnologías semánticas en dominios biomédicos. En este trabajo se describen las líneas de investigación más relevantes en las que hemos trabajado en los últimos años, y cuyo objetivo principal es la consecución de interoperabilidad semántica entre sistemas de información sanitarios. Asimismo, también se comentarán objetivos de investigación para los próximos años en los que las técnicas inteligentes desempeñarán un papel importante.

**Index Terms**—Artificial Intelligence, Knowledge Engineering, Electronic Medical Records

## I. PRESENTACIÓN DEL GRUPO DE INVESTIGACIÓN

El grupo *Tecnologías para el Modelado, Procesamiento y Gestión de Conocimiento* (TECNOMOD) se creó en la Universidad de Murcia en el año 2003 constituyendo la ingeniería ontológica, la web semántica y las tecnologías del lenguaje sus áreas principales de investigación. Actualmente el grupo de investigación tiene 15 integrantes (profesores, becarios y contratados), y cuenta con colaboradores a nivel nacional e internacional, incluyendo antiguos doctorandos del grupo cuya contribución a los resultados que se presentan en este trabajo ha sido fundamental. La mayoría de los integrantes del grupo de investigación son informáticos, si bien disponemos de dos investigadores cuya formación y actividad es sanitaria y un biotecnólogo. Hemos aplicado nuestra investigación en diversos dominios como el turismo, las finanzas, la política o la educación, pero la biomedicina ha sido el área donde más investigación hemos realizado. En este documento nos ceñiremos a la investigación realizada exclusivamente en este ámbito, y especialmente destacaremos dos líneas: interoperabilidad semántica y aseguramiento de la calidad de ontologías y terminologías biomédicas.

TECNOMOD es miembro del Instituto Murciano de Investigación Biosanitaria (IMIB-Arrixaca), que está acreditado por el Instituto de Salud Carlos III. IMIB-Arrixaca está vinculado al Hospital Universitario Virgen de la Arrixaca. Tenemos colaboraciones en marcha con varios grupos de investigación del instituto y una colaboración permanente con la Plataforma de Informática Biomédica y Bioinformática del mismo.

## II. CONTEXTO

La Web Semántica [2] es un espacio natural para la integración de datos y la interoperabilidad semántica entre sistemas, imponiendo un entorno de trabajo en el que cada sistema emplea el significado de los datos en diferentes contextos [10]. Las tecnologías semánticas posibilitan la descripción del contexto lógico de la información a intercambiar, mientras que se permite que cada sistema mantenga su máxima independencia. Las ontologías constituyen el nivel fundamental de la Web Semántica desde el punto de vista de representación formal del conocimiento, de ahí que parte del éxito de la Web Semántica recaiga en la calidad de las ontologías, por lo que el desarrollo de métodos que permita asegurar la calidad de las mismas es un objetivo crítico para entornos de interoperabilidad. En los últimos años, las tecnologías de la Web Semántica han ganado popularidad para la consecución de interoperabilidad semántica entre sistemas de información sanitarios, especialmente desde que el proyecto Semantic Health [24] recomendó su uso para dichos fines. Posteriormente, la FP7 Network of Excellence SemanticHealthNet<sup>1</sup> propuso que la formalización ontológica debería ser fundamental para permitir el intercambio y la cooperación entre los sistemas de historia clínica

<sup>1</sup><http://www.semantichealthnet.eu>



electrónica (HCE) y los sistemas de ayuda a la decisión (SAD). Dicho rol de las ontologías en escenarios de interoperabilidad impone una serie de requisitos sobre las ontologías [8]:

- Facilitar la representación, compartición, reutilización de conocimiento para modelos de información y de inferencia.
- Clasificar y recuperar datos de HCE según las reglas establecidas en guías y protocolos.
- Tener garantizada su calidad formal y estructural.

En este trabajo agrupamos nuestras líneas de investigación en dos grupos:

- Interoperabilidad semántica en salud (Sección III): investigación orientada a conseguir el intercambio de información entre sistemas sanitarios, incluyendo HCE y SAD.
- Aseguramiento de calidad de ontologías (Sección IV): investigación orientada a evaluar el cumplimiento de los requisitos de interoperabilidad por parte de las ontologías y terminologías biomédicas.

La investigación que se describirá en las próximas secciones ha sido posible gracias a financiación fundamentalmente pública. A continuación se enumeran los proyectos más relevantes de los últimos diez años.

- **Plataforma para la adquisición y compartición de información y conocimiento para comunidades de investigación clínica en red II** (TSI2007-66575-C02-02). Ministerio de Educación y Ciencia. 01/10/2007-31/12/2010.
- **Herramientas inteligentes para enlazar historias clínicas electrónicas y sistemas de ensayos clínicos II** (TIN2010-21388-C02-02). Ministerio de Ciencia e Innovación. 01/01/2011-31/12/2014.
- **Modelos de información y conocimiento clínicos para enlazar los sistemas de historia clínica electrónica y de ayuda a la decisión clínica II** (TIN2014-53749-C2-2-R). Ministerio de Economía y Competitividad. 01/01/2015-31/12/2018.
- **Semantic Interoperability for Health Network**. Unión Europea. 03/10/2012-30/11/2014. Red de Excelencia del programa FP7 en la que miembros de TECNOMOD participaron como expertos externos en el paquete de trabajo 4 "Harmonised resources for EHR/PHR and aggregation".
- **Gene regulation ensemble effort for the knowledge commons** (CA COST Action CA15205). Unión Europea. 08/09/2016 - 07/09/2020. Participación de TECNOMOD como líder del paquete de trabajo de ontologías y vocabularios controlados.
- **Unraveling in utero determinants predicting lung function in infants: a step for prenatal prevention of asthma** (PIE15/00051). Instituto de Salud Carlos III. 01/01/2016-31/12/2018.
- **Infraestructura y tecnologías de interoperabilidad para aplicaciones de Learning Health Systems I**

(TIN2017-85949-C2-1-R). Ministerio de Economía, Industria y Competitividad. 01/01/2018 - 31/12/2020.

### III. INTEROPERABILIDAD SEMÁNTICA EN SALUD

#### A. Transformación de datos y modelos clínicos

Las recomendaciones del proyecto SemanticHealth incluyeron el uso combinado de estándares de historia clínica electrónica, ontologías y terminologías biomédicas para facilitar el intercambio significativo de datos. Nuestro grupo de investigación desarrolló PoseacleConverter<sup>2</sup> como demostrador de las posibilidades de intercambio de datos entre distintos formatos de modelos clínicos (openEHR, ISO 13606 o CEM) y de datos entre openEHR e ISO 13606.

El trabajo comenzó en el año 2008 y el primer objetivo fue la transformación de arquetipos (como tipo de modelo clínico) entre los estándares openEHR e ISO 13606 [14]. La transformación de modelos se basó en la formalización de correspondencias entre las ontologías que representaban los modelos de información de los estándares. Esto permitió definir unas reglas que guiaron la transformación estructural de los arquetipos. Este método se extendió para transformar datos entre dichos estándares explotando las correspondencias anteriores [13]. El uso de ontologías permitió también comprobar en ambos esfuerzos la consistencia lógica de los modelos y datos obtenidos mediante el uso de razonadores. La figura 1 describe el proceso seguido a nivel de datos. Se puede apreciar en la parte inferior el extracto de datos fuente ISO 13606, que es transformado en un extracto openEHR usando las correspondencias definidas a partir de las ontologías. Para la aplicación de la transformación de datos no es necesaria la existencia del arquetipo destino, el cual se crea automáticamente. Esto genera una traza de transformación que es empleada para la generación del extracto destino.

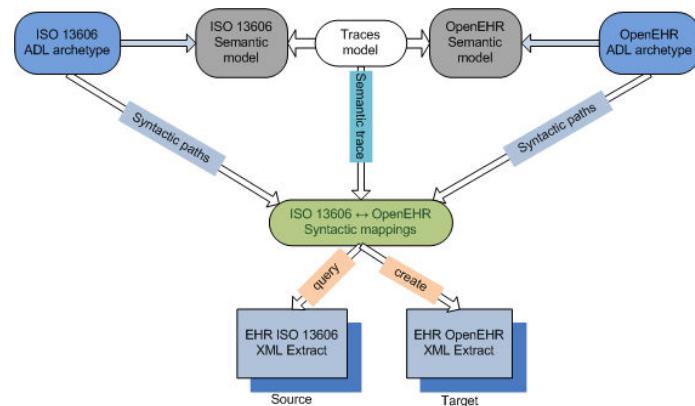


Fig. 1. Esquema de transformación de datos

Posteriormente el enfoque de transformación de modelos clínicos evolucionó cuando diseñamos el proceso para obtener arquetipos openEHR a partir de modelos clínicos CEM [11]. El nuevo enfoque se desarrolló exclusivamente con tecnologías OWL. En este caso las relaciones entre los modelos de

<sup>2</sup>miuras.inf.um.es/PoseacleConverter



referencia se expresaron como axiomas OWL y se crearon plantillas de transformación. De esta manera, la transformación se concibió como un proceso generativo de axiomas OWL, y el proceso estaba asistido por un razonador. La figura 2 ejemplifica este proceso de generación: (1) representación OWL del modelo clínico CEM; (2 y 3) plantillas de generación para los dos ítems incluidos en el modelo; (4) ontología del modelo de información openEHR, que se utiliza para la definición de las plantillas; (5, 6 y 7) representación parcial en OWL del modelo transformado a openEHR.

Estos métodos se evaluaron y validaron con las colecciones de arquetipos openEHR disponibles en CKM y modelos CEM disponibles en la librería de Intermountain disponibles en su momento.

#### *B. Validación semántica de arquetipos*

El intercambio automático de datos a través de sistemas informáticos basado en modelos clínicos, como pueden ser los arquetipos, requiere la corrección formal de los mismos. Nuestro grupo de investigación desarrolló Archeck<sup>3</sup> [15], que es un método basado en OWL para la validación de arquetipos. Los arquetipos son representados como clases OWL y la validación semántica de los mismos se realiza mediante el uso de razonadores automáticos. Para ello usa una ontología que define los tipos de restricciones que se pueden definir para un arquetipo y que vienen especificados en el modelo de arquetipos. Además, Archeck extiende esta ontología para poder definir métricas de calidad para arquetipos basadas en OWL. El método Archeck se aplicó a dos repositorios públicos de arquetipos openEHR, el repositorio gestionado en la herramienta Clinical Knowledge Manager (CKM) y el repositorio de programa National Health Service (NHS) del Reino Unido. La evaluación de los repositorios se centró en los arquetipos que definen relaciones de especialización, 81 en CKM y 212 en NHS. La validación de los repositorios encontró que un 22,2% de los arquetipos especializados en el repositorio CKM y un 21,2% en NHS contenían errores. De ellos, el 3% de los errores en CKM fueron identificados gracias a la métrica de calidad descrita anteriormente. Si bien puede considerarse un valor bajo esto también se debió a que los muchos de los arquetipos evaluados no contenían enlaces terminológicos, lo cual es un indicador negativo de calidad.

#### *C. Interoperabilidad de modelos de información, dominio e inferencia*

El trabajo en esta línea es el fruto de proyectos coordinados con la Universidad Politécnica de Valencia (UPV) y la Universidad Jaime I (UJI). UPV es experto en modelos de información, UJI en modelos de inferencia y nuestro grupo en modelos de dominio. El trabajo realizado en esta línea pretende conseguir aplicar las guías clínicas computerizadas a datos estandarizados de la HCE. Para ello se presentan las ontologías como elemento mediador que proporcione la semántica compartida entre ambos modelos.

<sup>3</sup>[miuras.inf.um.es/archek](http://miuras.inf.um.es/archek)

*1) Flujos de transformaciones interoperables. Aplicación a la clasificación de pacientes:* El primer trabajo en esta línea afrontó cómo abordar problemas de clasificación de pacientes a partir de protocolos estandarizados mediante las tecnologías de estándares de HCE y las ontologías [7]. Esto determinó un flujo de trabajo cuya entrada son datos no normalizados de HCE y la salida es la clasificación del paciente, tal y como se puede ver en la figura 3. En este caso trabajamos también en colaboración con el Programa de cribado de cáncer de colon y recto de la Región de Murcia. El objetivo específico era determinar el nivel de riesgo de los pacientes del programa. Para ello trabajamos con las guías europea y americana de cáncer de colon y recto. Las tecnologías semánticas contribuyen en dos etapas del proceso, como son la representación de los datos en formato semántico y la clasificación del paciente usando razonadores. Para ello también fue necesario formalizar las reglas de los protocolos en OWL. Este trabajo se sigue desarrollando a través de la plataforma CLIN-IK-LINKS [12], que busca generalizar y facilitar la definición y ejecución de transformaciones de datos interoperables.

*2) Enriquecimiento de guías clínicas computerizadas:* Una línea reciente de trabajo busca formalizar el conocimiento de las guías clínicas computerizadas (GCC) con conceptos existentes en ontologías y terminologías biomédicas, trabajo que se está realizando con colaboración con la UJI. El primer trabajo realizado ha sido identificar los conceptos de ontologías de BioPortal a partir de un conjunto de GCC seleccionado y disponible en formato PROforma [9]. SNOMED CT fue el recurso semántico con mayor número de resultados, por lo que es la que estamos usando en los estudios y evaluaciones preliminares del método desarrollado [22]. La figura 4 muestra cómo procesamos la GCC. Una vez extraído el texto a procesar, se buscan alineamientos usando métodos basados en OntoEnrich (ver Sección IV-B). El resultado es un conjunto de recomendaciones de conceptos asociados a la guía clínica.

### IV. ASEGURAMIENTO DE CALIDAD DE TERMINOLOGÍAS Y ONTOLOGÍAS

#### *A. Evaluación de calidad de ontologías*

OQuaRE<sup>4</sup> [5] es un *framework* para la evaluación de ontologías basado en la norma ISO/IEC 25000:2005 para la definición de requisitos y evaluación de la calidad de productos software, también conocida como SQuaRE [1]. OQuaRE realiza la evaluación de la calidad de ontologías con tres niveles de granularidad: características, subcaracterísticas y métricas de calidad. Actualmente el modelo de calidad de OQuaRE incluye 8 características, 29 subcaracterísticas y 19 métricas. Cada característica tiene un conjunto de subcaracterísticas asociadas que, a su vez, tienen asociadas un conjunto de métricas. Toda la información sobre el modelo de calidad está disponible en el sitio web de información sobre OQuaRE<sup>5</sup>.

Las métricas de OQuaRE tienen una función  $f(x)$  cuyo dominio es una ontología, pero el rango de cada métrica puede

<sup>4</sup><http://sele.inf.um.es/oquare>

<sup>5</sup><http://miuras.inf.um.es/oquarewiki>

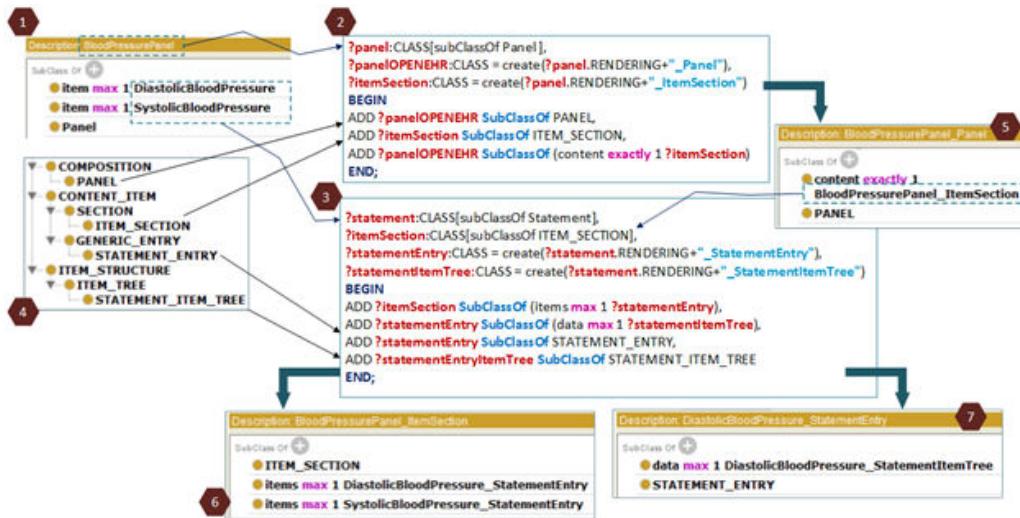


Fig. 2. Esquema de transformación de datos

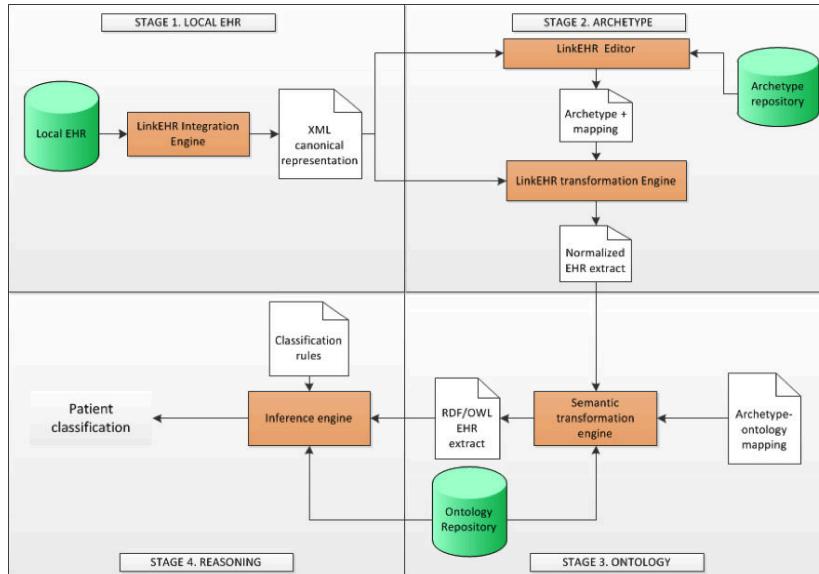


Fig. 3. Flujo de transformación de datos

variar. Tener una visión global de la evaluación de una ontología requiere combinar los resultados de todas las métricas. Es por ello que OQuaRE define dos tipos de funciones de escalado de métricas, que permiten abstraer al método de los distintos rangos de cada métrica:

- **Estática**, basada en recomendaciones y buenas prácticas.  $n(f(x))$  está predefinida y está basada en intervalos fijos continuos que partitionan el rango de  $f(x)$  en  $k$  categorías.
- **Dinámica**, que usa datos experimentales como referencia.  $n(f(x))$  aplica el método de clustering  $k$ -means para partitionar el rango de  $f(x)$  en  $k$  intervalos continuos no prefijados que contienen todas las observaciones incluidas en los datos experimentales. Para maximizar la compactación de las ontologías de cada categoría, min-

imizando la varianza intra-cluster, y para maximizar la separación entre las categorías, maximizando la varianza inter-cluster en cada iteración se recalculan los nuevos  $k$  centroides y la asignación de cluster se realiza asociando cada  $R_{\theta_j}$  al centroide más cercano. El algoritmo de clustering requiere la información de si los valores altos de  $f(x)$  se corresponden con las categorías más altas del factor. Esto se debe a que valores altos de una métrica no tienen por qué representar siempre una buena propiedad de una ontología.

Además, cabe mencionar que la versión actual de OQuaRE usa  $k = 5$  para todas las métricas. La escala estática de OQuaRE ha sido aplicada para analizar ontologías [3], mientras que la dinámica ha sido empleada para estudiar la evolución de ontologías. Para ello se han usado como datos

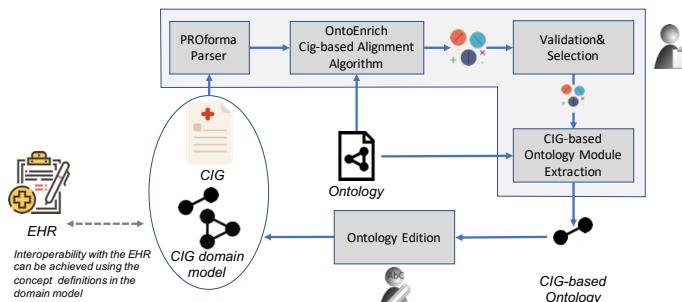


Fig. 4. Diagrama de procesamiento del contenido de la guía clínica computarizada

experimentales los valores de las métricas para cada versión de la ontología procesada [6], [17]. El conjunto de datos experimentales también puede estar compuesto por una versión de cada ontología, sirviendo entonces para medir similitudes y diferencias entre ellas.

#### B. Enriquecimiento de ontologías

OntoEnrich<sup>6</sup> es un método para el enriquecimiento de ontologías biomédicas basada en el análisis léxico de sus etiquetas [16], ya que las ontologías son ricas en contenido léxico para humanos, el cual no se pone siempre a disposición de las máquinas en forma de axiomas lógicos. Por tanto, el objetivo principal de OntoEnrich es explotar la semántica oculta de las ontologías [25]. OntoEnrich se basa en el concepto de regularidad léxica (RL), que se define como un conjunto de tokens consecutivos repetidos en diferentes etiquetas de la ontología. Cabe mencionar que la entrada habitual a OntoEnrich es una ontología en formato OWL, si bien en distintos casos de uso la herramienta que implementa la metodología ha sido adaptada para aceptar otros tipos de entradas, como ficheros de texto donde cada línea incluya cada etiqueta a analizar. La figura 5 muestra las principales etapas de la metodología:

- **Procesamiento de la ontología y cálculo de regularidades léxicas.** En primer lugar permite obtener el conjunto completo de regularidades léxicas (RL) utilizando como parámetro de entrada un *threshold* mínimo que permite podar las búsquedas. También se establecen relaciones super-sub entre regularidades léxicas. Se realiza también una descripción cuantitativa de cada RL. Se buscan alineamientos a partir de las RD en la propia ontología o en otras externas para promover la reutilización de conceptos entre la comunidad biomédica, y estos algoritmos buscan clases que contengan la RL o cuya etiqueta coincida con la RL [21].
- **Cálculo de métricas avanzadas.** En esta etapa se calculan métricas avanzadas que relacionan las RL con diferentes aspectos semánticos de la ontología, como localización o modularidad de las RL y los productos

<sup>6</sup><http://sele.inf.um.es/ontoenrich>

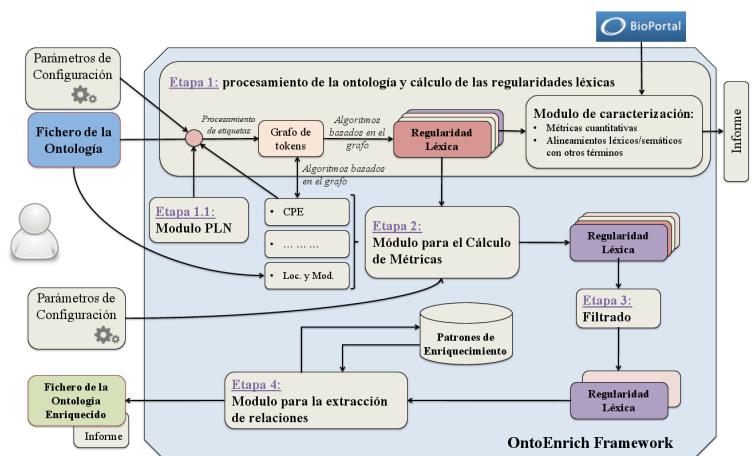


Fig. 5. Descripción de la metodología aplicada por OntoEnrich

cruzados, que informa sobre la potencialidad de enriquecimiento de una RL.

- **Filtrado de regularidades.** OntoEnrich permite utilizar las métricas para definir filtros que reducen el conjunto de RL a aquellas que cumplen ciertas propiedades. Se podría filtrar a aquellas RL que poseyeran determinado valor de modularidad, localización, etc. También se pueden filtrar las subregularidades para quedarnos con las regularidades más largas, ya que esto quiere decir que son más específicas.
- **Extracción de relaciones y axiomas.** En este paso se permite la definición de patrones axiomáticos a partir de las RL anteriores. El método es capaz de crear automáticamente patrones a partir de relaciones de subclase. Como resultado de la ejecución de dichos patrones se obtendría la ontología enriquecida.

#### C. Aplicaciones basadas en OQuaRE y OntoEnrich

Los métodos OQuaRE y OntoEnrich se han aplicado para analizar las ontologías de repositorios de ontologías biomédicas como OBO Foundry y BioPortal con el objeto de caracterizar dichas ontologías desde el punto de vista de optimidad para su uso en escenarios de interoperabilidad. Los aspectos específicos que han contribuido a estudiar han sido:

- Evaluación de ontologías [4], [6], [17].
- Evaluación de la aplicación de principios de diseño de ontologías, como delineación de contenido, riqueza de relaciones y nombrado sistemático [20], [21].
- Análisis de reutilización de axiomas y axiomas ocultos en ontologías [18].
- Enriquecimiento de la Gene Ontology [23].
- Detección de axiomas potencialmente erróneos y detección de posibles axiomas en SNOMED CT [19], [26].

#### V. PRÓXIMOS OBJETIVOS

En estas líneas de investigación pretendemos aplicar a corto plazo técnicas de inteligencia artificial para los siguientes



objetivos:

- Desarrollar el concepto de *Learning Health Systems* incluyendo técnicas de machine learning dentro de los flujos de procesamiento de datos interoperables.
- Analizar ontologías a partir de conjuntos de datos de HCE.
- Extracción de módulos de ontologías para el enriquecimiento de guías clínicas computerizadas y de otras ontologías.
- Aprendizaje del valor óptimo de  $k$  para cada métrica empleada en OQuaRE en la función de escalado dinámica.
- Clasificación y analizar tipos y grupos de ontologías a partir de los valores de sus métricas.
- Identificación de relaciones entre métricas de las ontologías y principios de diseño.
- Aprendizaje automático de patrones axiomáticos a partir de regularidades léxicas.

#### AGRADECIMIENTOS

Este trabajo ha sido posible gracias a la financiación del Ministerio de Educación y Ciencia (TSI2007-66575-C02-02), Ministerio de Ciencia e Innovación (TIN2010-21388-C02-02), Ministerio de Economía, Industria y Competitividad (TIN2014-53749-C2-2-R, TIN2017-85949-C2-1-R), así como el Fondo Europeo de Desarrollo Regional (FEDER) a través de los proyectos citados.

#### REFERENCIAS

- [1] ISO/IEC 25000:2005, Software Engineering - Software Product Quality Requirements and Evaluation (SQuaRE) - Guide to SQuaRE, 2005. [Online; accessed 01-June-2017].
- [2] T. Berners-Lee, J. Hendler, O. Lassila, et al. The semantic web. *Scientific american*, 284(5):28–37, 2001.
- [3] A. Duque-Ramos, M. Boeker, L. Jansen, S. Schulz, M. Iniesta, and J. T. Fernández-Breis. Evaluating the good ontology design guideline (goodod) with the ontology quality requirements and evaluation method and metrics (oquare). *PLOS ONE*, 9(8):1–14, 08 2014.
- [4] A. Duque-Ramos, J. T. Fernández-Breis, M. Iniesta, M. Dumontier, M. E. Aranguren, S. Schulz, N. Aussenac-Gilles, and R. Stevens. Evaluation of the oquare framework for ontology quality. *Expert Systems with Applications*, 40(7):2696–2703, 2013.
- [5] A. Duque-Ramos, J. T. Fernández-Breis, R. Stevens, and N. Aussenac-Gilles. OQuaRE: A SQuaRE-based approach for evaluating the quality of ontologies. *Journal of Research and Practice in Information Technology*, 43(2):159–176, 2011.
- [6] A. Duque-Ramos, M. Quesada-Martínez, M. Iniesta-Moreno, J. T. Fernández-Breis, and R. Stevens. Supporting the analysis of ontology evolution processes through the combination of static and dynamic scaling functions in oquare. *Journal of Biomedical Semantics*, 7(1):63, 2016.
- [7] J. T. Fernández-Breis, J. A. Maldonado, M. Marcos, M. d. C. Legaz-García, D. Moner, J. Torres-Sospedra, A. Esteban-Gil, B. Martínez-Salvador, and M. Robles. Leveraging electronic healthcare record standards and semantic web technologies for the identification of patient cohorts. *J Am Med Inform Assoc*, 20(e2):e288–96, Dec 2013.
- [8] J. T. Fernández-Breis, M. Quesada-Martínez, and A. Duque-Ramos. Can existing biomedical ontologies be more useful for ehr and cds? In *International Workshop on Knowledge Representation for Health Care*, pages 3–20. Springer, 2016.
- [9] J. Fox, N. Johns, and A. Rahmazadeh. Disseminating medical knowledge: the proforma approach. *Artif Intell Med*, 14(1-2):157–81, 1998.
- [10] C. Goble and R. Stevens. State of the nation in data integration for bioinformatics. *Journal of biomedical informatics*, 41(5):687–693, 2008.
- [11] M. d. C. Legaz-García, M. Menárguez-Tortosa, J. T. Fernández-Breis, C. G. Chute, and C. Tao. Transformation of standardized clinical models based on owl technologies: from cem to openehr archetypes. *J Am Med Inform Assoc*, 22(3):536–44, May 2015.
- [12] J. A. Maldonado, M. Marcos, J. T. Fernández-Breis, E. Parcerio, D. Boscá, M. D. C. Legaz-García, B. Martínez-Salvador, and M. Robles. A platform for exploration into chaining of web services for clinical data transformation and reasoning. *AMIA Annu Symp Proc*, 2016:854–863, 2016.
- [13] C. Martínez-Costa. *Modelos de representación y transformación para la interoperabilidad semántica entre estándares de Historia Clínica Electrónica basados en arquitectura de modelo dual*. PhD thesis, Universidad de Murcia, 2011.
- [14] C. Martínez-Costa, M. Menárguez-Tortosa, and J. T. Fernández-Breis. An approach for the semantic interoperability of iso en 13606 and openehr archetypes. *J Biomed Inform*, 43(5):736–46, Oct 2010.
- [15] M. Menárguez-Tortosa and J. T. Fernández-Breis. Owl-based reasoning methods for validating archetypes. *Journal of biomedical informatics*, 46(2):304–317, 2013.
- [16] M. Quesada-Martínez. *Methodology for the enrichment of biomedical knowledge resources*. PhD thesis, Depto. de Informática y Sistemas. Univ. de Murcia, 2015.
- [17] M. Quesada-Martínez, A. Duque-Ramos, M. Iniesta-Moreno, and J. T. Fernández-Breis. Preliminary analysis of the obo foundry ontologies and their evolution using oquare. *Informatics for Health: Connected Citizen-Led Wellness and Population Health*, 235:426, 2017.
- [18] M. Quesada-Martínez and J. T. Fernández-Breis. Studying the reuse of content in biomedical ontologies: An axiom-based approach. In *Conference on Artificial Intelligence in Medicine in Europe*, pages 3–13. Springer, 2017.
- [19] M. Quesada-Martínez, J. T. Fernández-Breis, and D. Karlsson. Suggesting missing relations in biomedical ontologies based on lexical regularities. *Stud Health Technol Inform*, 228:384–8, 2016.
- [20] M. Quesada-Martínez, J. T. Fernández-Breis, and R. Stevens. Lexical characterization and analysis of the bioportal ontologies. In *Conference on Artificial Intelligence in Medicine in Europe*, pages 206–215. Springer Berlin Heidelberg, 2013.
- [21] M. Quesada-Martínez, J. T. Fernández-Breis, and R. Stevens. Lexical characterisation of bio-ontologies by the inspection of regularities in labels. *Current Bioinformatics*, 10(2):165–176, 2015.
- [22] M. Quesada-Martínez, M. Marcos, F. Abad-Navarro, B. Martínez-Salvador, and J. T. Fernández-Breis. Towards the semantic enrichment of computer interpretable guidelines: a method for the identification of relevant ontological terms. In *AMIA Annual Symposium Proceedings*, 2018.
- [23] M. Quesada-Martínez, E. Mikroyannidi, J. T. Fernández-Breis, and R. Stevens. Approaching the axiomatic enrichment of the gene ontology from a lexical perspective. *Artificial intelligence in medicine*, 65(1):35–48, 2015.
- [24] V. Stroetman, D. Kalra, P. Lewalle, A. Rector, J. Rodrigues, K. Stroetman, G. Surjan, B. Ustun, M. Virtanen, and P. Zanstra. Semantic Interoperability for Better health and Safer Healthcare [34 pages]. (January), 2009.
- [25] A. Third. Hidden semantics: what can we learn from the names in an ontology? In *Proceedings of the Seventh International Natural Language Generation Conference*, pages 67–75. Association for Computational Linguistics, 2012.
- [26] P. van Damme, M. Quesada-Martínez, R. Cornet, and J. T. Fernández-Breis. From lexical regularities to axiomatic patterns for the quality assurance of biomedical terminologies and ontologies. *J Biomed Inform*, Jun 2018.