

Universidad Técnica Particular de Loja

Ciencias de la Computación

Integrantes:

* Leonardo Aguilar
* Royer Masache

Componente:

Ciencia de Datos

Ciclo académico:

octubre 2021 / febrero 2022



# Consumo, enriquecimiento y visualización Scientific Knowledge Graph

## **Scientific Knowledge Graph**

Los grafos de conocimiento son una gran red de entidades y relaciones generalmente expresados en tripletas RDF y que son relevantes en temas específicos y organizaciones. Partiendo de esa definición, Buscaldi et al. (2019) afirman que “Los grafos de conocimiento científico (SKG) se centran en el dominio académico y, por lo general, contienen metadatos que describen publicaciones de investigación como autores, lugares, organizaciones, temas de investigación y citas”.

Buscaldi et al. (2019) también manifiestan las siguientes limitaciones acerca de este tipo de grafos:

Típicamente representan el contenido de los trabajos como texto no estructurado (título, resumen, a veces el texto completo). Por tanto, un desafío importante en este campo se refiere a la generación de SKG que contienen también una representación explícita del conocimiento presentado en las publicaciones y potencialmente describe entidades tales como enfoques, aplicación, formatos, etc. La mayor parte de la información relevante para completar dicho KG podría derivarse del texto de las publicaciones de investigación. Sin embargo, la integración de esta información en un gráfico de conocimiento coherente sigue siendo un desafío abierto.

## **ENDPOINTS**

Existen diversas organizaciones que se encargan de crear grandes conjuntos de datos RDF con información sobre publicaciones científicas y entidades relacionadas entre las que cabe destacar a **Microsoft Academic Knowledge Graph (MAKG) y Scholary Data**, cuyos datos pueden ser consultados a través de SPARQL y sus endpoints específicos.

Diagram

Description automatically generatedA continuación se presenta un ejemplo de esquema de un SKG tomado de <https://makg.org/schema-linked-dataset-descriptions/>

Graphical user interface, text, application, email

Description automatically generatedEl **endpoint** público de SPARQL de Microsoft Academic Knowledge Graph se proporciona en <https://makg.org/sparql>

Graphical user interface, text, application, email

Description automatically generatedEl **endpoint** de Scholary Data donde es posible consultar datos a través de SPARQL se proporciona en <http://www.scholarlydata.org/sparql/>

## **Aplicaciones, proyectos, tecnologías**

Existe un amplio y diversos campos para las aplicaciones de los gráficos de conocimiento científico, entre las que Auer (2020) destaca que los SKG se centran en el dominio académico y el conocimiento de la investigación (temas de investigación, tareas, tecnologías) en este espacio, así como sus relaciones recíprocas. Se han construido grafos de conocimiento científico ampliamente en ciencias como la biomedicina y la informática.

Existe una variedad de proyectos y tecnologías relacionados con el uso de SKG, entre los que cabe destacar **Whyis,** que según McCusker et.al (s.f) “es el primer framework para crear gráficos personalizados de conocimiento basados en la procedencia. Los gráficos de conocimiento de Whyis se basan sobre nanopublicaciones, que simplifica y estandariza la producción de conocimiento estructurado y basado en la procedencia en gráficos de conocimiento”

McCusker et.al (s.f) también afirma que no de los proyectos que se ha desarrollado a través de este framework es:

**BioKG,** un gráfico de conocimiento de biología probabilística, poblado con contenido de proteínas y medicamentos bien utilizados de las ontologías DrugBank, Uniprot y OBO Foundry. Los desarrolladores de gráficos de conocimiento pueden usar Whyis para configurar canalizaciones personalizadas de conservación de conocimiento utilizando importadores de datos y scripts de carga, transformación y extracción semántica. Whyis también contiene una capacidad de metaanálisis de conocimiento para usar en la exploración de gráficos personalizables La arquitectura flexible basada en nanopublicaciones de Whyis permite graficar el conocimiento los desarrolladores Diagram

Description automatically generatedintegran, amplían y publican conocimientos de heterogéneos fuentes en la web.

Otro proyecto que cabe destacar es **cLODg2**, que, tal como afirma un Generador de Datos Abiertos Vinculados que implementa una metodología para producir Linked Data para describir una conferencia científica y sus publicaciones, participantes y eventos. Para generar metadatos de conferencias asumimos que tiene datos iniciales disponibles (con algún formato específico).

Diagram

Description automatically generated

Puede encontrar su repositorio en GitHub aquí <https://github.com/anuzzolese/cLODg2>

Dentro de las tecnologías existentes McCusker et al. (s.f) menciona que existen diversos frameworkes tales como:

“Stardog que incluye razonamiento OWL, mapeo de silos de datos en RDF y reglas personalizadas. Ontowiki proporciona una interfaz de usuario además de una base de datos RDF que rastrea el historial, permite a los usuarios navegar y editar conocimientos, y admite extensiones de interfaz de usuario. Callimachus permite a los desarrolladores crear interfaces de usuario por tipo de objeto mediante RDFa. Virtuoso Openlink Data Spaces es una herramienta de publicación de datos vinculados que proporciona un conjunto de datos predefinidos, herramientas de importación y un conjunto fijo de vistas sobre los datos vinculados que crea. Vitro6 admite la creación de nuevas clases e instancias de ontología, pero no permite que los usuarios creen interfaces personalizadas”

## BIBLIOGRAFÍA:

Buscaldi, D. et al. (2019). Mining Scholarly Publications for Scientific Knowledge Graph Construction. [*http://oro.open.ac.uk/61160/1/ESWC2019-Poster.pdf*](http://oro.open.ac.uk/61160/1/ESWC2019-Poster.pdf)

McCusker et al. (s.f). Developing Scientific Knowledge Graphs Using Whyis. [*http://ceur-ws.org/Vol-2184/paper-06.pdf*](http://ceur-ws.org/Vol-2184/paper-06.pdf)

Auer, S (25 de agosto del 2020). *SCIENTIFIC KNOWLEDGE GRAPHS.* SKG202. <https://skg.kmi.open.ac.uk/SKG2020/>