

RDF: CONCEPTOS Y SINTAXIS

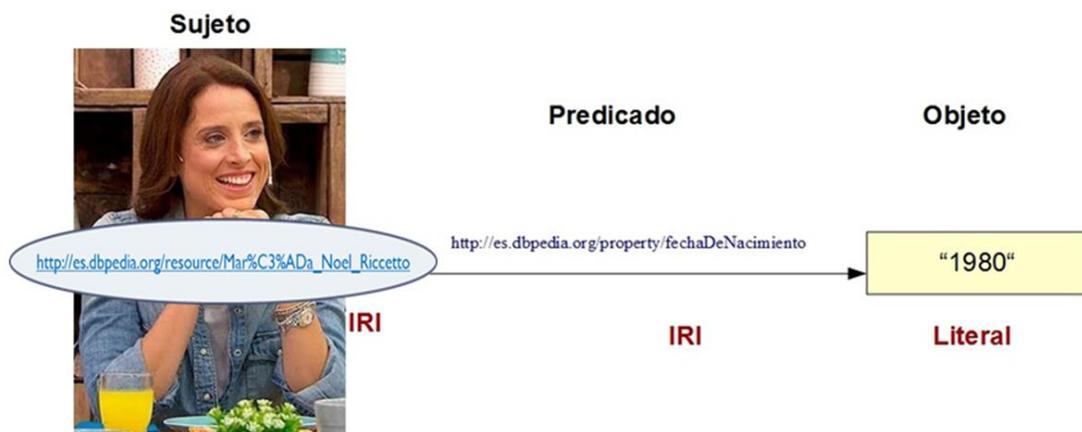
MATERIAL OBLIGATORIO: RDF 1.1 Concepts and Abstract Syntax
 W3C Recommendation 25 February 2014 <https://www.w3.org/TR/rdf11-primer/>

2.1 Definición del Modelo Abstracto de RDF

El conjunto de términos RDF consiste del conjunto de IRIs I , el conjunto de nodos en blanco B y el conjunto de nodos RDF literales L , siendo los tres conjuntos disjuntos entre sí.

Una tupla $\langle s, p, o \rangle$ perteneciente al conjunto $(I \cup B) \times I \times (I \cup B \cup L)$ se llama RDF Tripla o RDF Sentencia, donde s es el sujeto, p es el predicado y o es el objeto. Al conjunto de triplas se le llama RDF Grafo.

Ejemplo Tripla/Sentencia RDF que modela la siguiente realidad: **María Noel Riccetto nació en 1980.**



Los nodos literales (L) describen valores de datos. Como no tienen una existencia separada no precisan ser identificados de forma uniforme. En la representación del grafo de la figura anterior utilizamos recuadros para diagramarlos.

Los literales permiten representar strings (con o sin etiquetas de idioma) y otros valores de tipo de datos (enteros, fechas, etc.).

Los tipos de datos pueden expresarse mediante tipos de datos de XML Schema.

El espacio de nombres para literales en XML Schema es: <http://www.w3.org/2001/XMLSchema#>

EJEMPLOS:

“Semantics”^^<<http://www.w3.org/2001/XMLSchema#string>>

“1161.00”^^<<http://www.w3.org/2001/XMLSchema#float>>

“2020-03-02”^^<<http://www.w3.org/2001/XMLSchema#date>>

Ejemplos de etiquetas de idioma para los strings:

“Semantik”@de ---> Semantik está escrito en alemán

“Semantics”@en ---> Semantics está escrito en inglés

“Semántica”@es ---> Semántica está escrito en español

* Ver sobre tipos de datos de XML Schema en: <https://www.w3.org/TR/rdf11-concepts/>

Los nodos en blanco (B) (o también llamados existenciales) son nodos que denotan la existencia de un individuo con atributos específicos pero sin proveer un identificador o una referencia. Son nodos anónimos ya que no se les asigna un identificador (por ejemplo, en lugar de crear un identificador interno, tenemos la opción de usar nodos en blanco) y son utilizados para modelar información incompleta o compleja.

Ejemplo de información incompleta:

Las conferencias LACLO-2020 y CLEI-2020 se organizan juntas, por lo tanto se realizan en conocida. A pesar de no conocer la fecha se modela usando un nodo en blanco para la fecha la realidad de que coinciden en la fecha como lo muestra la Figura 2.

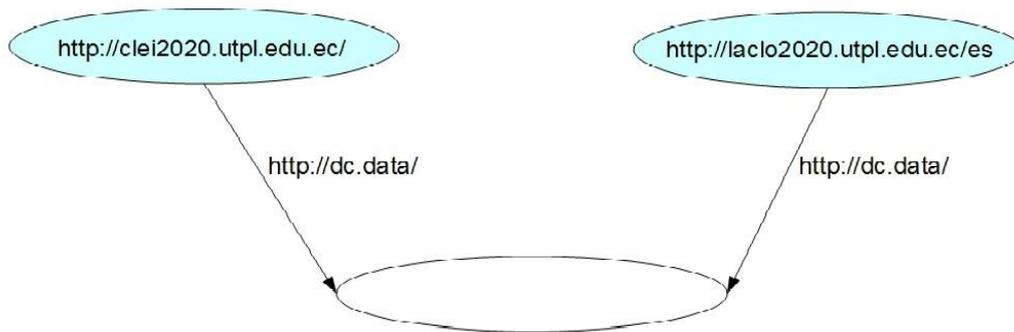


Fig.2 : Ejemplo de uso de nodo en blanco para información incompleta.

Una opción para este ejemplo sería simplemente omitir las propiedades de las fechas, en cuyo caso perdemos la información de que estos eventos se realizan en una fecha y que esas fechas coinciden. Otra opción podría ser crear una URI nueva que represente la fecha, pero semánticamente esto se vuelve indistinguible de que haya una fecha conocida. Por lo tanto una solución es usar el concepto de nodos existenciales o nodos en blanco, representados aquí como un círculo en blanco.

Ejemplo de información compleja:

Dentro de la conferencia CLEI-2020 se realizan los simposios de Ingeniería de Software, Informática y Sociedad; y Procesos de Negocio, una forma de modelarlo es usar un nodo en blanco para representar el concepto de una colección de simposios como muestra la Figura 3.

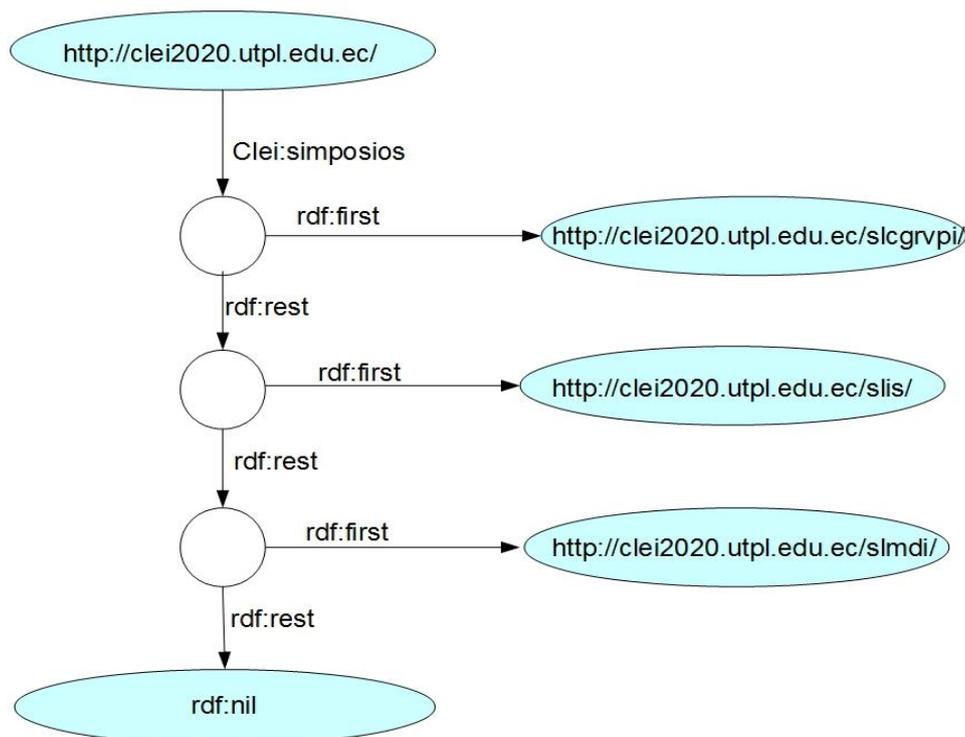


Fig. 3: Ejemplo uso de nodos en blanco para modelar listas en RDF.

Puede haber varios nodos anónimos en un grafo RDF. Cada nodo en blanco tendrá su propio identificador con alcance local al grafo en el cual está definido.

Aunque los nodos existenciales pueden ser convenientes, su presencia puede complicar las operaciones en los grafos, como decidir si dos grafos de datos tienen la misma estructura módulo nodos existenciales [1, 2]. Para evitar estos problemas algunos autores recomiendan aplicar métodos de skolemización en nodos existenciales, reemplazándolos con etiquetas canónicas, propuesta [2, 3]. Otros autores más bien llaman a atención a minimizar el uso de los nodos en blanco [4].

Si te interesa profundizar en el tema de los nodos en blanco te recomendamos la publicación de Hogan *et. al.* [5]. En la sección Referencias encuentras una selección de artículos sobre el tema.

REFERENCIAS:

[1] Richard Cyganiak, David Wood, and Markus Lanthaler. 2014. RDF 1.1 Concepts and Abstract Syntax, W3C Recommendation 25 February 2014. W3C Recommendation. World Wide Web Consortium. <https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>

[2] Aidan Hogan. 2017. Canonical Forms for Isomorphic and Equivalent RDF Graphs: Algorithms for Learning and Labelling Blank Nodes. *ACM Transactions on the Web* 11, 4 (2017), 22:1–22:62. <https://doi.org/10.1145/3068333>

[3] Tom Heath and Christian Bizer. 2011. *Linked Data: Evolving the Web into a Global Data Space (1st Edition)*. Synthesis Lectures on the Semantic Web: Theory and Technology, Vol. 1. Morgan & Claypool. 136 pages.

[4] Dave Longley and Manu Sporny. 2019. RDF Dataset Normalization, A Standard RDF Dataset Normalization Algorithm, Draft Community Group Report 27 February 2019. W3C Community Group Draft Report. <http://json-ld.github.io/normalization/spec/>

[5] Aidan Hogan, Marcelo Arenas, Alejandro Mallea, and Axel Polleres. 2014. Everything you always wanted to know about blank nodes. *Journal of Web Semantics* 27–28 (2014), 42–69. <https://doi.org/10.1016/j.websem.2014.06.004>

[6] Yamaguchi A, Yamamoto Y (2019) Split4Blank: Maintaining consistency while improving efficiency of loading RDF data with blank nodes. *PLOS ONE* 14(6): e0217852. <https://doi.org/10.1371/journal.pone.0217852>