

# Ingeniería de reglas de asociaciones en Grafos de Conocimiento: Identificación de mejoras para lenguajes visuales

Christian Gimenez<sup>1</sup>[0000-0002-8347-2526]

Germán Braun<sup>1</sup>[0000-0003-0769-6680] y Laura Cecchi<sup>1</sup>[0000-0001-5236-6715]

Grupo de Investigación en Lenguajes e Inteligencia Artificial (GILIA)  
Facultad de Informática, Universidad Nacional del Comahue  
{christian.gimenez, lcecchi, german.braun}@fi.uncoma.edu.ar

TESIS DE MAGISTER EN CIENCIAS DE LA COMPUTACIÓN  
(EN DESARROLLO)

**Palabras Clave:** Web Semántica . Grafos de conocimiento . Reglas de asociaciones . *Mapping rules* . Lenguajes visuales

Las ontologías y grafos de conocimientos son ampliamente usados para la Web Semántica y la Inteligencia Artificial [5]. Estas formas de representación de conocimiento se utilizan en la medicina<sup>1</sup> [12,11], la genética<sup>2</sup>, la industria del petróleo [18], la biología marina [19], la ciberseguridad [8], etc. Sin embargo, los sistemas informáticos actuales almacenan la información en bases de datos relacionales, en formatos basados en XML o JSON, en archivos de datos de tipo CSV o TSV, entre otros. Para que estos datos puedan ser usados en la Web Semántica, deben utilizarse motores de asociación, que se conectan con estas fuentes y producen grafos de conocimientos con la estructura requerida por los usuarios. Estos motores, en su mayoría, han prevalecido aquellos que también se alimentan de reglas de asociación (*mapping rules*) para definir cómo generar los elementos de los grafos de conocimiento y qué información de la fuente utilizar [4].

La creación de estas reglas es un proceso realizado habitualmente por ingenieros ontológicos que requiere de diversos conocimientos: el dominio, la estructura de la fuente de datos, la información que los diferentes *stackholders* desean consultar, la ontología y el vocabulario destino, etc. Por consiguiente, el diseño de estas reglas debe poder ser fácil de comunicar entre los ingenieros y a las distintas personas involucradas en el proceso y a los futuros usuarios.

Para realizar estos diseños, se puede utilizar un lenguaje visual ya que facilita la comprensión del modelo. En la ingeniería de software, lo habitual es utilizar lenguajes de este tipo, por ejemplo: UML [15], BPMN [14] y MER [3]. Si bien,

---

<sup>1</sup> Véase <https://www.snomed.org> y <https://bioportal.bioontology.org/ontologies/SNOMEDCT> para más información. Visitado el día 28 de septiembre del 2023.

<sup>2</sup> Véase <https://geneontology.org>. Visitado el día 28 de septiembre del 2023.

existen lenguajes visuales para representar reglas de asociación, estos no son estándares y son implementados por herramientas donde la sintaxis y semántica dependen fuertemente de la utilidad de las mismas. Cabe mencionar que, los lenguajes de modelado de la ingeniería de software, no son utilizados por estos lenguajes visuales, por ejemplo, para representar la fuente de datos.

Por ello, se propone un metodología para identificar mejoras sobre un lenguaje visual para la ingeniería de asociaciones. Se hace énfasis en ampliar la efectividad cognitiva y la expresividad de los mismos para comunicar las reglas de forma visual.

Esta metodología comienza con la descripción del lenguaje visual, describiendo su sintaxis visual y la semántica asociada. Para ello, se requiere relevar y analizar artículos que describan el lenguaje o la herramienta que lo utiliza y sus casos de estudio. Además, se utilizan ejemplos para interpretar la asociación entre cada sentencia del diagrama y su representación en las reglas de asociación.

Luego, se estudia la descripción realizada con respecto a los principios de *Physics of Notations* (PoN) planteados en [13]. Estos, describen criterios que persiguen ampliar la expresividad del lenguaje, reducir ambigüedades, gestionar la complejidad de los diagramas de gran tamaño y proponer dialectos para diferentes audiencias (p. ej.: personas con o sin experiencia en el uso del lenguaje) y medios (p. ej.: digital, papel, impresión, etc.).

Así, una vez analizado el lenguaje, se sugieren posibles modificaciones para que éste cumpla en mayor medida con los principios de PoN. Asimismo, se plantea la forma en que cada opción posible a modificar afecte, no solo al criterio que se hace foco, sino también a los otros, proveyendo recomendaciones para evitar impactos negativos.

En conclusión, se provee un procedimiento que busca describir a un lenguaje, analizarlo según los principios de PoN y, en base a este resultado, proponer mejoras y consideraciones para ampliar la expresividad y la comprensión de los diagramas producidos con el lenguaje. Considerando también el objetivo particular del lenguaje mismo, y a su vez la audiencia a la cual se desea presentar los diagramas.

Para comprender el funcionamiento del procedimiento propuesto se aplica esta metodología a lenguajes visuales existentes, particularmente Map-on [16], SQuaRE [1,2], Juma [10,9], KR2RML/Karma [17] y MapVOWL/RMLEditor [6].

Se espera que la utilización de esta metodología permita la creación de lenguajes más comprensibles y que facilite la comunicación de las reglas de asociación entre los ingenieros y los diferentes *stakeholders*. Asimismo, esta investigación promueve mejoras en herramientas gráficas como crowd<sup>3</sup> para brindar soporte *Ontology-Based Data Access* (OBDA) con el fin de popular esquemas ontológicos con instancias por medio del uso de diseños visuales de reglas de asociación fáciles de entender.

---

<sup>3</sup> <https://crowd.fi.uncoma.edu.ar> visitado el día 30 de octubre del 2023.

## Referencias

1. Bak, J., Blinkiewicz, M.: Square: A visual tool for creating R2RML mappings. In: Kawamura, T., Paulheim, H. (eds.) Proceedings of the ISWC 2016 Posters & Demonstrations Track co-located with 15th International Semantic Web Conference (ISWC 2016), Kobe, Japan, October 19, 2016. CEUR Workshop Proceedings, vol. 1690. CEUR-WS.org (2016), <http://ceur-ws.org/Vol-1690/paper62.pdf>
2. Bak, J., Blinkiewicz, M., Lawrynowicz, A.: User-friendly visual creation of R2RML mappings in square. In: Ivanova et al. [7], pp. 139–150, <http://ceur-ws.org/Vol-1947/paper13.pdf>
3. Chen, P.: Entity-Relationship Modeling: Historical Events, Future Trends, and Lessons Learned, pp. 296–310. Springer (2002), [http://bit.csc.lsu.edu/~chen/pdf/Chen\\_Pioneers.pdf](http://bit.csc.lsu.edu/~chen/pdf/Chen_Pioneers.pdf)
4. Dimou, A.: R2RML and RML comparison for RDF generation, their rules validation and inconsistency resolution. CoRR **abs/2005.06293** (2020), <https://arxiv.org/abs/2005.06293>
5. Harmelen, F.V., Lifschitz, V., Porter, B. (eds.): Handbook of Knowledge Representation. Elsevier (2008)
6. Heyvaert, P., Dimou, A., De Meester, B., Seymoens, T., Herregodts, A.L., Verborgh, R., Schuurman, D., Mannens, E.: Specification and implementation of mapping rule visualization and editing: Mapvowl and the rmleditor. Journal of Web Semantics **49**, 31 – 50 (2018). <https://doi.org/https://doi.org/10.1016/j.websem.2017.12.003>, <http://www.sciencedirect.com/science/article/pii/S1570826817300707>
7. Ivanova, V., Lambrix, P., Lohmann, S., Pesquita, C. (eds.): Proceedings of the Third International Workshop on Visualization and Interaction for Ontologies and Linked Data co-located with the 16th International Semantic Web Conference (ISWC 2017), Vienna, Austria, October 22, 2017, CEUR Workshop Proceedings, vol. 1947. CEUR-WS.org (2017), <http://ceur-ws.org/Vol-1947>
8. Jia, Y., Qi, Y., Shang, H., Jiang, R., Li, A.: A practical approach to constructing a knowledge graph for cybersecurity. Engineering **4** (02 2018). <https://doi.org/10.1016/j.eng.2018.01.004>
9. Junior, A.C., Debruyne, C., O’Sullivan, D.: Juma: An editor that uses a block metaphor to facilitate the creation and editing of R2RML mappings. In: Blomqvist, E., Hose, K., Paulheim, H., Lawrynowicz, A., Ciravegna, F., Hartig, O. (eds.) The Semantic Web: ESWC 2017 Satellite Events - ESWC 2017 Satellite Events, Portorož, Slovenia, May 28 - June 1, 2017, Revised Selected qPapers. Lecture Notes in Computer Science, vol. 10577, pp. 87–92. Springer (2017). [https://doi.org/10.1007/978-3-319-70407-4\\_17](https://doi.org/10.1007/978-3-319-70407-4_17), [https://doi.org/10.1007/978-3-319-70407-4\\_17](https://doi.org/10.1007/978-3-319-70407-4_17)
10. Junior, A.C., Debruyne, C., O’Sullivan, D.: Using a block metaphor for representing R2RML mappings. In: Ivanova et al. [7], pp. 1–12, <http://ceur-ws.org/Vol-1947/paper01.pdf>
11. Liu, Y., Hildebrandt, M., Joblin, M., Ringsquandl, M., Raissouni, R., Tresp, V.: Neural multi-hop reasoning with logical rules on biomedical knowledge graphs. In: Verborgh, R., Hose, K., Paulheim, H., Champin, P.A., Maleshkova, M., Corcho, O., Ristoski, P., Alam, M. (eds.) The Semantic Web. pp. 375–391. Springer International Publishing, Cham (2021)
12. Mann, M., Ilievski, F., Rostami, M., Aastha, Shbita, B.: Open drug knowledge graph. In: Chaves-Fraga, D., Dimou, A., Heyvaert, P., Priyatna, F., Sequeda, J.

- (eds.) Proceedings of the 2nd International Workshop on Knowledge Graph Construction co-located with 18th Extended Semantic Web Conference (ESWC 2021) (2021), <https://ceur-ws.org/Vol-2873/>
13. Moody, D.: The "physics" of notations: Toward a scientific basis for constructing visual notations in software engineering. *IEEE Trans. Softw. Eng.* **35**(6), 756–779 (Nov 2009). <https://doi.org/10.1109/TSE.2009.67>, <https://doi.org/10.1109/TSE.2009.67>
  14. Object Management Group: Business process model and notation (BPMN). Tech. rep., Object Management Group (2011), <http://www.omg.org/spec/BPMN/2.0/>, versión 2.0
  15. Object Management Group: OMG unified modeling language (OMG UML). Tech. rep., Object Management Group (2017), versión 2.5.1
  16. Sicilia, Á., Nemirovski, G., Nolle, A.: Map-on: A web-based editor for visual ontology mapping. *Semantic Web* **8**(6), 969–980 (2017). <https://doi.org/10.3233/SW-160246>, <https://doi.org/10.3233/SW-160246>
  17. Slepicka, J., Yin, C., Szekely, P.A., Knoblock, C.A.: KR2RML: an alternative interpretation of R2RML for heterogenous sources. In: Hartig, O., Sequeda, J.F., Hogan, A. (eds.) Proceedings of the 6th International Workshop on Consuming Linked Data co-located with 14th International Semantic Web Conference (ISWC 2015), Bethlehem, Pennsylvania, USA, October 12th, 2015. CEUR Workshop Proceedings, vol. 1426. CEUR-WS.org (2015), <http://ceur-ws.org/Vol-1426/paper-08.pdf>
  18. Soyulu, A., Kharlamov, E., Zheleznyakov, D., Jiménez-Ruiz, E., Giese, M., Skjæveland, M.G., Hovland, D., Schlatte, R., Brandt, S., Lie, H., Horrocks, I.: Optiquevqs: A visual query system over ontologies for industry. *Semantic Web* **9**(5), 627–660 (2018). <https://doi.org/10.3233/SW-180293>, <https://doi.org/10.3233/SW-180293>
  19. Zárate, M.D., Braun, G.A., Fillottrani, P.R., Delrieux, C., Lewis, M.: *BiGe-Onto*: An ontology-based system for managing biodiversity and biogeography data. *Appl. Ontology* **15**(4), 411–437 (2020). <https://doi.org/10.3233/AO-200228>, <https://doi.org/10.3233/AO-200228>