

# Keynote: Powering Biomedical Artificial Intelligence with a Holistic Knowledge Graph

Catia Pesquita<sup>1</sup>

<sup>1</sup>LASIGE, Faculdade de Ciencias da Universidade de Lisboa

## Abstract

Biomedical AI applications increasingly rely on multi-domain and heterogeneous data, especially in areas such as personalised medicine and systems biology. Biomedical Ontologies are a golden opportunity in this area because they add meaning to the underlying data which can be used to support heterogeneous data integration, provide scientific context to the data augmenting AI performance, and afford explanatory mechanisms allowing the contextualization of AI predictions. In particular, ontologies and knowledge graphs support the computation of semantic similarity between objects, providing an understanding of why certain objects are considered similar or different. This is a basic aspect of explainability and is at the core of many machine learning applications. However, when data covers multiple domains, it may be necessary to integrate different ontologies to cover the full semantic landscape of the underlying data. The talk presented recent work on building an integrated knowledge graph that is based on the semantic annotation and interlinking of heterogeneous data into a holistic semantic landscape that supports semantic similarity assessments. The talk discussed the challenges in building the knowledge graph from public resources [1], the methodology we are using [2] and the road-ahead in biomedical ontology and knowledge graph alignment as AI becomes an integral part of biomedical research [3, 4, 5, 6].

**Acknowledgements.** CP is funded by the FCT through LASIGE Research Unit (ref. UIDB/00408/2020 and ref. UIDP/00408/2020), and also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017453.

## References


- [1] J. D. Ferreira, D. C. Teixeira, C. Pesquita, Biomedical ontologies: coverage, access and use (2021).
- [2] M. C. Silva, D. Faria, C. Pesquita, Matching multiple ontologies to build a knowledge graph for personalized medicine, in: European Semantic Web Conference, Springer, 2022, pp. 461–477.
- [3] R. T. Sousa, S. Silva, C. Pesquita, Evolving knowledge graph similarity for supervised learning in complex biomedical domains, BMC bioinformatics 21 (2020) 1–19.
- [4] C. Pesquita, Towards semantic integration for explainable artificial intelligence in the biomedical domain., in: HEALTHINF, 2021, pp. 747–753.


---

*SeWebMeDa-2022: 5th International Workshop on Semantic Web solutions for large-scale biomedical data analytics, May 29, 2022, Hersonissos, Greece*

✉ [clpesquita@ciencias.ulisboa.pt](mailto:clpesquita@ciencias.ulisboa.pt) (C. Pesquita)

ORCID [0000-0002-1847-9393](https://orcid.org/0000-0002-1847-9393) (C. Pesquita)

 © 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

- [5] S. Nunes, R. Sousa, C. Pesquita, Predicting gene-disease associations with knowledge graph embeddings over multiple ontologies. arxiv 2021, arXiv preprint arXiv:2105.04944 (????).
- [6] R. M. Carvalho, D. Oliveira, C. Pesquita, Knowledge graph embeddings for icu readmission prediction, BMC Medical Informatics and Decision Making 23 (2023) 12.