

FROM BIOMEDICAL DATA TO KNOWLEDGE

Rita T. Sousa
rita.sousa@uni-mannheim.de

Academic Speed Dating on Data Science





CHALLENGES IN BIOMEDICAL DATA INTEGRATION

Databases



- Fragmentation.
- Different models.
- Different identifiers.

Literature

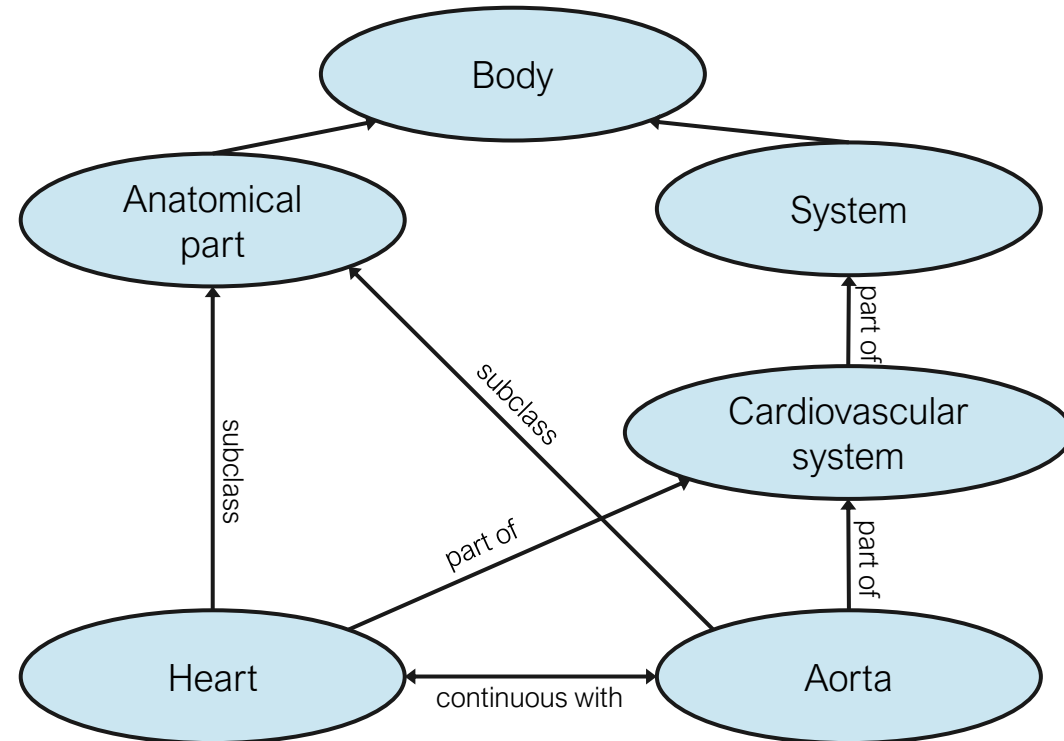


- Natural language is ambiguous.
- Computational analysis is complex.

STRUCTURING KNOWLEDGE IN ONTOLOGIES

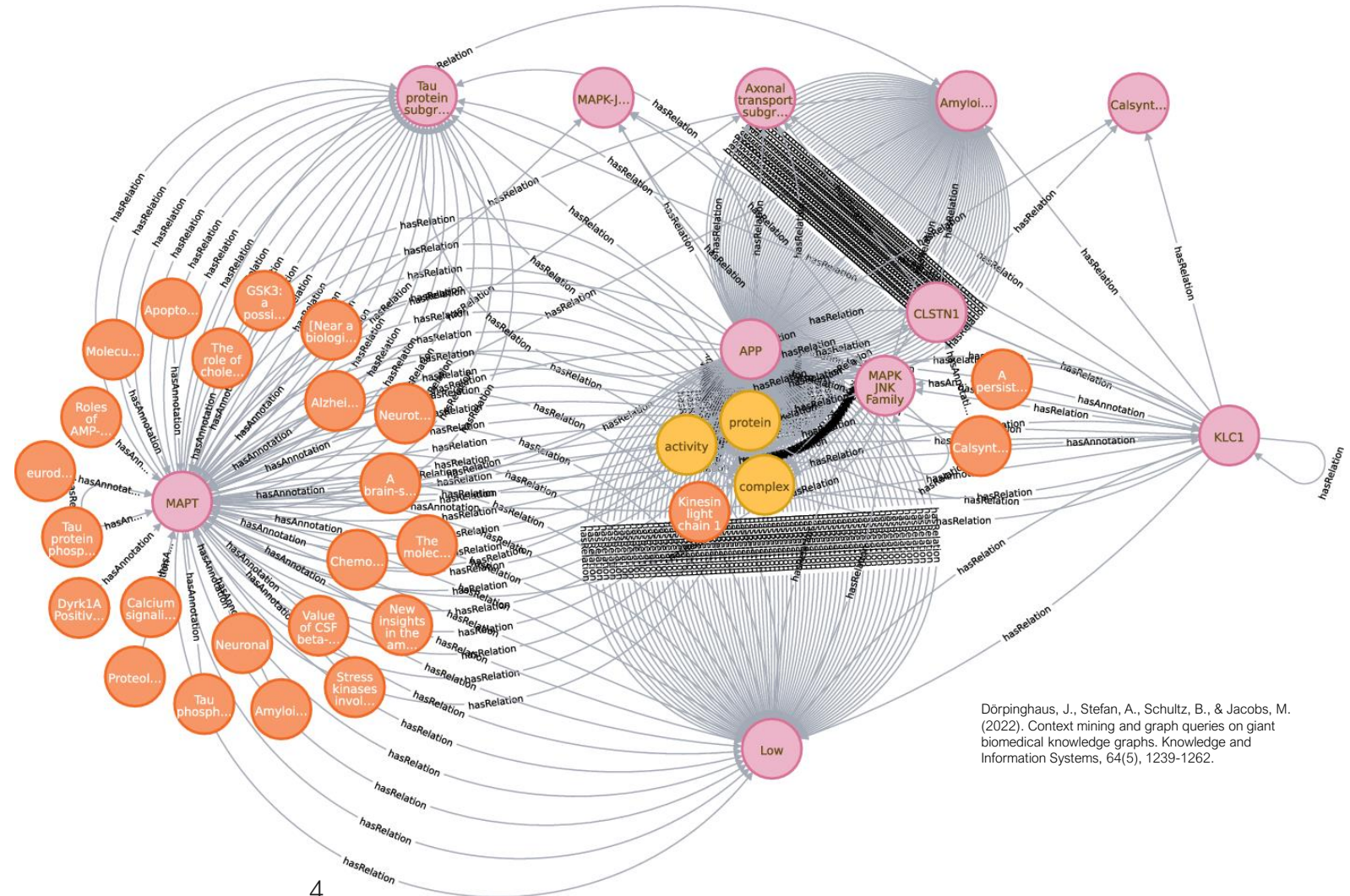
“Ontologies are a technique used to represent and share knowledge about a domain, by modeling the entities existing in that domain and the relationships between them.” (Bodenreider & Stevens, 2006)

- Allow for complex relationships and logical rules.
- Model the semantics (meaning) of a domain.
- Understood by humans and machines.



BIOMEDICAL ONTOLOGIES AND KNOWLEDGE GRAPHS

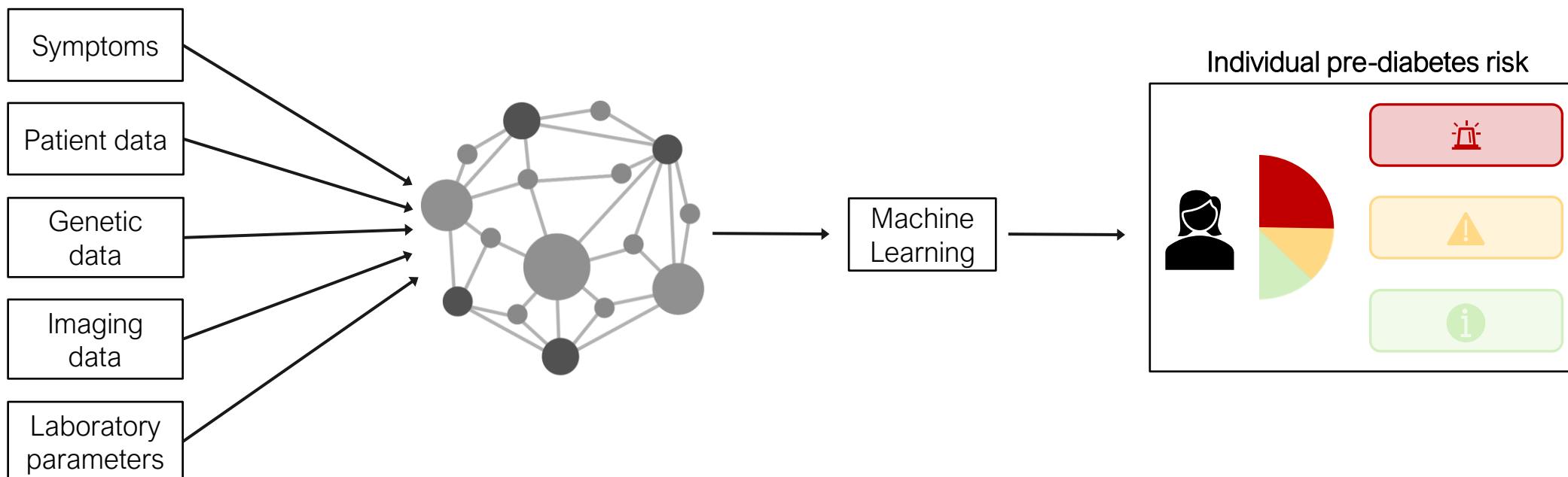
- There are multiple biomedical ontologies that can be incorporated into knowledge graphs.
- Knowledge graphs have been explored for many biomedical applications such as finding new treatments for existing drugs, diagnosing patients, identifying associations between diseases and genes, predicting interactions between proteins, etc.



Dörpinghaus, J., Stefan, A., Schultz, B., & Jacobs, M. (2022). Context mining and graph queries on giant biomedical knowledge graphs. *Knowledge and Information Systems*, 64(5), 1239-1262.

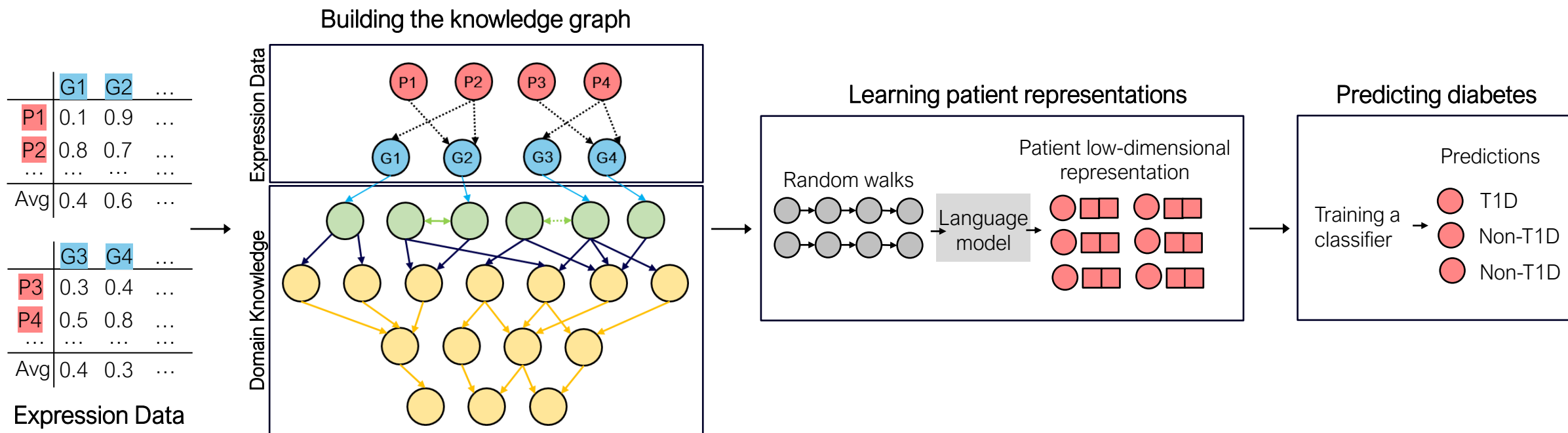
KI-DIABETES DETECTION PROJECT

The goal is to integrate data from various sources in a knowledge graph and apply machine learning methods to improve the early-stage detection of Diabetes.



HANDLING GENE EXPRESSION DATA

Novel approach that integrates heterogeneous gene expression data through knowledge graphs for improving diabetes prediction.





THANK YOU FOR YOUR ATTENTION.



rita.sousa@uni-mannheim.de



[@RitaTorresSousa](https://twitter.com/RitaTorresSousa)



<https://ritatsousa.github.io/>

Academic Speed Dating on Data Science

