Keynote: Towards Biomedical Neurosymbolic AI: From Knowledge Infrastructure to Explainable Predictions

Michel Dumontier[1]

[1]Maastricht University, Maastricht, Limburg, NL

Abstract

The increased availability of biomedical data, particularly in the public domain, offers the opportunity to better understand human health and to develop effective therapeutics for a wide range of unmet medical needs. However, data scientists remain stymied by the fact that data remain hard to find and to productively reuse because data and their metadata i) are wholly inaccessible, ii) are in non-standard or incompatible representations, iii) do not conform to community standards, and iv) have unclear or highly restricted terms and conditions that preclude legitimate reuse. These limitations require a rethink on how data can be made machine and AI-ready - the key motivation behind the FAIR Guiding Principles. Concurrently, while recent efforts have explored the use of deep learning to fuse disparate data into predictive models for a wide range of biomedical applications, these models often fail even when the correct answer is already known, and fail to explain individual predictions in terms that data scientists can appreciate. These limitations suggest that new methods to produce practical artificial intelligence are still needed.

In this talk, I will discuss our work in (1) building an integrative knowledge infrastructure to prepare FAIR and "AI-ready" data and services along with (2) neurosymbolic AI methods to improve the quality of predictions and to generate plausible explanations. Attention is given to standards, platforms, and methods to wrangle knowledge into simple, but effective semantic and latent representations, and to make these available into standards-compliant and discoverable interfaces that can be used in model building, validation, and explanation. Our work, and those of others in the field, create a baseline for building trustworthy and easy-to-deploy AI models in biomedicine.

SeWebMeDa-2024: 7th International Workshop on Semantic Web solutions for large-scale biomedical data analytics, May 26, 2024, Hersonissos, Greece

michel.dumontier@maastrichtuniversity.nl (M. Dumontier)
https://www.maastrichtuniversity.nl/mj-dumontier (M. Dumontier)
0000-0003-4727-9435 (M. Dumontier)