Abstract Authors:

John A. Bachman, Benjamin M. Gyori

Abstract Text:

Computational models representing interconnected systems are a critical tool for understanding complex phenomena and for designing effective interventions. However, the routine use of models in scientific discovery and decision-making has been limited by the difficulties involved in maintaining models with up-to-date information, and performing analysis with respect to scientific questions. Each of the 3,000 new articles published daily in biomedicine contains fragments of information that, when assembled, could lead to surprising novel conclusions. However, no human can systematically monitor this flood of information, let alone interpret the impact of each finding in the context of prevailing knowledge.

The Ecosystem of Machine-Maintained Models with Automated Analysis (EMMAA) is a framework to automatically maintain, validate and analyze scientific models built from literature. EMMAA builds and updates models incrementally by proactively monitoring newly published articles, processing them with multiple natural language processing systems and assembling the new knowledge into a set of models with an initial focus on cancer biology. Each newly updated model is then systematically checked for its ability to causally explain a suite of empirical observations. EMMAA addresses these challenges by coupling recent advances in the automated assembly of models from literature to a model analysis system that identifies possible causal explanations for observations derived from data. Together, these two capabilities allow scientists to move from skimming scientific abstracts as they appear to gaining a systematic view of how newly published discoveries are impacting the state of knowledge in their field.