KnowledgeSphere: An Automated and Integrative Framework for Drug Repurposing Empowered by Knowledge Graph and AI

Liang-Chin Huang[1], Ying Li[2], Jingcheng Du[1], Kyeryoung Lee[1], Jingqi Wang[1], Frank Manion[1], Xiaoyan Wang[3]

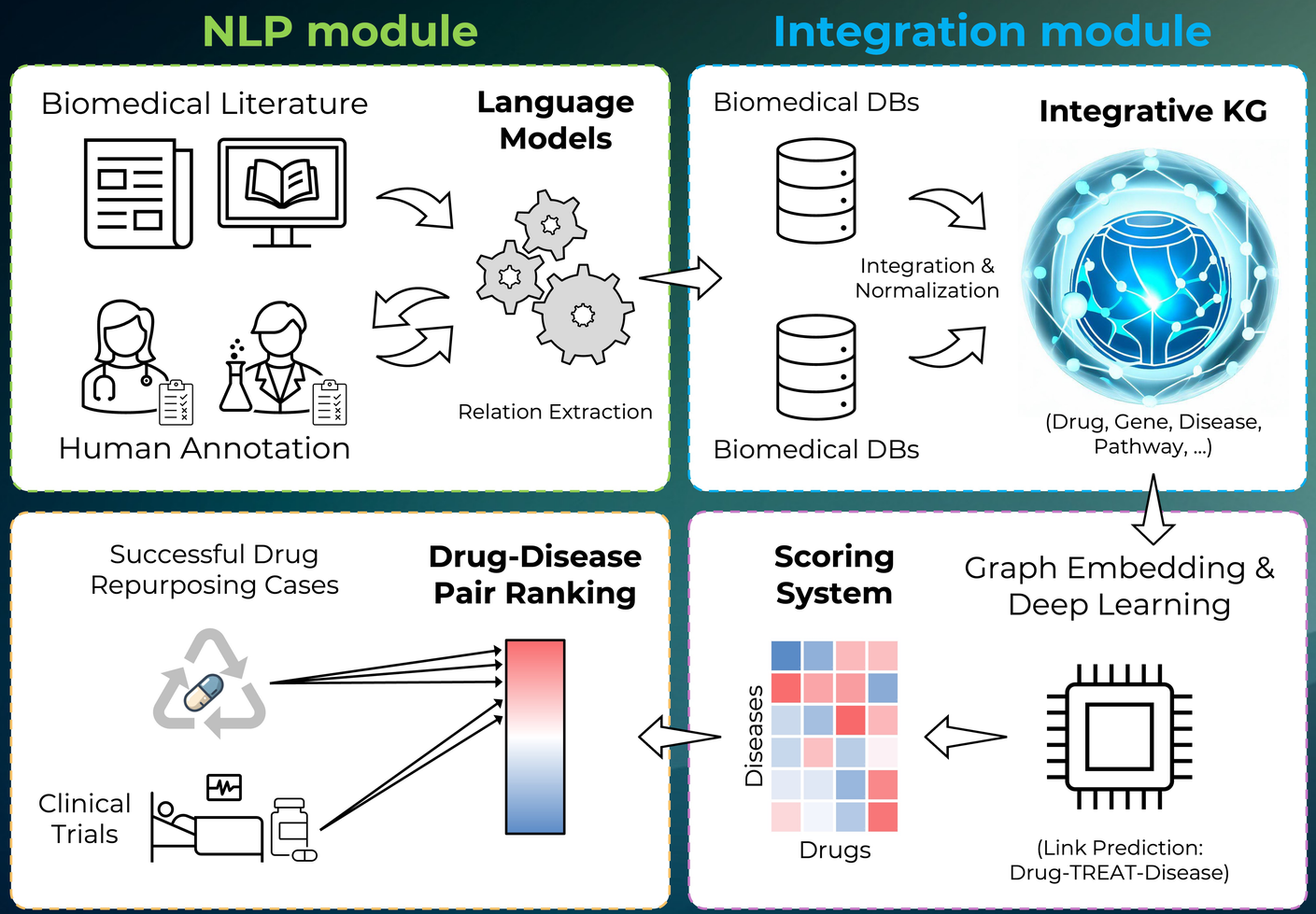
[1]Melax Tech, Houston, TX, USA, [2]Regeneron Pharmaceuticals, Tarrytown, NY, USA, [3]Melax Tech, Stamford, CT, USA

OBJECTIVES

De novo drug development is an extremely time-consuming process with a high risk of failure and tremendous resource requirements. For these reasons, a drug-repurposing strategy has gained significant attention as a cost-effective alternative strategy. Drug repurposing can rapidly repurpose FDA-approved drugs to other indications than the approved indications. Until now, drug repurposing is serendipitous, such as repurposing Sildenafil for erectile dysfunction, and sometimes requires extensive manual reviews of related literature, clinical trials, and relevant clinical data. With an ever-growing amount of literature and information, an automatic method is much needed.

METHODS

We developed KnowledgeSphere, a framework for processing and integrating diverse knowledge bases. It includes four modules: i) NLP module: building natural language processing (NLP) pipelines to extract biomedical knowledge from literature, ii) integration module: incorporating data from various sources, iii) knowledge module: building deep learning-based models and scoring systems, and iv) evaluation module: validating results.



RESULTS

- First, we built NLP pipelines (BERT-based models refined by utilizing human annotations) to extract biomedical entities and relations from **35 million PubMed abstracts**.
- Second, we integrated the data from manually curated biomedical resources into our literature-based knowledge graph. As a result, our knowledge graph consists of **20 thousand entities** (drugs, diseases, genes, etc.) and **10 million relations** ("inhibits", "treats", "stimulates", etc.).
- Third, we trained deep learning-based knowledge graph embedding models and then **predicted the "treats" relations** for each drug-disease pair.
- Finally, in our evaluation module, we excluded all relations involving **14 successful drug repurposing cases** [1] collected from review articles during knowledge graph embedding. After applying link prediction for all 14 successful pairs of drugs and their new indications, we found that all are **ranked top 0.5%** across all diseases.

Year	Drug name	Original indication	New indication	Score	Rank (%)
1987	Zidovudine	Malignant Neoplasms	Acquired Immunodeficiency Syndrome	6.774	0.20%
1988	Minoxidil	Hypertensive disease	Alopecia	8.154	0.02%
1998	Sildenafil	Angina Pectoris	Erectile dysfunction	6.252	0.24%
1998	Thalidomide	Morning Sickness	Erythema nodosum leprosum	2.926	0.48%
2000	Celecoxib	Inflammatory pain	Adenomatous Polyposis Coli	6.030	0.26%
2002	Atomoxetine	Parkinson Disease	Attention deficit hyperactivity disorder	6.948	0.19%
2004	Duloxetine	Mental Depression	Urinary Stress Incontinence	3.780	0.42%
2006	Thalidomide	Morning Sickness	Multiple Myeloma	7.983	0.10%
2007	Raloxifene	Osteoporosis	Malignant neoplasm of breast	5.920	0.26%
2010	Fingolimod	Graft Rejection	Multiple Sclerosis	9.182	0.01%
2012	Dapoxetine	Central nervous system depression	Premature Ejaculation	3.312	0.45%
2012	Topiramate	Epilepsy	Obesity	7.402	0.14%
2014	Ketoconazole	Mycoses	Cushing Syndrome	7.892	0.11%
2015	Aspirin	Analgesia	Colorectal Carcinoma	5.565	0.29%

CONCLUSIONS

In conclusion, KnowledgeSphere, a framework leveraging deep learning-based knowledge graph embedding models, enables the representation of complex and interconnected knowledge from various sources and potentially opens the way for drug repurposing at scale.

REFERENCE

1. Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C. and Norris, A., 2019. Drug repurposing: progress, challenges and recommendations. *Nature reviews Drug discovery*, 18(1), pp.41-58.

Please get in touch with the corresponding author via email: xiaoyan.wang@melaxtech.com