Constructing Knowledge Graphs of Depression (Extended Abstract)*

Zhisheng Huang[1], Jie Yang[2], Frank van Harmelen[1], and Qing Hu[1]

[1] VU University Amsterdam, the Netherlands
[2] Beijing Anding Hospital, China

1 Introduction

Major depressive disorder (MDD) has become a serious problem in modern society. Using antidepressants has been considered the dominant treatment for MDD. Psychiatric doctors confront the challenge to make clinical decision efficiently by gaining a comprehensive analysis over various knowledge resources about depression.

In this paper we propose an approach to constructing a knowledge graph of depression using semantic web technology to integrate those knowledge resources, achieving a high degree of inter-operability. With a single semantic query over integrated knowledge resources, psychiatric doctors can be much more efficient in finding answers to queries which currently require them to explore multiple databases and to make a time-consuming analysis on the results of those searches.

The term "Knowledge Graph" is widely used to refer to a large scale semantic network of entities and concepts plus the semantic relationships among them. Such knowledge graphs are very *a-specific* in terms of the diseases that they cover, and are often prohibitively large, hampering both efficiency for machines and usability for people. In this paper we propose an approach to the construction of *disease-centric knowledge graphs*,which are focussed on a single disease or coherent group of diseases.

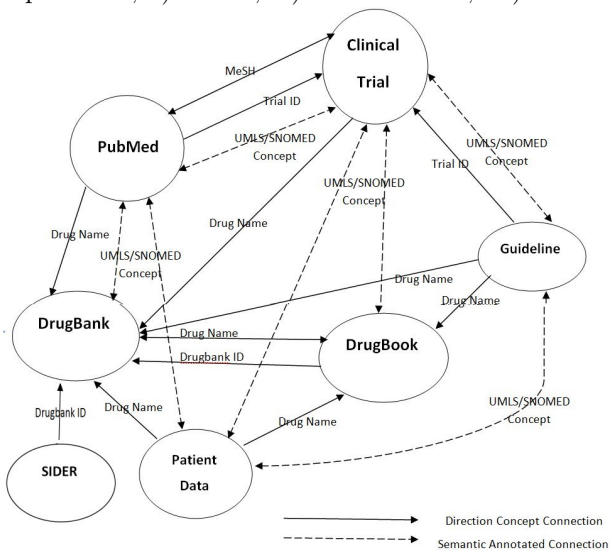
We illustrate our general idea by integrating various knowledge resources about depression (e.g., clinical trials, antidepressants, medical publications, clinical guidelines, etc.). We call the generated knowledge graph *DepressionKG* for short. DepressionKG is represented in RDF/NTriple format [2]. DepressionKG provides a data infrastructure to explore the relationship among various knowledge and data-sources about depression. We show how it provides support for clinical question answering and knowledge browsing.

2 Knowledge Resources and Integration

Following commonly used technology, we construct our knowledge graph as an RDF graph The current version of DepressionKG (version 0.60) consists of the

* The full version of this paper appears in the proceedings of 2017 International Conference on Health Information Science

RDF representation of the following knowledge resources: i)Clinical trials; i-i)PubMed; iii) Guidelines; iv) Drugbank; v)Drugbook; vi)Wikipedia Side Effects of antidepressants; v)SIDER; vi)SNOMED CT; vii)Patient Data.



We use the following methods to integrate the various knowledge resources. (i) *Direct Entity identification*; (ii) *Direct Concept identification*; (iii) *Semantic Annotation with an NLP tool*; and (iv) *Semantic Queries with regular expressions*. The figure shows the connectivity of DepressionKG.

3 Conclusions

In this paper, we have proposed an approach to making a knowledge graph of depression, and we have shown how various knowledge resources concerning depression can be integrated for semantic inter-operability. We have provided several use cases for such a knowledge graph of depression. From those use cases, we can see that by using a knowledge graph with its semantic search, it is rather convenient for us to detect relationship which cover multiple knowledge resources. Our conclusions are (i) that it is indeed possible to make disease-centric subgraphs without having to include the entire original graph, and (ii) that realistic clinical queries can still be answered over such disease-specific knowledge graphs without substantial loss of recall.

References

1. S. Ait-Mokhtar, B. D. Bruijn, C. Hagege, and P. Rupi. Intermediary-stage ie components, D3.5. Technical report, EURECA Project, 2014.
2. R. Cyganiak, D. Wood, and M. Lanthaler. RDF 1.1 concepts and abstract syntax, 2014.