Babylon Health's Medical Knowledge Graph: Why, What, and How

Claudia Schulz, Damir Juric, Jetendr Shamdasani, Martin Coste, Szymon Wartak, Aleksandar Savkov, Nils Hammerla, and Mohammed Khodadadi

Babylon Health, London SW3 3DD, UK
https://www.babylonhealth.com

Babylon Health offers digital healthcare services through a mobile application. Besides video consultations with doctors, the app is equipped with a chatbot able to provide medical information and perform triage.

To realise this digital healthcare experience, various services within Babylon generate, exchange, and consume health data and clinical knowledge. For example, text understanding services process the user's chatbot input and recognise relevant medical terms therein. These terms may need to be matched with predefined symptoms and risk factors that form the input to the triaging engine, or with past diagnoses stored in user profiles. Furthermore, consultation notes are labelled with clinical codes to allow for easy processing and retrieval.

To ensure that a uniform medical terminology is used throughout Babylon's services and to allow reasoning over medical data, a large medical knowledge graph was created [1], realised as the materialisation of an RDF triple store.

Building the Babylon Knowledge Graph As illustrated in Figure 1, the Babylon Knowledge Graph was constructed by integrating information from multiple existing medical ontologies. To ensure accurate alignment of these heterogeneous ontologies, we developed a novel iterative approach for ontology integration [3], where SNOMED served as the seed ontology. Our approach tackles the problem of hierarchy incompatibility between ontologies by prioritising the current ontology over the one to be integrated thus minimising the amount of inconsistent mappings being dropped.

We furthermore augmented the integrated ontologies with information extracted from biomedical papers. Since SNOMED and other ontologies mainly provide hierarchical relationships between concepts, this information extraction was focused on other types of relationships between medical concepts, e.g. that between diseases and their symptoms or risk factors.

Knowledge Graph Validation To ensure the syntactic and semantic validity of the Babylon Knowledge Graph, we apply both automatic [4] and manual evaluation methods. Checking for (undesirable) cycles and investigating average path length are two of the straight-forward automatic syntax analyses performed. An example of ensuring semantic validity is to check that there is no label overlap between concepts. A manual evaluation of a sample of the Babylon Knowledge Graph was performed with the help of doctors, focusing on the correctness of the hierarchy and concept labels.

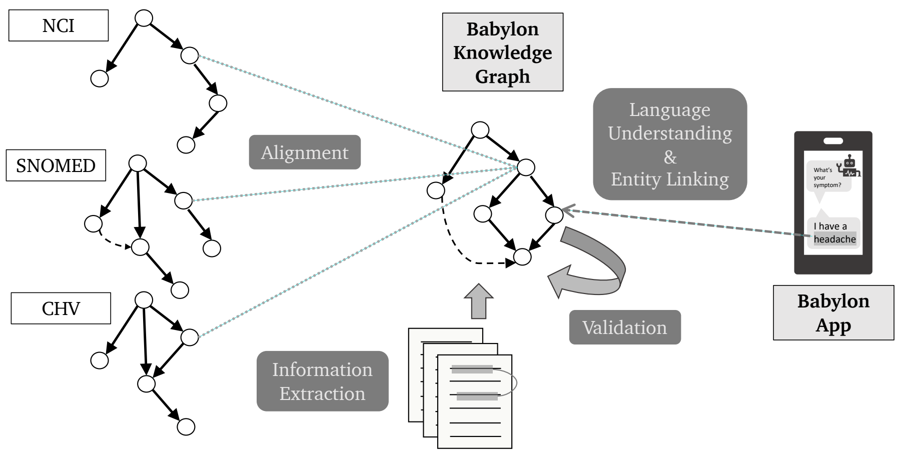

Fig. 1. Creation and usage of the Babylon Knowledge Graph

Knowledge Graph Powered Triage To enable the understanding and processing of user input in a chatbot conversation, we employ various Natural Language Processing (NLP) techniques: dependency parsing combined with entity linking identifies medical terms in text and links them to concepts in the Babylon Knowledge Graph. These concepts are then matched with symptoms and risk factors represented in Babylon's triage system. However, the combination of multiple concepts in the user input may correspond to a single concept in the triage system and triage concepts may be more general than the very specific symptoms provided by a user. For example, the concepts 'severe pain' and 'forehead' may have been identified in the user input, which have to be matched to the single and more general triage concept 'headache'. To realise this advanced reasoning, we developed a hybrid reasoning method that combines NLP methods with logic-based subsumption [2].

References

1. Barisevičius, G., Coste, M., Geleta, D., Juric, D., Khodadadi, M., Stoilos, G., Zaihrayeu, I.: Supporting Digital Healthcare Services Using Semantic Web Technologies. In: ISWC'18. pp. 291–306 (2018)
2. Juric, D., Stoilos, G., Wartak, S., Khodadadi, M.: Reasoning with Textual Queries: A Case of Medical Text. In: ISWC'18 Poster (2018)
3. Stoilos, G., Geleta, D., Shamdasani, J., Khodadadi, M.: A Novel Approach and Practical Algorithms for Ontology Integration. In: ISWC'18. pp. 458–476 (2018)
4. Stoilos, G., Geleta, D., Wartak, S., Hall, S., Khodadadi, M., Zhao, Y., Alghamdi, G., Schmidt, R.A.: Methods and Metrics for Knowledge Base Engineering and Integration. In: WOP'18. pp. 72–86 (2018)