

[The PhD Workshop at the 41st International Conference on Very Large Data Bases](#)**VLDB 2015 PhD Workshop**

August 31, Kohala Coast, Hawaii, USA

Reviews For Paper**Paper ID** 9**Title** Intuitive and Interactive Query Formulation to Improve the Usability of Query Systems for Heterogeneous Graphs**Masked Reviewer ID:** Assigned_Reviewer_1**Review:**

Question	
Overall Rating	Strong Accept
Detailed Comments	<p>The thesis focuses on the problem of query formulation on top of heterogeneous graphs. The task of query formulation already being daunting for relational queries, becomes critical for graph databases. The thesis is structured into two parts: the first presents GQBE, a query-by-example approach that uses only entity tuples, while the second focuses on an interactive query formulation tool to specify exact query graphs. The paper is well written, and the work is corroborated by the implementation of two research prototypes.</p> <p>Just a minor comment, about related work. This paper may be an interesting reading:</p> <p>Interactive Path Query Specification on Graph Databases, EDBT 2015.</p>

Masked Reviewer ID: Assigned_Reviewer_2**Review:**

Question	
Overall Rating	Accept
Detailed Comments	<p>Problem description is clear: addressing the problem of improving the query formulation capability of query systems for large heterogeneous graphs.</p> <p>Importance is well-argued: proliferation of heterogeneous graph data and difficulty to formulate queries (hard to grasp entire schema)</p> <p>Existing solutions not sufficient: well-addressed in the related work</p> <p>Outline of new solution is sufficiently detailed in Section 3: query by example for graphs, general idea is well explained (+ semantics semantics), query processing is addressed. Ranking of results is investigated and experiments are performed on Freebase. It is not mentioned whether the complete freebase data is used (I suspect not).</p>

Masked Reviewer ID: Assigned_Reviewer_3**Review:**

Question	
Overall Rating	Accept
	<p>Querying graph databases can be challenging because of a user's lack of familiarity with the query language (often confusing semantics) and the schema semantics of the particular dataset. The proposal addresses these problems by advocating query-by-example as a means to bootstrap query construction and then an interactive visual query editor to further refine the query.</p>

Detailed Comments	<p>The problem is clearly an important one. All the buzz around Big Data has resulted in people getting interested in data analysis, but not many of the potential users have the necessary skills to master database query languages. However, such users do understand the data and are motivated by stories they can write based on the data. Further, they have the necessary motivation to be early adopters of technologies that make their data analysis efforts simpler.</p> <p>The first part of the proposal is well-motivated and well-written and clearly illustrates the problems and ideas behind possible solutions. The writing can be improved for the second part. For example, the difference between active and passive (?) modes is not really clear (or necessary).</p> <p>At a high level, the line of work has many more applications and author can probably be more ambitious. For example, think more about visualizations. Tuple-based interfaces, or even visual query editors, are not the future. The people who would most benefit from interactive query formulation would also likely be better served by doing their slicing and dicing through other interfaces that support interactive visualization, e.g., charts or maps or timelines. A visual query editor might not be that useful for them because it still would require them to understand one of the many graph-structured query languages. A good starting point to learn further about that direction of work (interactive visualization) would be to read Jeff Heer's PhD thesis.</p>
-------------------	---