We would like to thank all reviewers and the meta reviewer for their insightful comments which helped us to significantly improve the paper. We are encouraged by reviewers’ acknowledgement of this paper. In this document, we outline the revisions we made in response to each review comment. We have used the notation [Wx], [Dx], [Rx] to denote weak points, detailed comments and revision requirements, respectively.

Assigned\_Reviewer\_1

W1) While the paper is extremely well thought out, the overall motivation for this problem needs to be fleshed out more. While this problem itself may be NP-hard, the authors do not articulate how common this problem is, or what the overall impact is. As the authors mention, “Freebase offers an entrance page showing 6 major entity domains” then perhaps first demonstrate that intuitive lternatives do not suffice, such as Pagerank representation of the graph (ObjectRank, Balmin @ VLDB 2004) and schema summarization etc? Alternatively, perhaps show that this method is \*faster\*(i.e. time-to-task\* than human curation?

W2) The paper could do with more comparisons against reasonable baselines / comparable algorithms; the evaluations (Fig 6 & 7) simply compare the two proposed algorithms, but do not compare this against a reasonable existing alternative.

D1) The details of the user study need to be fleshed out. Tor example, what does it mean to say “important” in “The workers were asked which of the 2 entity types in the pair is more important.”; all responses are valid (no outliers) ?, etc. Other papers to look at, esp. in the context of search:  
\* "Qunits: queried units in database search", Nandi and Jagadish CIDR 2009  
\* "Query Biased Snippet Generation in XML Search", Chen et al SIGMOD 2009

R1) The focus of this paper could be improved significantly if the authors focused on the qualitative aspects and the human curation aspects instead of the performance experiments. Conceptually, the task is not time-intensive since once the preview is generated, it won’t change frequently even there is more new edits coming in considering the size of entity graph.

R2) The current user study is not enough to tell how this will indeed help users to make a decision to select the entity graph. It is only used to evaluate the ranking. The paper needs to show that the technique provided is a good qualitative alternative to the baseline / naive case.

Assigned\_Reviewer\_3

W1) Selecting sample tuples is very likely an important part of the puzzle and it was disappointing to see the authors use random sampling.

W2) User study and evaluation are weaker in comparison to the rest of technical presentation.

D1) It is hard to appreciate the quality comparison, when the Freebase schemas used are unknown (as the author suggests, the published pages are no longer accessible). At the very least the authors should list the entity types and the non-key attributes for a few of the tested domains.

D2) The authors assume that the six entity types and their attributes as published by Freebase can be used a golden set. Is there a justification for this? Freebase does not necessarily have the preview goals laid out in the paper. It would be more convincing if the paper included a evaluation in which the Freebase schema was also rated in the user study and then the overlap measure with the solutions generated by the different algorithms.

D3)The entity type importance correlation (Section 6.1.3) does not appear to necessarily relevant to the author's preview table goals. There has to be justification for why that comparison is a reasonable one.

D4) Why is Precision@K a good measure for preview quality? At least an informal justification is necessary. Further, the charts in Figure 5 are hard to read and understand.

D5) There is no explanation for the non-key attribute scoring in Table 4 not described. In order to measure non-key attribute quality, you could fix the key attributes to be the same as in the gold standard and then let the algorithm only pick the non-key attributes. Then, you can the same metric used to measure quality of key attributes can be used to compare the precision of the non-key attributes.

D6) An alternate evaluation strategy might be a have a set of 10 experts (e.g., grad students with some data management understanding) independently create preview tables for the said datasets under the outlined design constraints (k tables, n attributes). These can be instead compared with the results of the algorithm (and correlation/overlap measured).

D7) Though the paper emphasizes the ability to obtain multiple variations of a preview schema (concise, concise + dense, and concise + diverse), the experimental section does not address the objective / subjective quality of the previews computed over the variations (except for the example in Table 6).

D8)The sample preview schemas in Table 5 and Table 6 do appear rather interesting and convincing, which makes the weakness of the rest of the evaluation, unfortunately, more pronounced. Also, the examples in Table 5 seem much more convincing than the motivating examples in Table 2.

R1) The user evaluation has be more convincing. Specially, it would make sense to quantitatively compare the preview tables computed by the algorithms against hand-crafted schemas created by a set of experts.

R2)The differences between the variations in criterion (concise + diverse + dense) would ideally also be evaluated.

Assigned\_Reviewer\_5

W1) In the experiments, there should be a direct comparison between the schema graph (or summary of the schema graph) and the preview tables from the usefulness and understandability of the users (i.e., data workers) point of view.

W2) The schema graphs used in the experiments are not big. Considering the fact that preview tables are useful when users are dealing with big schema graphs, this is not acceptable.

W3) The related work section is too short and shallow. The conclusion should include the possible future work.

D1)The problem studied in this paper looks interesting. The concept of preview tables are modeled mathematically and the right parameters to control the size of the preview tables are introduced. However, the motivation of using preview tables is not supported in the experiment.

D2)The contribution of this paper is that it claims providing preview tables is more useful than providing schema graph (or summary of the schema graph) for the purpose of understanding the entity graph. However, the authors fail to show this in the experiments. In the experiments, the accuracy of the solution is considered only by looking at the ranking orders of candidate keys. Even in the user study, users are asked to compare which pair of the 2 entity types are more important. However, the right experiment would show how useful the schema graph (or summary of the schema graph) is versus the preview tables. Instead of asking the users of AMT, actual data workers should be asked to evaluate the usefulness of the preview tables.

R1) In the experiments, compare the usefulness of the preview tables vs schema graph (or summary of the schema graph). This can be done by a user study among data workers (and not AMT users).

R2) In the experiments, a sample of the schema graph summary should also be shown beside each preview table sample.

R3) The related work should be expanded. More discussions about similar work that has been done in the past should be added.

R4) Since the concept of the preview tables for entity graphs are introduced for the first time in this paper, the conclusion should include a thorough discussion about the possibilities of future work.