Dear reviewers and EDBT chair,

Find attached the following your requests we updated the files:

- **Golfarelli EDBT_2021_Demo.pdf**: the camera ready version of our paper
- **Golfarelli EDBT_2021_Demo_Diff.pdf**: the camera ready version of our paper where we highlighted the changes with respect to the original version.

Compatibly with the space constraints, we responded to the reviewers' comments. Find below a detailed list.

We hope that the paper will meet your approval.

The authors

# Reviewer #1

**Questions**

- **1. Recommendation**
  - Accept
- **2. List 3 Strong Points**
  - S1:                   The               paper                  is                    well-written
    S2: The demo addresses an interesting and non-trivial use case
    S3: This paper clearly indicates which research papers and approaches it builds upon.
- **3. List 3 Weak Points**
  - W1: No code, link to website, or video

    We added the link to the online demo. The video will be published on the conference date.

  - W2: The demonstration scenario focuses on usability in an abstract way rather than describing how exactly users can interact with the system and what data/scenario is used.

    In "Section 4 - Demo Proposal" we describe two scenarios: guided and unguided. For each scenario we explain the interaction protocol and the goal we want to achieve. In particular, we rephrased the guided scenario to be more specific. However, due to the sake of brevity, we do not have space to put more concrete examples as done in Section 3 to explain the visual and interaction metaphor.

- **4. Summary of recommendation**

- ○ Conversational OLAP is a very relevant, challenging, and interesting problem. This paper outlines a system that helps users formulate OLAP queries using natural language in an interactive way. Having done a user study of the system is quite nice but due to the brevity of the discussion in this paper, it is difficult to understand the details. It would have been nice if the authors had described what exactly is planned to be demonstrated in terms of data, user interaction, etc. The description could be improved but overall I think this could be a solid demo paper.

# Reviewer #2

**Questions**

- ● **1. Recommendation**
  - ○ Accept
- ● **2. List 3 Strong Points**
  - ○ S1. Interesting demo; conversational OLAP will be interesting to a large audience at the conference.
  - ○ S2. Use of cubes to generate the KB enables a large class of analytical queries
- ● **3. List 3 Weak Points**
  - ○ W1. The **OLAP operator is not clear**. How does the interpretation gets generated incrementally from a full interpretation tree? The description is insufficient.

    We have extended the discussion around the OLAP operator. The interpretation of a full query into SQL is described in "Section 2.2.3 - SQL generation".

  - ○ W2. The paper is hard to follow.

    Following the reviewer's remarks, we rephrased several parts of the paper to simplify and clarify the presentation.

  - ○ W3. There are several key concepts that are either not described at all, or not described in enough detail. Most notably, DFM and OLAP operator concepts are central yet it is not clear how these work in COOL.

    We added the description of the DFM formalism as well as a more detailed definition of OLAP operator.

- ● **4. Summary of recommendation**
  - ○ The demo proposes a conversational OLAP system. They propose to create a KB from the cube definition and use textual similarity for matching tokens to entities. The system has the ability to disambiguate, and the final results are shown as visualizations.

    This could be an interesting demo. However, the whole concept of OLAP

operators, and how they incrementally change the interpretations are not covered with enough detail. Moreover, DFM plays a key role in the interpretation but not it's not described at all in the paper. These are important aspects of the demo proposal.

Entity matching is based on textual similarity, which has limited accuracy. Especially, if the names in the database are not meaningful names, these similarity measures have limited use.

We agree with the reviewer. However, while the "minimal" lexicon is indeed imported directly from the cube, we remark that COOL supports the possibility to integrate (ad-hoc) vocabularies to answer this specific scenario.

The paper's language is somewhat hard to follow with grammatical mistakes.

We have carried out careful grammar checks.

# Reviewer #3

**Questions**

- **1. Recommendation**
  - Accept
- **2. List 3 Strong Points**
  - 1. Providing a NLP interface to perform OLAP queries is a very useful service, as a lot of data reside in DWs and a large portion of Data Scientists are not SQL literate
  2. The proposed system tries to automatically identify the key entities, while providing an helpful interactive interface for the user to help in the process
  3. Their preliminary test with 40 users has some encouraging results
- **3. List 3 Weak Points**
  - 1. The offline phase (where key entities are extracted) should be described with more details. As is now it seems that either the created KB is very restricted or the manual process is very cumbersome

    We detailed the offline phase to explain how metadata are automatically collected from a cube in a ROLAP engine.

  - 2. I think the paper is missing some quantitate metrics to assess the usability functionality that they describe as the target of the demo

    Quantitative metrics on user experience as well as efficiency are reported at the end of Section 3.

- ○ 3. As I describe above and in the summary I think some more explanation is needed in parts of the paper

  *Following the reviewer's remarks, we rephrased and enriched several parts of the paper.*

- **4. Summary of recommendation**
  - ○ I think this is in general a well written purpose providing a satisfactory solution to a very real/relevant problem

    Concerning the tokenization/mapping during the online phase, the usage of n-grams as opposed to a word-based is a little strange as one may end up combining non-sensical words based on the assigned similarity score.

    *We will investigate this issue in an extended version of the paper. N-grams can indeed produce combinations of words that do are unlikely to the entities in the knowledge base.*

    As far as the paper presentation is concerned I think it is in general well written, with the exception of week point (1) above as well as a lack of description for the need of two separate phases (engage and navigate). From my experience a navigation phase using free text can easily become confusing to the untrained user, and thus make him waste time trying to rewrite the query in an appropriate format. A GUI may be more appropriate. I think the authors should comment on that to clarify

    *The reviewer makes an interesting suggestion. However, having a single phase would have required a single and more complex grammar; by designing two distinct phases (engage and navigate) we keep the overall approach simpler, diving the problem of interpreting a single query rather than an OLAP operator. Tests with users did not presented particular issues with the navigation phase (e.g., if a user is "lost" it is possible to restart the session). As to the GUI, Figures 5 and 6 show COOL's GUI.*

    Some typos:
    ---------------
    Section 1: "The goal of the demonstration is to let the audience to evaluate" --> "The goal of the demonstration is to let the audience evaluate"

    "to assist the user in easily formulate" --> "to assist the user in easily formulating"

    *We fixed the typos.*