R1.

**W1:** It's a good idea that you mentioned semantic heterogeneity!

In AliCoCo, semantic heterogeneity is mainly handled by the class taxonomy.

In the layer of primitive concepts, we allow each surface form to belong to several classes, which means there can be several primitive concepts with the same name but different IDs (meanings). During the construction, we used a sequence labeling model to disambiguate between different classes (section 4.1).

In the layer of e-commerce concepts, we express each e-commerce concept by several primitive concepts with classes to identify its accurate meaning (section 5.3). This process is also an example of using AliCoCo to disambiguate raw texts.

In our revised version, we added a short explanation in section 4.1, and a short comparison of several KGs towards handling semantic heterogeneity in related work.

**W2:** We mentioned "lack of context" in section 5.2.2 and section 5.3, where we try to pick out good e-commerce concepts from candidates and further understand their semantic meanings by linking to primitive concepts. The main challenge is that surface forms are too short to provide enough semantic information for understanding, comparing to normal text classification and NER tasks.

We agree with you that user data may contain helpful information, however, the definitions of these two tasks are actually user-independent. Besides, user data is actually too noisy to infer much useful information for these two tasks.

**W3:** In section 8, we explained the main difference between AliCoCo and Amazon's Product Graph (rephrased in revised version). Although we want to make more empirical comparisons, unfortunately, the information about Amazon's Product Graph we can gather is limited. For possible future comparison, we listed detailed statistics of AliCoCo in table 2.

**D:** Refer to W2.

R2.

**W1:** We have improved English-written in the revised version.

**W2:** Thanks for your suggestion, but we think "functional concepts" can be ambiguous since we already defined a class called "Function" in the taxonomy and here we want to emphasize that those concepts are used for e-commerce purposes.

**D1:** Yes, good question! We are aware of that. In the first version of AliCoCo, we planned to keep things simple, since the current Boolean true/false relationships support our online business well. In the subsequent version, we will bring the score into the system.

**D2:** We have corrected the errors in our revised version. Thank you!

R3.

**W1:** We agree that each component although sophisticated is not incredibly novel. The main contribution of this paper is to put together an end-to-end large comprehensive knowledge graph that is actually deployed for real-world e-commerce system in Alibaba. Each component is non-trivial and the proposed methods are effective and actually novel at times, e.g. using external knowledge to overcome the difficulties of limited contexts.

To summarize our main contribution more clearly, we added a paragraph at the end of section 1 in our revised version.

**W2:** We have improved the English in our revised version and will continuously work on that if the paper is accepted. Thanks for your advice!

**W3:** Thanks for your understanding! In the future, we will consider releasing samples of the evaluation data.

**D1:** We have updated figure 1 and revised section 2 to make ideas clearer. Thanks for your suggestions!

**D2:** We have corrected the errors in our revised version. Thank you!

**D3:** We have omitted the detailed description of BiLSTM-CRF in our revised version. Thanks for your advice!

**D4:** The sampling strategy is actually different from classical uncertainty sampling, where we not only consider the uncertain samples but also samples with high confidence. We have added a reference and rephrased the description in our revised version to make it clearer.