# 论文修改意见汇总

## 题目

* ☐ 题目应该体现出文章的特色。从数据或者技术特点（基于NLP）入手，最好的方式是通过数据特点入手体现特色。

## 引论

主要问题：引论部分过于平淡，体现不出工作价值。逻辑缺陷大

### Major

* ☐ 限定用户的身份，程序开发人员和普通人员。
* ☐ 定量描述DockerHub的容器镜像数目和拉去次数，来表明搜索的事情是常有发生的，因而需提高搜索效率及其必要性。
* ☐ 需要说明需要得到精准标签的理由。
* ☐ 描述其他工作的目的在于体现现阶段研究刚起步，突出我们的工作价值。
* ☐ 解释数据是什么，为什么要考虑这两类数据。
* ☐ 方法的创新性是如何体现出来的。为什么要进行扩展。
* ☐ 贡献描述过于平淡，仅体现了本文的工作。

### Minor

* ☐ 语病处理。

## Motivation

### Major

Motivation部分的问题在于描述还是过于虚，没有完整的举一个示例，从搜索到结果的展示。目前考虑两个解决方案：

* ☐ 完善示例。
* ☐ 合并到引论部分。

目前是在Motiviation中把长文本和短文本描述的文本特征描述出来，这一点在引论部分就没有过多地讲述，导致了逻辑的缺陷。

### Minor

* ☐ 语病处理。
* ☐ 补充为什么要考虑其他社区的数据。

## 相关工作

### Major

* ☐ 目前的思路与现有的方式差距较大，最后一段修改逻辑。

### Minor

* ☐ 语病处理。

## 主体方法

### Major

* ☐ LDA方法的工作过程，选择该方法的原因，按照短文本的思路进行并添加伪代码。

### Minor

* ☐ 语病处理。
* ☐ 整体方法框架图加上Data process层。
* ☐ 数据处理过程图细化。

## 实验

### Major

* ☐ 加大实验规模。
* ☐ 阐述为什么使用EnTagRec++。

### Minor

* ☐ 语病处理。

## 整体

* ☐ 完善表格