**Meeting Agenda and Minutes** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Meeting Title**: Project Kick-Off Meeting  
 **Date:04/08/2022**  
 **Time: 5:15 pm – 6:15pm**  
 **Attendees: Ahmad Asgharian Rezaei, Chengyang He, Huahu Wen, Yipei Liu, Yiyang Xie** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Item: result overview Allocated time: 30** minutes **Notes: Clear what we have**

**Item:** Communications **Allocated time: 10** minutes **Notes: Identify the jobs of each team member**

**Item:** Next Steps  
 **Allocated time: 50** minutes  
 **Notes:**

1. **The** **download code snippet for downloading the pdf files from different sources**
2. **Extracting the uncategorized tables from the PDF Files**
   1. **Right now we have a primary algorithm for reading tables from images.**
3. **Listing the all the families of tables in target DOIs.**
4. **Building a machine learning model for predicting the category of a table.**
5. **Predicting each table and putting it in one of the categories.**
6. **Having a function to build each family of tables.**
7. **Having a function to read each family of tables and create the unified table.**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Other items:** **Leader: Chengyang**  
**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**