

Terms and Application Form of the JDDC 2.1 Corpus

INTRODUCTION

JDDC 2.1 corpus is a large-scale multimodal multi-turn dialogue dataset collected from JD.com, a mainstream Chinese E-commerce platform, containing about 246 thousand dialogue sessions, 3 million utterances, and 507 thousand images, along with product knowledge bases and image category annotations. The dataset is collected from the online real data, and thus, we must ensure that the dataset is used only for research purposes. We appreciate that you could strictly observe all the terms hereof.

Terms

The copyright of this dataset belongs to JD.com. The dataset is ONLY used for research purposes. Commercial purposes are strictly prohibited. Reproduction, transmission, or distribution of this dataset are prohibited. **An individual with access is not allowed to share the dataset with other persons or institutions. The access is to be terminated when the applicant is no longer at the current institution. (Chinese translation of the key point: 有权使用该数据集的个人不允许与其他人或机构共享数据集。当申请人不在当前机构时,该数据集的使用权将被终止。)** Violations of the above terms will result in legal liability. If you are interested in our dataset, please fill out the application form below and send it to lihaoran24@jd.com. We will send the download link of the dataset to the applicant.

Application Form

I, _____ (name), working currently
at _____ (affiliation)

acknowledge my understanding and accepting of the above terms.

Signature _____

Date _____

Accepted by the organization (To be completed by your supervisor):

Name _____

Institution _____

Title _____

Signature _____

Date _____

Please print, complete, and c.c. this form to your supervisor.

To maintain compliance with the terms and simplify the approval procedure, please send the application forms with your organization email. We would appreciate that you can provide proofs of your identity, such as your homepage, semantic scholar ID, lab website, etc.