**儿童故事常识推理与寓意理解**

**（The Dataset of Commonsense Reasoning and Moral Understanding Evaluation in Children's Stories，CRMUS）**

**共享协议 （2024 年 3月版本）**

甲方：山西大学计算机与信息技术学院

乙方：

CRMUS是基于四百多个经典伊索寓言故事构建的常识推理与寓意理解数据集。对与故事密切相关的常识知识进行推理问答，以及挖掘故事内在的深层寓意，能让语言模型更好地模拟人类对文本的理解，提高其对故事的理解能力和推理能力，这对于推进故事阅读理解任务有着重要意义。为了促进机器阅读理解技术的研究与发展，甲方将CRMUS免费共享给乙方，用于科学研究。

**特别说明**

1.如果乙方希望将共享数据用以商业产品，请先发送邮件联系我们说明情况。

2.未来我们还会继续共享更多数据，旧的数据也可能发生更新（修改错误等）。如果需要新的共享数据，通常不需要重新签署协议（除非协议内容有所更新）。对于新的共享数据，即使无需重新签署协议，乙方仍需遵守本协议条款。

CRMUS数据集共享资源详情信息如下表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集划分 | Train | Dev | Test | Total |
| 总的问题数/选项数 | - | 652/2608 | 2748/10992 | 3400/13600 |
| 常识推理 | - | 400 | 1692 | 2092 |
| 寓意理解 | - | 252 | 1056 | 1308 |

本着互相促进及尊重知识产权的原则，乙方必须遵守下述条款：

1. [CRMUS只免费提供给用户用于](mailto:GCRC只免费提供给用户用于非盈利性科学研究使用。如果用于商业产品，请联系谭红叶老师，联系邮箱hytan_2006@126.com。)**[非盈利性](mailto:GCRC只免费提供给用户用于非盈利性科学研究使用。如果用于商业产品，请联系谭红叶老师，联系邮箱hytan_2006@126.com。)**[科学研究使用。如果用于商业产品，请联系谭红叶老师，联系邮箱tanhongye@sxu.edu.cn。](mailto:GCRC只免费提供给用户用于非盈利性科学研究使用。如果用于商业产品，请联系谭红叶老师，联系邮箱hytan_2006@126.com。)
2. 该资源仅在签署协议的组织内部使用，不以任何方式将数据集扩散给第三方。
3. 高校或研究所获得CRMUS后，在和第三方公司进行非商业性质的、以发表文章为目的的科研合作中，允许使用该数据集。如果科研合作终止，公司需要取得山西大学的许可，并签署协议，方可单独使用。
4. 数据或论文引用条款：如果文章或其他科研成果中使用到CRMUS，对外发布时，应公开声明使用了山西大学提供的数据，同时发信给tanhongye@sxu.edu.cn，说明相关情况。
5. 代码公开建议条款：我们希望通过共享数据，推动机器阅读理解领域的研究。代码公开非常重要，可以保证研究的良性发展。如果公开发表的论文（包括Arxiv文章），且实验中使用了CRMUS数据集，建议发布完整的实验代码，包括实验配置等信息。
6. 乙方了解并确认，山西大学提供的数据都将是以“现状”、“原样”提供的，在适用法律所允许的最大范围内明确否认和排除任何其它担保，无论是采用书面还是口头形式或者以明示或暗示方式，包括任何适用于某一特定用途或目的的质量性、适销性或适合性保证、涉及任何专利有效性的担保或任何第三方知识产权的非侵权承诺。在法律允许的最大范围内，山西大学不因为向乙方公司/单位授权使用该数据而承担任何赔偿责任或其他法律责任等不利后果。

乙方负责人需完整、清晰地了解以上协议内容，并严格遵守。

🞎我已完整阅读并同意该协议。

🞎拒绝并退出。