

ChiMed 和 ChiMST 数据集使用协议

此协议是您（希望使用、并下载数据的用户）与我们（华盛顿大学语言学系 BioNLP 研究小组）之间的协议。

您同意以下条款，来下载数据：

1. 本数据的原始数据版权为“39 健康网”所有；标注的分词和医疗术语标签为我们所有。我们依据与“39 健康网”之间签署的协议来发布此原始数据。
2. 任何个人和团体未经“39 健康网”授权，不得将此数据共享给第三方。
3. 本数据仅限非商业性质的研究使用。
4. 如果使用本数据发表成果，应在成果显著位置注明“原始数据来源 39 健康网”，并引用以下文章
 - （1）Yuanhe Tian, Weicheng Ma, Fei Xia, and Yan Song. 2019. ChiMed: A Chinese medical corpus for question answering. In *Proceedings of the 18th BioNLP Workshop and Shared Task in conjunction with ACL-2019*, pages 250–260. <https://www.aclweb.org/anthology/W19-5027/>
 - （2）Yuanhe Tian, Han Qin, Fei Xia and Yan Song. 2022. ChiMST: A Chinese Medical Corpus for Word Segmentation and Medical Term Recognition. In *LREC 2022*, pages 5654-5664. <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.607.pdf>
5. 如果在本数据上做额外的标注（比如在问答记录上添加额外的标签），标注的版权为研究者所有。如需发布标注结果，应只发布标注标签，不应发布数据的文本。如果需要发布带有标签的数据，可以联系我们，并由我们协助完成。届时，我们将在我们的网站上发布您标注的数据，并在显眼位置写明这项工作由您的团队完成。

数据使用申请表

请选择申请使用的数据

- ☐ ChiMed
☐ ChiMST

姓名: _____

单位: _____

邮箱地址，我们只接受以 edu 为结尾的邮箱地址: _____

在下面签名，辨明您同意上述“ChiMed 和 ChiMST 数据集使用协议”，并确认“数据使用申请表”中的信息准确无误。

签字: _____

日期: _____