

# 木兰-白玉兰开放数据许可协议 1.0

## 结构内容分离版 MBODL 1.0(Draft)

### 对协议采纳者的声明

白玉兰开源开放研究院不是一家律师事务所，也不对外提供法律服务或建议。提供「木兰-白玉兰开放数据许可协议 V1.0 (Mulan-Baiyulan Open Data License V1.0，以下缩写为 MBODL V1.0) 并不导致「律师—当事人」关系或其他法律关系的建立。

该协议文本及相关信息皆依其「现状」(as is)提供。白玉兰开源开放研究院("Baiyulan Open AI")对于其协议文本，或通过该协议文本提供的数据或相关信息不提供任何担保，在最大可能程度内，也不对因使用其协议或通过该协议提供的数据或信息而造成的损失承担损害赔偿责任。

「木兰-白玉兰开放数据许可协议 V1.0」条款：<http://baiyulan.org.cn/mbodl/1.0/>

### 使用方式

在需要发布并采用本协议的数据集的相关文件、文档和网页等资料中，附上下述声明并附上协议链接

「本数据集授权在木兰-白玉兰开放数据许可协议 V1.0 下，数据集中的数据内容相应授权在XXXX」

注：

1. 木兰-白玉兰开放数据许可协议仅对数据集结构做授权
2. 数据集中的内容如本身有其授权许可请引用其原有的授权许可
3. 数据集中的内容如原本未有授权许可，建议将其授权在知识共享 CC-4.0 国际版下

协议的使用前提，请发布者自检：

- √ 数据发布者对数据拥有所有权或用益权
- √ 数据发布者应依照相关数据安全管理制度做好数据审查，所发布数据应不包含有涉及国家秘密、国家安全、社会公共利益、商业秘密和个人信息的内容
- √ 数据发布者希望将数据公开、免费发布
- √ 数据发布者希望仅对数据使用者加以「注明来源」的使用、分享限制
- √ 数据发布者希望协议不具备「传染性」，以宽松方式传播

### 协议正文

木兰-白玉兰开放数据许可协议 1.0 (草案) MBODL1.0(Draft)

#### 前言

目的：木兰-白玉兰开放数据协议旨在提供一组标准的许可协议，从而加速增长可用于人工智能训练的数据资源。

价值：木兰-白玉兰开放数据协议根据「开放协作、以人为本、合法利用」的原则，为创建良好的数据流通和利用环境，并鼓励数据流通趋向自由开放，产出的代码、模型和成果开源共享，探索人工智能对数据开源开放的新模式。

**适用性：** 本协议适用于数据发布者具备对数据的所有权或者用益权的情况。本协议不适用于包含个人信息的数据。

## 定义

1. 授权人：指依本许可协议提供数据的自然人或法人
2. 您：指依据本协议行使其所获得授予之权利的个人或机构。"您的" 有相应的含义
3. 数据集：指经由某种系统性或方法汇编而成的数据集合
4. 数据：指任何以电子或者非电子形式对信息的记录
5. 数据内容：即数据所记录的信息内容
6. 数据结构：即某种系统性或方法性的选取、编排、组织数据的方法
7. 获取：指访问、浏览或/和下载数据的行为
8. 使用：是「数据处理」行为的一种，包括但不限于复制、审阅、公开展示等行为以及计算使用
9. 计算使用：指对数据的
  - 分析：指通过对数据予以解读并利用工具等发现模式、趋势、关联性等的操作
  - 训练：指向未训练模型提供数据，从而通过权重、参数、结构等调节而输出已训练模型的操作。
  - 抽象表示：指将数据转化为另一种表现形式的操作，新表现形式下每一个数据元素能够模拟或映射原有数据。
10. 模型：指基于数据预测未来趋势、总结过往模式或得到其他抽象结论的机器学习算法、人工智能算法、或其他学习型技术的组合：
  - 未训练模型：指模型尚未输入数据，其权重、参数和架构均未蕴含来自数据的分析
  - 已训练模型：指未训练模型通过数据的输入，其权重、参数和架构蕴含来自数据的分析
11. 成果：指任何通过数据的使用所获得的输出或产出，其包括但不限于基于本许可协议下数据所训练而得到的已训练模型，基于本许可协议下数据分析而得到的结论等。成果自身不应当附带超过计算使用所需的最小量的数据。
12. 新增：指针对本许可协议下数据，作出新增数据的行为。新增并不算作一种「成果」。
13. 修改：指针对本许可协议下数据，作出删减、修正或重新编排的行为。修改并不算做一种「成果」。
14. 改良数据：指对本许可协议下数据做「新增」或「修改」后所产生的新数据
15. 分享：指针对本许可协议下所授权数据集的全部或部分，或其改良数据，向本许可协议约定范围外的第三方提供物理拷贝或虚拟访问的行为。

## 许可范围

1. 您被授予全球范围内，免费的、非排他的、不可撤销的、永久的（除非符合「期限与终止」所规定的情况）许可。该许可准许您的行为如下：
  - 获取
  - 使用
  - 分享
2. 本许可协议不适用于任何专利权或商标权许可。
3. 为避免疑义，若著作权的「例外和限制」（如合理使用 Fair Dealing and Fair Use ）适用于您对本许可协议下的数据的使用，本许可协议将不适用，您也无须遵守本许可协议之条款。

## 许可限制

您被许可的行为明确受以下条件限制：

### 1. 注明来源

a. 若使用、分享在本许可协议下的数据或改良数据或成果，您必须清楚标识「数据来源」信息，建议的形式如下：

[数据提供机构] [年份][数据、数据及其素材的发布名称与版本号][数据URI或链接]

此数据依「木兰-白玉兰开放数据许可协议 1.0版」向公众发布，您在遵守本协议条款各规定前提下利用数据。

「木兰-白玉兰开放数据许可协议 1.0版」条款：<http://baiyulan.org.cn/mbodl/1.0/>

b. 若您的成果涉及多个来源的数据，且列举所有数据来源不具备可操作性，则您可以使用一个单独的文档列举并标识所有数据来源，并在您的产出作品或衍生数据中提供指向该文档的链接。

## 期限与终止

您在利用本协议下的数据的过程中，如有下列行为之一，则您依据此许可协议享有的权利自动终止：

1. 违反本许可协议条款的行为
2. 侵犯商业秘密、个人隐私等他人合法权益的行为
3. 违反中华人民共和国法律法规等规范性文件的其他行为

## 免责声明

1. 本许可协议不构成、或不得被解释为允许您声明或主张：您或您对授权数据的使用与数据的权利人及相关的数  
据管护、数据发布等机构相关联，或得到其认可或被授予正式地位。
2. 您利用本许可协议下的数据，而受到任何损害或损失，或因此导致第三人遭受损害或损失而因此遭到赔偿请  
求，除法律、法规另有规定外，数据的权利人及相关的数  
据管护、数据发布等机构不承担任何赔偿或补偿的  
责任。

## 法律依据

本条款的解释、效力、履行及其他未尽事宜，以中华人民共和国法律为依据。