
Plan Overview

A Data Management Plan created using DMPonline

Title: 2024□□□□□□□□□□

Creator:da dajun

Affiliation: The University of Hong Kong

Funder: Arts and Humanities Research Council (AHRC)

Template: AHRC Template

ID: 167324

Last modified: 05-01-2025

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

2024□□□□□□□□ - Initial DMP

Ethics and responsible research and innovation (RRI)

What are the ethical or RRI implications and issues relating to the proposed work?

Question not answered.

2024年度项目数据管理计划 - Detailed DMP

Data Summary

1. Briefly introduce the types of data the research will create. Why did you decide to use these data types?

本项目计划于2024年启动，旨在研究人工智能在医疗诊断中的应用。项目团队决定使用以下数据类型：1. 临床数据：包括患者病史、诊断记录、治疗方案等。2. 影像数据：包括X光片、CT扫描、MRI等。3. 文本数据：包括医生问诊记录、患者咨询记录等。4. 结构化数据：包括实验室检测结果、患者基本信息等。5. 非结构化数据：包括医生手写病历、患者咨询记录等。6. 元数据：包括数据源、数据格式、数据更新时间等。7. 模型数据：包括训练好的模型、模型评估结果等。8. 实验数据：包括模型训练过程中的数据、模型测试结果等。9. 用户数据：包括用户使用系统的记录、用户反馈等。10. 系统数据：包括系统运行日志、系统配置等。以上数据类型将用于构建人工智能模型，以提高医疗诊断的准确性和效率。

Data Collection

2. Give details on the proposed methodologies that will be used to create the data. Advise how the project team selected will be suitable for the data/digital aspects of the work, including details of how the institution's data support teams may need to support the project

数据收集将主要通过以下方法进行：1. 临床数据：通过医院信息系统（HIS）和电子病历系统（EMR）收集。2. 影像数据：通过医院影像系统（PACS）收集。3. 文本数据：通过医生问诊记录系统收集。4. 结构化数据：通过实验室信息系统（LIS）收集。5. 非结构化数据：通过医生手写病历系统收集。6. 元数据：通过数据源系统收集。7. 模型数据：通过模型训练系统收集。8. 实验数据：通过模型测试系统收集。9. 用户数据：通过用户使用系统收集。10. 系统数据：通过系统运行日志系统收集。项目团队将与合作机构的数据支持团队紧密合作，确保数据的完整性和安全性。

Short-term Data Storage

3. How will the data be stored in the short term?

短期数据存储将采用以下方法：1. 临床数据：存储在医院信息系统的数据库（如Oracle、MySQL）中。2. 影像数据：存储在医院影像系统的存储设备（如磁带库、光盘库）中。3. 文本数据：存储在医生问诊记录系统的数据库中。4. 结构化数据：存储在实验室信息系统的数据库中。5. 非结构化数据：存储在医生手写病历系统的数据库中。6. 元数据：存储在数据源系统的数据库中。7. 模型数据：存储在模型训练系统的数据库中。8. 实验数据：存储在模型测试系统的数据库中。9. 用户数据：存储在用户使用系统的数据库中。10. 系统数据：存储在系统运行日志系统的数据库中。

3a What backup will you have in the in-project period to ensure no data is lost?

项目期间将采用以下备份策略：1. 临床数据：每天备份一次，备份到磁带库。2. 影像数据：每周备份一次，备份到光盘库。3. 文本数据：每天备份一次，备份到医生问诊记录系统的数据库。4. 结构化数据：每天备份一次，备份到实验室信息系统的数据库。5. 非结构化数据：每天备份一次，备份到医生手写病历系统的数据库。6. 元数据：每天备份一次，备份到数据源系统的数据库。7. 模型数据：每天备份一次，备份到模型训练系统的数据库。8. 实验数据：每天备份一次，备份到模型测试系统的数据库。9. 用户数据：每天备份一次，备份到用户使用系统的数据库。10. 系统数据：每天备份一次，备份到系统运行日志系统的数据库。

Long-term Data Storage

4. How the data will be stored in the long term?

长期数据存储将采用以下方法：1. 临床数据：存储在医院信息系统的数据库（如Oracle、MySQL）中。2. 影像数据：存储在医院影像系统的存储设备（如磁带库、光盘库）中。3. 文本数据：存储在医生问诊记录系统的数据库中。4. 结构化数据：存储在实验室信息系统的数据库中。5. 非结构化数据：存储在医生手写病历系统的数据库中。6. 元数据：存储在数据源系统的数据库中。7. 模型数据：存储在模型训练系统的数据库中。8. 实验数据：存储在模型测试系统的数据库中。9. 用户数据：存储在用户使用系统的数据库中。10. 系统数据：存储在系统运行日志系统的数据库中。

4a. Where have you decided to store it, why is this appropriate?

U

4b. How long will it be stored for and why?

Question not answered.

4c. Costs of storage - why are these appropriate? Costs related to long term storage will be permitted providing these are fully justified and relate to the project Full justification must be provided in Justification of Resources (JoR)

Question not answered.

Data Sharing

5. How the data will be shared and the value it will have to others

0000000000PPT000000000000000000000000PPT000000000000;00000000

5a. How the data will enhance the area and how it could be used in the future?

PPT ↔ PPT

5b. Releasing the data - advise when you will be releasing and justify if not releasing in line with AHRC guidelines of a minimum of three years. If the data will have value to different audiences, how these groups will be informed?

Question not answered.

5c. Will the data need to be updated? Include future plans for updating if this is the case.

□□□□□□□□□□2024□□□□□□□□□□□□□□□□

