

数据原则

- 确保原始数据是正确的。
- 修复根数据而不是生成数据。
- 使用大量检查脚本来确保数据正确性。
- 检查数据的正确性和一致性。
- 某些数据需要完全正确，而其他数据可以大致正确。例如，翻译即使包含偶尔的拼写错误也是可以接受的。
- 使用定量指标来衡量数据质量。
- 错误的数据会导致更多错误的的数据。
- 数据包括数字、字符串、文本、数据结构、图片、视频、代码、文件以及其他信息类型。
- 在物理、科学和计算机科学的实验中保留原始数据。
- 数据应该是可重复的和可再现的。