

打通校内学习平台与数据中台：突破日志保留限制，实现跨学年学情数据长期沉淀与应用

案例提供部门：西浦学习超市

1. 案例背景

学校的校内在线学习平台 Learning Mall Core 基于 Moodle 搭建，面向全校师生提供课程学习、作业测验、教学互动等服务。由于平台采用“按学年运行”的模式，每个学年结束后都会对当年的系统进行归档（archive），为新学年提供一个相对“干净”的系统环境，用于承载新的课程与教学活动。

随着平台使用规模不断扩大，课程数量、访问量和日志数据急剧增长。为保障生产系统性能与存储空间，其中 Learning Mall Core 的日志表（如学习行为 log 表）在原有架构下仅保留最近 180 天的数据。这样一来，一旦跨越 180 天窗口，教师、学院以及职能部门就难以在 LMS 本身中获取完整的历史行为数据，更无法在同一数据视图中开展跨学年、跨届的学情分析和教学质量评估。

与此同时，学校信息技术部门建设了统一的数据中台，用于汇聚教务、教学、科研等多个业务系统的数据资产。如何在不打乱现有学年归档机制和日志保留策略的前提下，将 Learning Mall 的关键学情数据纳入数据中台，解决“日志只保留 180 天、系统每年归档导致数据割裂”的痛点，成为 Learning Mall 部门与 IT 数据团队需要共同攻克的问题。

2. 解决方案

本案例以“**业务节奏不变、数据能力升级**”为原则，通过以下几方面完成方案设计与实施：

1. 梳理学情数据模型，明确对接范围

○ 由 Learning Mall 部门与 IT 数据团队联合梳理 LMS 中与学情分析强相关的数据表和字段，重点包括：

- 用户与组织信息（用户表、学院/专业/班级关联等）；
- 课程与课程分类（课程表、学期信息等）；
- 选课与角色数据（师生角色、选课关系）；
- 成绩与测评数据（测验成绩、作业成绩等）；
- 行为日志表（访问、点击、活动参与记录等）。

○ 在此基础上，形成 **Learning Mall 学习数据域模型**，为后续数据中台建模与同步提供一致的数据基础。

2. 在 2024 年春季完成与数据中台的批量对接

○ 确定首批对接的数据范围，以**用户、课程、选课关系和日志数据**为核心。

○ 对历史数据执行一次性回溯导入，将既有学年的关键学情数据迁入数据中台。

○ 之后通过定时任务流程，按日或按批次进行增量同步，保证数据中台中的 Learning Mall 数据持续更新。

3. 设计“日志 180 天 + 中台长期沉淀”的分层存储机制

- 生产环境继续保留“日志 180 天”的策略,以保障 Moodle 系统运行稳定和数据库性能。
- 数据中台中建立 Learning Mall 专属的数据主题,通过大容量存储和冷热分层策略,实现学习行为日志的**长期保存与跨年查询**。
- 从架构上形成“**LMS 负责教学业务运行,数据中台负责长期数据沉淀与分析计算**”的职责分工。

4. 建立标准化的数据服务与访问机制

- 在数据中台中为 Learning Mall 数据建立统一的数据服务接口(如视图、接口或数据集),供后续学情看板、统计报表和专项分析使用。
- 配合数据治理要求,按角色配置不同的数据访问权限,避免直接访问 LMS 生产库。

5. 面向教学一线的实际应用——MTH007 课程案例

- 2024 年 10 月,MTH007 课程教师希望查询 **2023/24 学年 S1** 中学生的在线学习行为数据,用于课程总结和教学改进。
- 按照原有“日志仅保留 180 天”的策略,Moodle 生产系统中已无法直接获得该学期完整行为日志。
- 借助 2024 年春季完成的 LM 与数据中台对接,LM 团队与 IT 数据团队通过数据中台,成功拉取并整理出 MTH007 在 23/24 S1 学期的学习行为数据(包括登录次数、资源访问、测验参与等),为教师提供了跨学年的学情分析支持,成为本方案在一线教学场景中的典型落地案例。

3. 成果与效益

- **突破日志保留与系统归档限制，实现学情数据长期连续沉淀**
 - 即便 LMS 生产环境仍采用“日志保留 180 天、每学年归档”的运维策略，中台中的 Learning Mall 学习数据依然可以跨年、跨届连续沉淀。
 - 教学管理部门和学院能够在统一的数据视图中，开展多学年的纵向对比分析，支撑课程建设评估、教学改革效果评估等工作。
- **真实教学需求得到验证：支撑 MTH007 跨学年学情分析**
 - 在 MTH007 课程的实际需求中，如果仅依赖原有 LMS 日志，已无法满足教师对上一学年 S1 的行为数据查询诉求。
 - 借助数据中台，成功拉取并分析 23/24 学年 S1 的学习行为日志，实现对学生学习活跃度、资源使用、测验参与等的量化呈现，为教师提供了有依据的教学反思素材。
 - 该案例不仅解决了具体课程的实际问题，也让教师切实感受到数据中台建设对教学工作的**直接支持与价值**。
- **显著提升决策支持与数据应用能力**
 - 学校管理层、学院和职能部门可基于中台的数据主题，构建多维度学情看板和监测指标，及时了解各学院、各课程的在线学习情况和趋势。
 - 为未来构建学生学习预警模型、教学质量诊断模型等提供了必要的历史数据沉淀。

- **降低 LMS 运维压力，优化整体系统架构**
 - Moodle 生产库不再承担长期日志存储和复杂历史查询压力，可以更聚焦于稳定承载教学业务。
 - 数据中台统一承接 Learning Mall Core 的历史数据存储与分析计算任务，架构分工更清晰，有利于后续扩展更多教学系统的数据接入。
- **提升数据治理和安全合规水平**
 - 通过数据中台集中管理访问权限和数据脱敏策略，避免“各系统各自拉库”的零散做法。
 - 为学校整体数据治理体系的完善提供了实践样本。

4. 可复制性与推广价值

1. 适用于同类 LMS + 学年归档场景

- 对于同样采用 Moodle 或其他 LMS，且需要按学年归档、限制日志保留时间的高校，本案例提供了一条“**不改教学业务节奏、通过数据中台实现数据能力升级**”的通用路径。

2. 适用于其他教学业务系统的数据中台接入

- 本案例已经验证了从业务系统梳理数据模型 → 明确对接范围 → 建立中台数据域 → 设计分层存储与数据服务的完整流程。
- 该模式可以推广至教务系统、继续教育平台等其他教学业务系统，逐步形成以数据中台为核心的“**教学大数据底座**”。

3. 为后续 AI 学习分析与智能应用奠定基础

- 长期、连续的学习行为数据是建设 AI 学习分析、个性化推荐、风险预警等应用的前提。
- 本案例的实施为今后在学情 Agent、智能看板等方向的探索提供了坚实的数据基础，具有良好的扩展潜力和推广价值。

5. 下一步计划

1. 建设面向多角色的学情可视化看板

- 基于数据中台中已沉淀的 Learning Mall Core 数据，进一步设计和开发面向校领导、学院负责人、教师等不同角色的学情可视化看板，提供分层次、可钻取的在线学习分析视图。

2. 探索学生学习预警与教学质量评估模型

- 在保证数据安全与隐私前提下，尝试基于历史学习行为数据构建学生学习风险预警模型和课程质量评估模型，为教务和学院提供更前瞻的决策支持。

3. 与智能学情分析 Agent 深度结合

结合学校正在规划的学情分析智能 Agent，将数据中台中的 Learning Mall Core 学习数据作为底层数据源，为教师和管理者提供“问答式”的即时学情洞察服务，进一步释放数据价值。