

“Build with Us | Deep Dive: Speedrun Your First AIP Workflow”是 learn.palantir.com 上的一门实战导向课程，由 Ontologize 团队(前 Palantir 工程师组成)提供 1。该课程旨在指导用户在 Palantir Foundry 中，利用 AIP(人工智能平台) 将大量的非结构化数据转化为可交互的 AI 驱动应用程序 1, 2。

以下是根据来源对该标题及课程内容的详细解释：

1. 核心目标与定义

该课程的重点是展示如何构建一个端到端的 AIP 工作流。其核心任务是将 PDF 文档(非结构化数据)转化为知识图谱(Knowledge Graph)，并利用 AI 智能体(Agent)为用户提供基于事实的准确答案 1。

- “Speedrun”(快速起步)：意味着课程侧重于快速实践，带领用户迅速走完从原始数据到产出结果的全过程 1, 3。
- “AIP Workflow”(AIP 工作流)：特指利用 Palantir 的 AI 能力(如大语言模型 LLM)与底层数据架构(本体 Ontology)相结合的开发模式 2, 4。

2. 业务场景：结核病研究库

课程使用了一个具体的公益组织案例：

- 背景：一支专注于结核病(Tuberculosis)研究的团队需要高效导航大型研究文献库，以扩大药物获取途径 2。
- 目标应用：构建一个“文献综述应用(Literature Review App)”，用户可以用自然语言提问(例如：“为什么多重耐药结核病令人担忧？”)，系统会返回精准的证据和可视化关系图 2, 4。

3. 技术架构与操作步骤

该工作流展示了 AIP 如何与 Foundry 的其他工具链深度集成：

- 数据摄取(Data Ingestion)：将 PDF 文章上传为 Media Set(媒体集) 2, 5。
- 文档处理(Pipeline Builder)：
- PDF 文本提取：使用 OCR 或原始文本提取技术 6。
- 分块(Chunking)：将长文本拆分为更易于 LLM 处理的短块(Chunks) 7。
- LLM 抽取：利用大模型提取每块内容的摘要和实体(如疾病、治疗方法) 8, 9。
- 嵌入(Embeddings)：为文本块生成向量，以便进行语义搜索(Semantic Search) 10。
- 本体配置(Ontology)：创建“块(Chunk)”和“实体(Entity)”对象，并建立它们之间的多对多链接 11-13。
- 逻辑构建(AIP Logic)：使用 OAG(Ontology Augmented Generation, 本体增强生成) 模式。这种方法通过将 LLM 与组织自身的本体数据相结合，能够有效避免 AI “幻觉”，并提供可追溯的证据 14-16。
- 可视化与应用(Vertex & Workshop)：
- 在 Vertex 中创建知识图谱模板，展示实体间的关联 17, 18。
- 在 Workshop 中搭建最终应用，集成搜索框、关联图谱和 AI 生成的回答 19-21。

4. 课程的核心价值

通过这门课程，用户可以理解 AIP 的独特优势：它不只是一个聊天机器人。它通过将 LLM 的推理能力与受治理的本体数据(Ontology)结合，实现了一种能够处理复杂、专业数据且高度可靠的 AI 开发范式 4, 22。

总结：该标题代表了一次深度实践，教你如何将散乱的 PDF 文档 转化为一个具备语义搜索、知识图谱可视化和事实问答能力的智能化专业应用 21, 23。