

# Phase2: 复杂场景深入挖掘阶段

## Phase 2: AI 异步长耗时任务 (场景扩容)

**阶段目标：** 引入事件驱动架构（EDA），使大模型脱离单一的“秒回”聊天框，具备处理诸如“幻灯片智能解说配音”、“长视频生成”、“大体量文档总结”等耗时数十分钟任务的能力。

**各个语言/组件开发清单：**

- **前端层 (React 19 / Next.js)：**
  - 开发长任务大盘（Task Center）列表页。
  - 实现基于定时轮询或全局 WebSocket 的状态栏及进度条 UI 组件。
- **业务服务层 (Java / Spring Boot)：**
  - **异步任务引擎：** 提供创建任务的接口，按任务耗时与优先级投递至底层 MQ 对应 Topic。
  - **进度调度中心：** 编写专职消费者监听进度回传，执行“高频写 Redis 缓存 + 降频写 MySQL 落盘”策略。
  - **看门狗机制 (Watchdog)：** 编写 `@Scheduled` 定时任务，清理 OSS 过期文件并纠正超时僵尸任务。
- **智能与基础服务层 (Python / FastAPI)：**
  - **异步 Worker 节点集群：** 接入 MQ 消费者框架。
  - **多模态模型集成：** 封装视觉处理引擎与 TTS 语音合成，执行幻灯片图像理解与音频流生成等重负载 AI 任务。
  - **状态反馈模块：** 封装向进度 MQ ( `progress_update` ) 投递状态和错误信息的标准化 SDK。
- **基建中间件配置：**
  - **MQ (Kafka/RabbitMQ)：** 配置优先级隔离的 Topic 集群（如 `queue:vip_long` , `queue:progress` ）。
  - **MySQL：** 新增 `async_task` 核心调度表（包含乐观锁 Version 字段）。

### 🏆 Phase 2 交付标准 (Definition of Done):

- ☒ 用户发起耗时任务后，前端秒级获取 TaskID 并无阻塞返回其他页面。

- ☒ 任务在 Python 节点执行期间，前端进度条能平滑刷新（误差 < 10%）。
- ☒ 若强杀一个 Python 进程，未完成的任务能被 MQ 重新调度到其他节点执行，不丢失。
- ☒ 最终生成的音视频/多媒体产物能成功上传 OSS，并在前端提供下载或播放链接。