

# 审评需求分析

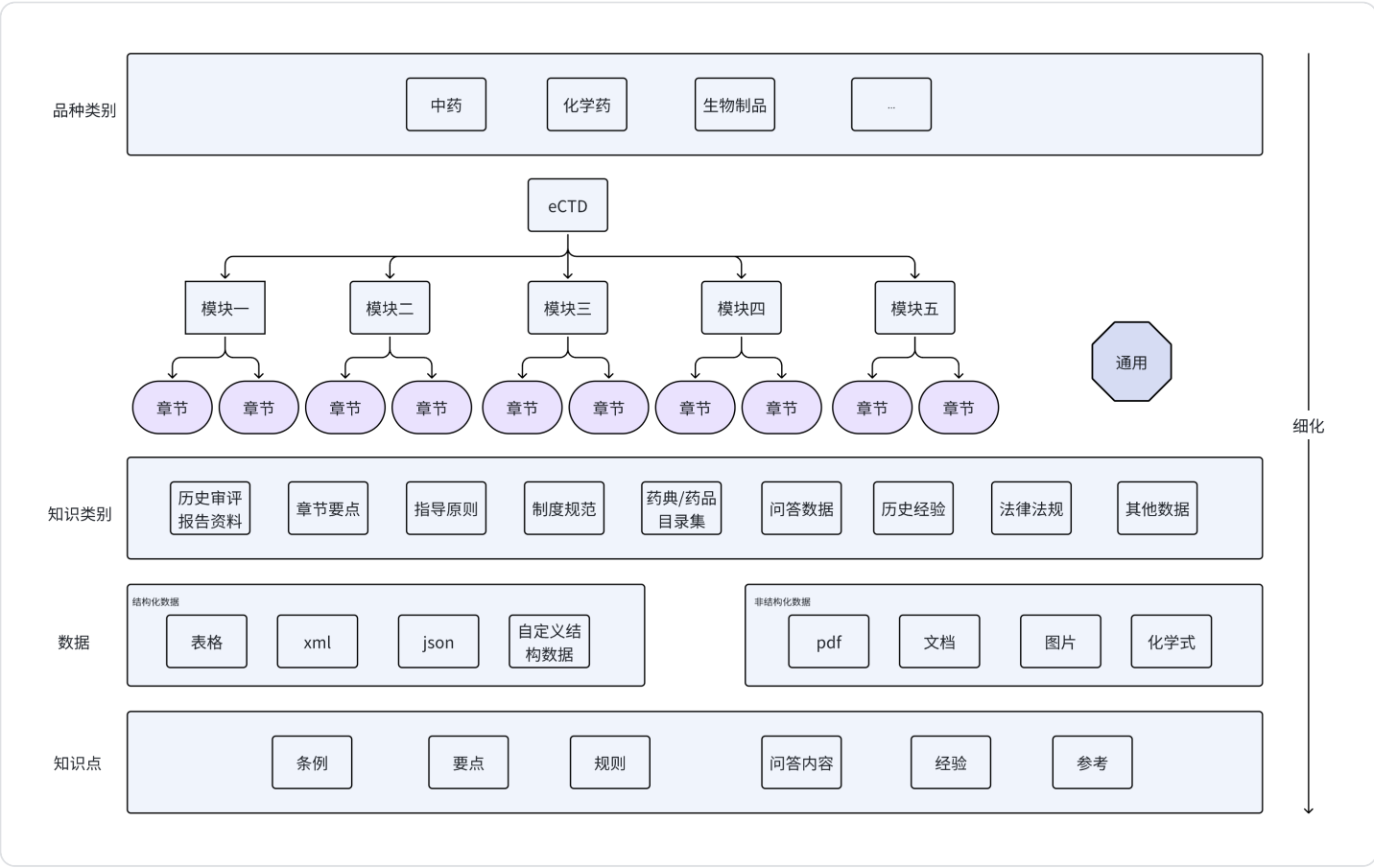
## 知识库设计

eCTD（电子通用技术文档）在结构上被划分为五个主要模块，每个模块进一步细分为多个章节，每个章节都专注于特定的内容。在药品审评过程中，通常会按照这些章节（或更细致的分类）来进行审评工作。药品主要分为中药、化学药品和生物制品三大类，同类药品在eCTD的各个章节审评过程中所依赖的参考资料和经验具有很高的可复用性。故知识库的搭建可以根据章节进行统计分类，以确保在审评过程中能够快速定位最相关审评依据以及相关数据，根据调研统计分析，审评数据可以分为如下大类数据：

1. 按照章节（或者更细粒度）拆分之后eCTD原始资料
2. 历史审评报告资料(仿制药可做参考)
3. 章节要点
4. 审评准则（CDE，FDA）：
5. 制度规范（CDE）
6. 法律法规（CDE，FDA）
7. 历史经验（使用系统时产生）
8. 问答数据
9. 药典/药品目录集
10. 外部其他数据

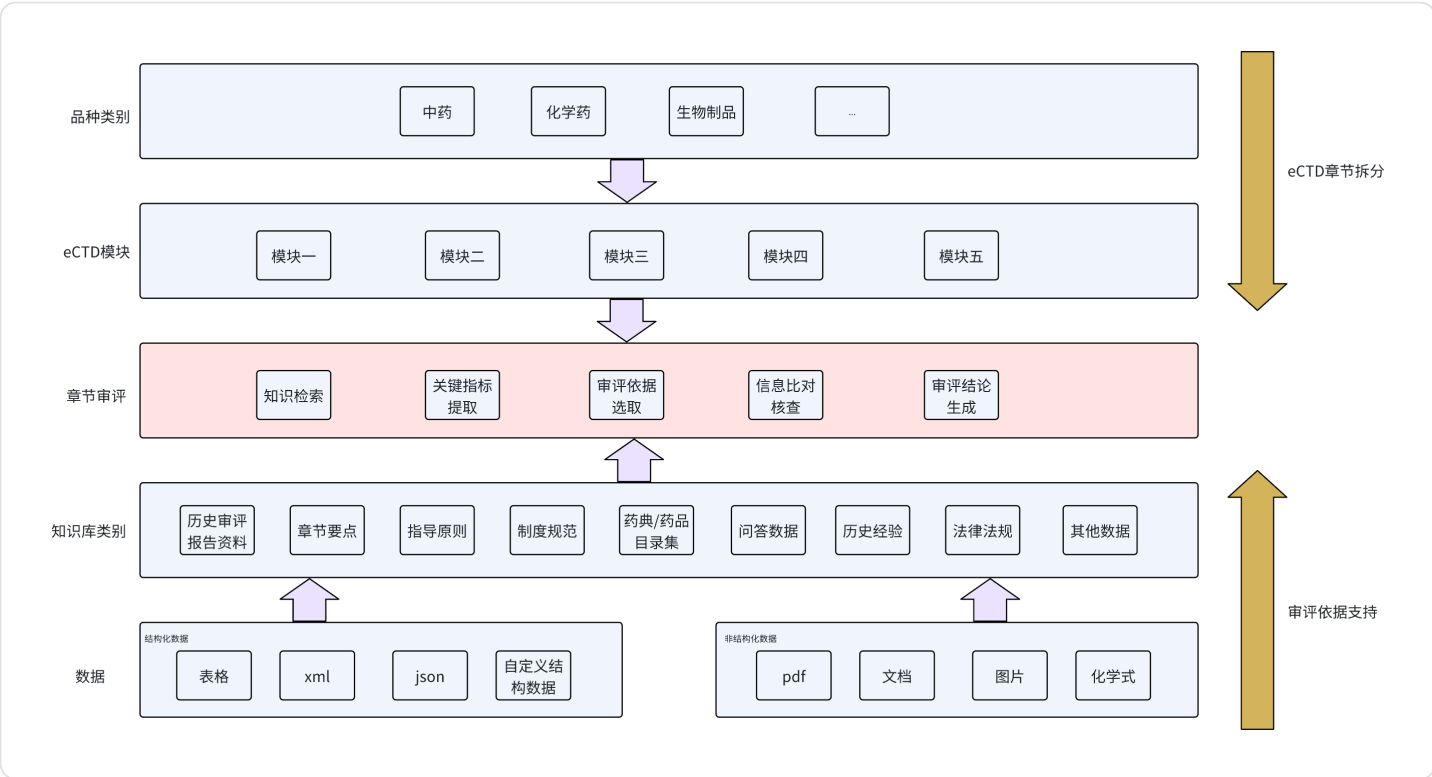
目前已整理（持续完善中）：[📁 审评知识分类示例](#)

知识库数据组织形式：



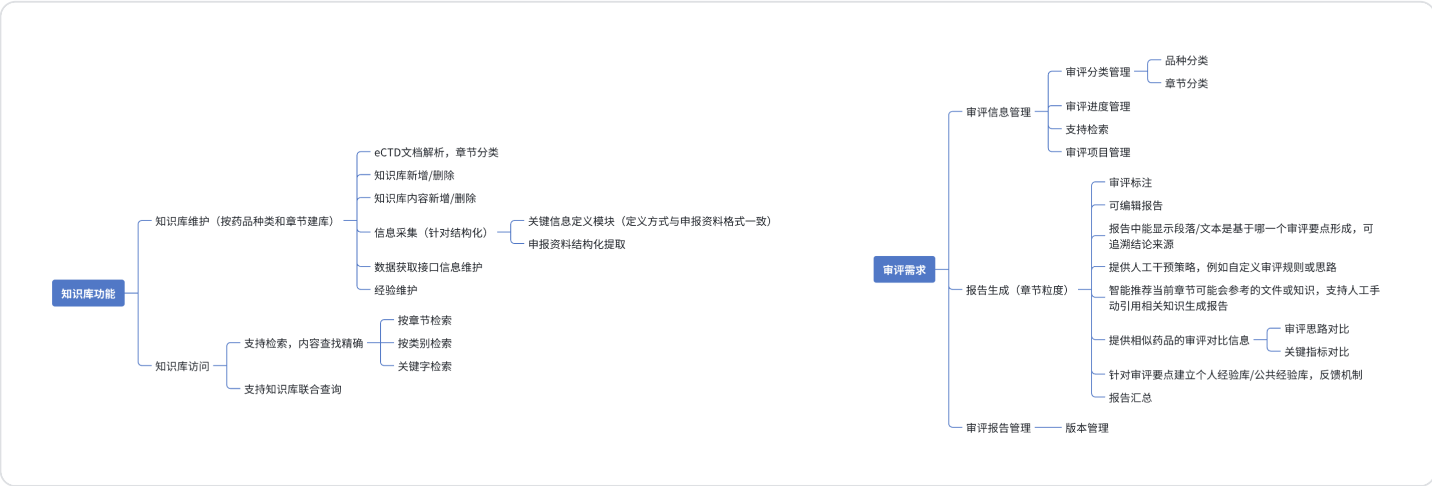
知识组织形式

知识库的结构设计以审评材料eCTD的组织结构作为驱动，更加符合审评场景的需求，站在审评人员的角度，在面对审评任务时首先需要区分当前审评的药品类型，不同类型的药物审评需要的知识和数据有所不同，在确定好具体的药品类别后，审评人员需要按照eCTD的结构去组织和查找报告的内容是否符合一定的标准和要求，所以二级类别可以按照审评资料的章节来组织分类，每个章节下的相关知识为该章节的审评提供依据（该类别的粒度划分可以参照审评人员真实审评情况进行划分），随后按照审评时不同的数据种类进行划分，每个类别下会存在结构化数据和非结构化数据。由于文档的粒度较大，更进一步，可以考虑将文档拆分成细粒度知识，使得审评依据聚焦于某一章节审评内容。



业务视角知识使用方式

# 功能设计



功能结构图

## 功能介绍

### 知识库需求管理

知识库需求管理模块用于知识库的维护和管理，通过整合外部及内部知识资源，优化审评过程中的信息获取。该模块包含以下功能：

#### 1. 知识库维护

- eCTD文档解析和分类
  - 支持eCTD文档解析，按类别和章节分类，提升文档管理的效率。
- 知识库新建/删除

- 知识库建设以eCTD结构章节划分，分为章节知识库和通用知识库。
- 知识库支持分级分类，在章节分类的基础上区分不同数据类型，在知识新增时，需要选择对应的章节、知识类别。
- 知识的粒度支持以文件为单位以及更细粒度的拆分
- 知识库新建/删除需获取权限
- **知识库内容新增/删除**
  - 新增知识
  - 允许删除已过期或不再需要的知识，以保持知识库的更新和精简。
  - 知识库内容的新增/修改需获取权限
- **结构化数据采集**
  - 用户可自定义信息抽取的schema字段
  - 基于自由文本的结构化数据提取。
- **数据接口维护**：为知识库提供数据获取接口，并对其进行维护，确保信息能够及时、准确地导入系统。
- **经验维护**：在审评过程中支持用户经验留存（包括审评思路/参考文件）

## 2. 知识库访问

- **支持检索**：
  - 支持按照章节检索当前章节相关的系列分类知识
  - 系统提供精确的关键词搜索和内容查找功能，使得审评人员能够快速找到所需的知识点。
- **联合查询**：系统支持跨知识库的联合查询，方便用户从不同的知识源中获取全面的信息支持。

## 审评需求管理

审评需求管理模块是为了优化和组织审评过程中的各项需求而设计的。该模块分为以下几个功能子模块：

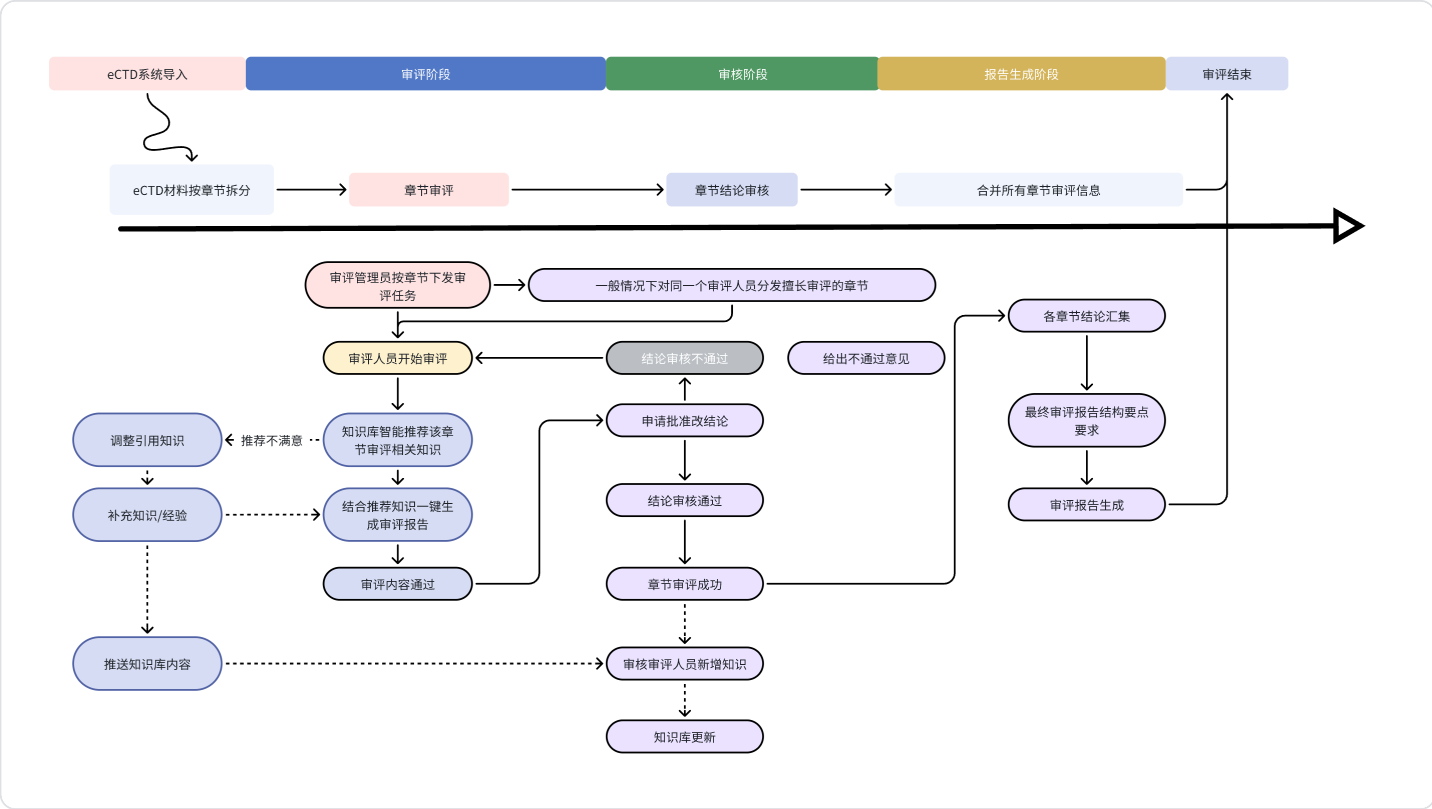
### 1. 审评信息管理

- **审评分类管理**：按照药品类别进行管理，单个药品按照章节拆分，可以快速定位到每一章节的相关信息。
- **审评进度管理**：实时追踪和管理审评过程的进度，确保每个审评阶段按时完成。
- **支持检索**：提供关键词或条件搜索功能，方便快速定位到相关的审评信息。
- **审评项目管理**：集中管理多个审评项目，便于审评人员全面了解项目进展和状态。

### 2. 报告生成（章节粒度）

- **审评标注：**针对审评的各个要点提供标注功能，使得审评者能够清晰地标记和总结报告中的重要信息。
  - **报告可编辑：**支持用户对生成的报告进行编辑和修改，满足个性化需求。
  - **报告显示逻辑：**报告中能够展示文本的段落或结构，且能够基于审评要点的形式形成，并能溯源至相应的结论来源。
  - **人工干预：**允许人工设定审评规则或路径，进行个性化定制，审评人员可以自由调整当前审评内容的参考依据。
  - **相关知识智能推荐：**
    - 默认推荐的知识参与报告生成
    - 支持人工手动选择相关知识生成报告
  - **对比分析：**支持提供相似药品或产品的审评对比信息，帮助审评者在对比中做出决策。例如，可以进行审评路径对比或关键指标对比。
  - **经验库建立：**为审评要点建立个人经验库或公共经验库，帮助总结反思机制，提升审评效率。
  - **报告汇总：**针对各个章节报告汇总成总报告。
3. **审评报告管理** 审批报告管理模块用于审评报告的存储、版本控制及审批流程的记录，确保报告的完整性和可追溯性。

设计原型（原型图界面待设计）



业务流程图

1. 首页 Dashboard

- **页面功能：**该页面是系统的首页仪表盘，提供审评项目、报告状态、知识库更新等内容的总览。
  - **页面布局：**
    - **顶部导航栏：**包含以下菜单项：
      - 审评需求管理
      - 知识库需求管理
      - 报告管理
      - 用户设置
      - 通知中心
    - **左侧边栏：**提供快速访问的主要模块按钮：
      - 审评信息管理
      - 报告生成
      - 知识库访问
    - **中央主要区域：**
      - **审评项目进度卡片：**展示正在进行的审评项目列表，每个卡片包括项目名称、当前进度条(%)、负责人、截止日期等。
      - **报告状态卡片：**展示最近生成的报告列表，包含报告名称、生成时间、状态（已提交、待修改、草稿等）。
      - **知识库更新：**最近更新的知识库条目，显示标题、更新时间、来源。
    - **右侧边栏：**快捷操作按钮，包含“新建审评项目”、“生成报告”、“查看审评历史”等。
- 

## 2. 审评信息管理页面

- **页面功能：**管理和查看所有的审评项目信息，支持搜索、过滤功能。
- **页面布局：**
  - **顶部搜索栏：**
    - **搜索框：**支持关键词输入（按项目名称、负责人、状态等条件检索）。
    - **高级过滤选项：**支持按审评状态（已完成、进行中、已取消）、日期范围、审评员等筛选条件过滤。
  - **审评项目列表：**
    - 列表视图：每个审评项目以行方式展示，包括以下字段：
      - 项目名称
      - 审评状态（进行中、已完成、暂停等）

- 负责人
- 当前进度（以百分比或进度条形式展示）
- 截止日期

- **项目操作按钮：**每个项目行旁边有“查看详情”、“编辑”、“删除”等按钮。

- **分页控件：**项目列表底部有分页按钮，允许用户切换页面查看更多项目。
  - **右上角按钮：**页面右上角有“创建新审评项目”按钮，点击后进入新项目创建页面。
- 

### 3. 审评标注与报告生成页面

- **页面功能：**生成和编辑审评报告，支持人工干预，报告标注与路径设定，并支持人工调整参考数据。
  - **页面布局：**
    - **左侧边栏：**
      - **文档列表：**展示当前审评项目的文档，支持折叠和展开功能。
      - **标注工具：**提供标注工具，用于高亮和标记文本中的关键审评要点。
      - **路径与策略选择：**提供选择审评路径和预定义策略的工具（下拉菜单或切换按钮），也支持自定义审评路径。
    - **中央编辑区：**
      - **文档查看区域：**显示完整的审评文档或结构化数据，支持拖动滚动查看全文。
      - **标注显示：**已标注内容在文档中以不同颜色高亮展示，用户可以点击标注区域查看详细的标注信息（如备注、来源等）。
      - **报告编辑器：**下方或侧边有报告生成编辑框，支持自由文本输入、表格添加、图片插入等编辑功能。
    - **人工干预策略设置：**
      - 提供预设的审评策略模板或用户自定义的规则，用户可以选择或编辑这些规则，并直接应用到审评路径中。
- 

### 4. 知识库管理页面

- **页面功能：**用于管理知识库，支持文档的解析、分类、删除和外部知识整合。
- **页面布局：**
  - **顶部搜索栏：**
    - **搜索框：**关键词检索，用户可以通过输入关键词查找知识库内容。

- **筛选选项：**支持按文档类型（eCTD文档、外部文档等）、创建时间、分类标签进行筛选。
  - **知识库条目列表：**
    - 每条目包含：
      - 知识标题
      - 分类
      - 创建时间
      - 来源
    - 条目右侧操作按钮：“查看”、“编辑”、“删除”等。
  - **文档管理工具：**
    - 支持eCTD文档解析和分类，用户可以上传新文档，系统会自动解析章节并进行分类。
    - **外部知识导入按钮：**允许导入外部知识库，并根据来源进行分类管理。
- 

## 5. 知识库查询页面

- **页面功能：**查询知识库中的信息，支持跨知识库的联合查询。
  - **页面布局：**
    - **顶部搜索框：**
      - 输入关键词，支持模糊查询或精确匹配或向量搜索。
      - **高级筛选选项：**根据文档来源、时间范围、分类标签进行过滤。
    - **结果展示区域：**
      - 以列表形式展示查询结果，内容包括知识标题、来源、创建时间、相关文档链接等。
      - 每个条目右侧提供“查看详情”按钮，点击后进入详细信息页面。
    - **联合查询选项：**在搜索框下方，用户可以选择不同的知识库，进行多源联合查询，结果按来源分类展示。
- 

## 6. 系统设置与用户管理页面

- **页面功能：**配置系统设置、管理用户权限。
- **页面布局：**
  - **左侧菜单：**系统设置、用户管理、操作日志等模块导航。
  - **用户管理区域：**
    - **用户列表：**展示系统中所有用户，包含用户名、角色（审评员、管理员等）、权限级别。

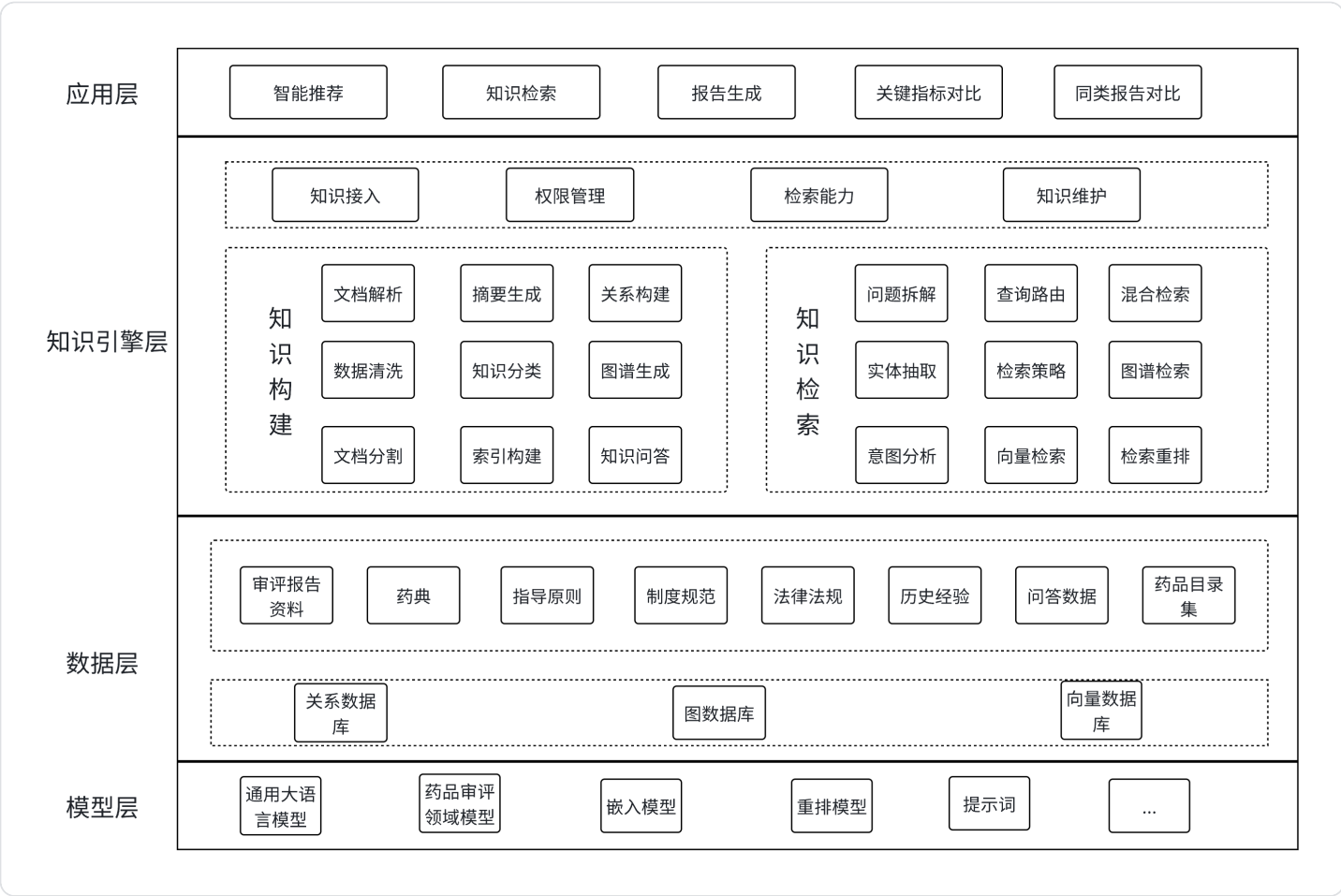


- 每个用户行的右侧提供“编辑权限”、“删除用户”等操作按钮。
  - **系统设置：**支持系统通知、审评策略模板配置等全局设置选项。
  - **操作日志：**展示最近的系统操作日志，包括用户登录、数据修改等记录。
- 

## 7. 数据接口管理页面

- **页面功能：**管理知识库和审评数据的接口配置，确保数据正常导入。
- **页面布局：**
  - **接口列表：**
    - 每个接口包含以下信息：接口名称、来源、数据类型、数据分类、状态（启用/禁用）。
    - 列表右侧提供“查看详情”、“编辑”、“删除”等操作按钮。
  - **接口配置页面：**
    - **接口详细信息：**展示接口的具体配置，包括数据同步频率、格式要求等。
    - **数据同步设置：**支持调整数据获取频率、自动同步设置等。
  - **接口日志：**展示该接口的调用记录和数据同步日志，帮助用户跟踪数据传输情况。

## 知识库技术方案设计



整体架构图

## 模型层

- 领域数据学习
- 嵌入模型
- 重排模型
- 提示词工程

## 数据层

数据层主要包含两部分内容，一是针对知识库需要确定包含的知识具有有哪些类别，同类别的数据统一组织建库，二是根据数据特点需要确定每类数据的索引构建方式和存储形式。数据层的实现主要关注以下方面：

### 数据预处理

- 数据类别分类
  - pdf（包含文字、表格、图片）
    - 标题提取
    - 表格提取
    - 图片提取

- 向量索引+倒排索引
- QA
  - 问题关键字提取
  - 意图
  - 向量索引+倒排索引
- 结构化数据（药典等）
  - 结构梳理
  - 结构化存储+参与命中部分倒排

## 知识引擎层

## 应用层