**湖南农业大学**

**专业综合训练**

**软件开发项目报告**

[数据脱敏治理中心]

**□□□□□□学生姓名：×××**

**□□□□□□学□□号：20××××××××××**

**□□□□□□年级专业及班级：2020级 计算机科学与技术 (1)班**

**□□□□□□校内指导老师及职称：xxx xxxx**

**校外指导老师及职称：xxx xxxx**

**□□□□□□学□□院：信息与智能科学技术学院**

提交日期：20xx年x月

目 录

[1 概述 1](#_Toc198336482)

[1.1 项目背景 1](#_Toc198336483)

[1.2 项目目标 1](#_Toc198336484)

[1.2.1短期目标（项目开发阶段）： 1](#_Toc198336485)

[1.2.2长期目标（产品迭代阶段）： 1](#_Toc198336486)

[2 需求分析 2](#_Toc198336487)

[2.1 用户需求 2](#_Toc198336488)

[2.1.1 用户角色划分 2](#_Toc198336489)

[2.1.2 用户需求说明 2](#_Toc198336490)

[2.1.3 用例图示例 3](#_Toc198336491)

[2.2 功能需求分析 3](#_Toc198336492)

[2.3 非功能需求分析 4](#_Toc198336493)

[3 系统概要设计 4](#_Toc198336494)

[3.1 系统整体架构设计 4](#_Toc198336495)

[3.2 系统技术架构设计 4](#_Toc198336496)

[3.3 系统功能结构设计 5](#_Toc198336497)

[3.4 数据库设计 6](#_Toc198336498)

[3.4.1 概念结构设计 6](#_Toc198336499)

[3.4.2 逻辑结构设计 8](#_Toc198336500)

[3.4.3 物理结构设计 8](#_Toc198336501)

[4 系统详细设计 8](#_Toc198336502)

[4.1 推荐模块设计 8](#_Toc198336503)

[4.1.1 界面原型设计 8](#_Toc198336504)

[4.1.2 接口设计 8](#_Toc198336505)

[4.2 订单模块设计 8](#_Toc198336506)

[4.2.1 界面原型设计 9](#_Toc198336507)

[4.2.2 接口设计 9](#_Toc198336508)

[5 系统实现 9](#_Toc198336509)

[5.1 推荐模块实现 9](#_Toc198336510)

[5.2 订单模块实现 9](#_Toc198336511)

[6 系统测试 9](#_Toc198336512)

[7 总结与展望 9](#_Toc198336513)

[参考文献 9](#_Toc198336514)

[数据脱敏治理中心]

# 概述

## 项目背景

随着数据合规要求日益严格，尤其是《中华人民共和国个人信息保护法》《网络安全法》等法规的实施，企业与组织在数据处理和存储过程中面临日益严峻的隐私保护挑战。传统的数据管理系统在处理敏感信息时缺乏有效的控制手段，容易引发数据泄露、越权访问等安全问题。

为了实现数据的最小可用原则与“脱敏后可用、还原不可逆”的数据使用目标，迫切需要构建一套统一、高效、可管可控的数据脱敏治理平台，对存储在数据库、文件系统、日志等多种场景下的敏感数据进行自动识别、策略化脱敏与访问控制，降低数据泄露风险，提升合规能力。

“数据脱敏治理中心”项目正是在这一背景下提出，旨在通过技术手段和治理机制双轮驱动，建立企业级的数据脱敏治理能力，实现敏感数据的全生命周期安全管理。

## 项目目标

“数据脱敏治理中心”项目旨在搭建一个面向政企用户的可扩展、模块化、具备集中管控能力的数据脱敏平台，核心目标如下：

### 1.2.1短期目标（项目开发阶段）：

1. 实现一套可视化的脱敏策略配置与管理系统，支持常见脱敏算法（如掩码、加密、替换、哈希等）；
2. 实现对主流数据库（如MySQL、PostgreSQL）中敏感字段的自动识别与分类标记；
3. 构建统一的数据访问中间层，支持基于策略的实时脱敏展示；
4. 完成权限管控、操作审计、日志追踪等基本治理功能；
5. 提供RESTful API，支持前后端分离开发。

### 1.2.2长期目标（产品迭代阶段）：

1. 支持更多数据源类型，如MongoDB、大数据平台（如Hadoop、Hive）及日志系统；
2. 支持可插拔的自定义脱敏插件机制；
3. 实现基于机器学习的敏感信息智能识别能力；
4. 建立数据脱敏效果评估与策略优化建议机制；
5. 实现与数据安全、数据血缘、数据资产等系统的联动，构建完整的数据治理生态。

通过该项目的实施，预期大幅提升系统的数据安全等级，实现企业数据资产合规使用、安全共享与有效治理，为今后相关系统建设提供技术与治理支撑。

# 需求分析

## 用户需求

为确保系统功能符合实际业务场景，本项目从用户角度出发，划分角色并梳理各角色的核心需求。

### 用户角色划分

表 1 用户角色表

|  |  |
| --- | --- |
| **角色** | **描述** |
| 系统管理员 | 负责系统的初始化配置、权限分配与用户管理 |
| 数据安全管理员 | 负责敏感字段识别、脱敏策略制定与数据访问审计 |
| 数据使用人员 | 面向业务系统的数据使用者，根据权限申请数据脱敏，并在申请通过后查看被脱敏的数据 |
| 审计人员 | 负责系统日志、安全事件和脱敏操作记录的审查 |

### 用户需求说明

* 系统管理员
* 添加/删除/修改用户与角色
* 分配权限与资源访问范围
* 管理系统参数与运行环境配置
* 数据安全管理员
* 识别并标记敏感数据字段
* 创建、修改和应用脱敏策略
* 配置访问控制策略与数据标签
* 审查数据使用记录，查看操作日志，生成审计报告
* 数据使用人员
* 发起数据脱敏以及使用申请
* 通过前端页面或 API 访问所需数据
* 根据权限查看或使用脱敏后的数据
* 审计人员
* 审查敏感数据使用记录、变更记录
* 导出系统运行与审计报告

### 用例图示例

图示

AI 生成的内容可能不正确。

图 1 用例图

## 功能需求分析

系统应至少具备以下核心功能模块：

表 2 核心功能模块

|  |  |
| --- | --- |
| **功能模块** | **功能描述** |
| 用户与权限管理 | 支持用户注册、角色配置、权限粒度控制、登录认证 |
| 敏感数据识别 | 自动扫描数据库字段，根据规则或模型识别敏感信息，如身份证、手机号、银行账号等 |
| 脱敏策略管理 | 支持多种脱敏方式（掩码、哈希、加密、替换、自定义插件），策略可配置、可调度 |
| 数据脱敏服务 | 支持实时/离线脱敏，提供 API 接口对接业务系统，返回脱敏后数据 |
| 日志审计与跟踪 | 记录所有策略应用、数据访问和系统操作日志，供审计人员查验 |
| 系统配置管理 | 包括日志保留周期、任务调度周期、脱敏算法参数等配置项 |
| 报表与可视化 | 输出数据脱敏情况统计、用户访问行为分析图表，支持导出 PDF/Excel |

## 非功能需求分析

表 3 非功能性需求模块

|  |  |
| --- | --- |
| **类别** | **需求描述** |
| 性能 | 系统支持并发 100+ 用户访问，数据处理响应时间 <1s（实时请求场景）；可扩展性强 |
| 安全性 | 系统应支持基于角色的访问控制（RBAC）、数据传输加密（HTTPS）、策略修改留痕 |
| 兼容性 | 支持 MySQL、PostgreSQL、Oracle 等主流数据库，API 兼容 RESTful 标准 |
| 易用性 | 提供操作指引与帮助文档，前端界面简洁，配置流程可视化、支持策略模板复用 |
| 可维护性 | 模块解耦，支持插件化策略扩展，日志详细，异常处理完善 |
| 可部署性 | 支持 Docker 镜像部署，适配主流 Linux 发行版，提供安装脚本与部署文档 |

# 系统概要设计

## 系统整体架构设计

图示

AI 生成的内容可能不正确。

图 2 系统整体架构设计图

## 系统技术架构设计

“数据脱敏治理中心”系统采用**前后端分离架构**，主要包括前端、后端和数据库三大部分：

* **前端部分**：使用 Vue3 框架开发，负责为用户提供友好的操作界面，功能包括数据申请、脱敏结果查看、导出数据等。前端与后端通过 RESTful API 接口进行数据交互，未来也可以基于此封装成 App 应用。
* **后端部分**：采用 Spring Boot 框架开发，负责处理系统的核心业务逻辑，如用户权限管理、脱敏策略匹配、数据处理等。同时整合 JWT 进行用户认证，保证系统安全。
* **数据库部分**：使用 MySQL 存储系统中的用户信息、脱敏规则、申请记录、审计日志等。

## 系统功能结构设计

本系统主要包括四类角色：系统管理员、数据安全管理员、数据使用人员、审计人员。系统初始仅由**系统管理员**登录使用，普通用户在注册后需由系统管理员分配角色与权限，才可成为特定类型用户并使用相应功能。

系统功能结构划分如下：

* **系统管理员功能模块**
  + 用户注册审批与角色分配
  + 用户管理（添加、修改、删除用户信息）
  + 权限配置与访问控制
  + 系统参数设置与运行配置
  + 日志查看与系统运行状态监控
* **数据安全管理员功能模块**
  + 敏感字段识别与管理
  + 脱敏策略创建与修改（如掩码、替换、加密等）
  + 数据访问策略配置（设置谁能访问、访问何种数据）
  + 查看操作日志、生成审计报告
* **数据使用人员功能模块**
  + 提交脱敏数据申请（上传数据、选择脱敏方式）
  + 查看脱敏处理后的数据
  + 下载脱敏结果数据
  + 查询历史申请与处理记录
* **审计人员功能模块**
  + 查看系统操作日志与使用记录
  + 审查敏感数据访问与策略调整记录
  + 导出审计报告



图 3 系统功能结构图

## 数据库设计

### 概念结构设计

**实体及关系描述：**

* **用户（User）**：系统登录用户，可上传原始数据、发起脱敏任务。
* **角色（Role）**：定义用户类型，例如：普通用户、系统管理员、数据安全员、审计人员等。
* **权限（Permission）**：指定角色所拥有的功能权限。
* **数据表（SensitiveData）**：保存用户上传的原始数据及其脱敏结果，区分访问权限。
* **脱敏策略（MaskingPolicy）**：定义脱敏处理方法。
* **脱敏任务（MaskingTask）**：记录每次脱敏的执行记录及状态。
* **操作日志（OperationLog）**：记录系统操作行为。
* **审计报告（AuditReport）**：生成审计人员对操作行为的汇总与分析报告。

为了方便绘制 E-R 以及划分各个实体之间的关系，先将其简单划分为下面三个模块，单独绘制各个模块的 E-R 图

#### 登录与权限模块

涉及用户身份认证、角色权限管理相关实体。包含实体Role、User、Permission

角色（Role）:（角色ID，角色描述，角色名）

外键：（角色）角色ID → 角色权限（角色ID）

属性含义：

* 角色ID：唯一标识一个角色类型，在 [1,4] 区间之中
* 角色名称：有且只有系统管理员、数据安全管理人员、数据使用人员（普通用户）、日志审计员四种，与角色ID [1,4] 相互对应，默认设置为数据使用人员（普通用户）。
* 描述：角色的职责说明。

模块意义：

将权限归类，用于统一分配给多个用户。

用户（User）：(\_用户ID\_, 用户名, 密码, 邮箱，角色ID)

外键：（用户）角色ID → 角色(角色ID)

属性含义：

* + - 用户ID：系统用户唯一标识。
    - 用户名：用户登录名。
    - 密码：用户登录密码。
    - 邮箱：用户邮箱，用于通知与找回密码。
    - 角色ID：用户所归属角色,默认设置为 4.

模块意义：

管理系统登录账号及其角色，控制访问权限的基础。

权限(Permisson)：（权限ID，权限描述）

属性含义：

* 权限编码：唯一权限标识（如：data:upload）。
* 权限描述：权限功能描述（如：上传数据）。

模块意义：

用于精细控制角色可执行的具体操作。

角色权限（Role\_Permission）:（角色ID，权限ID）【由关系转换而来】

外键：（角色权限）角色ID → 角色（角色ID）

（角色权限）权限ID → 权限（权限ID）

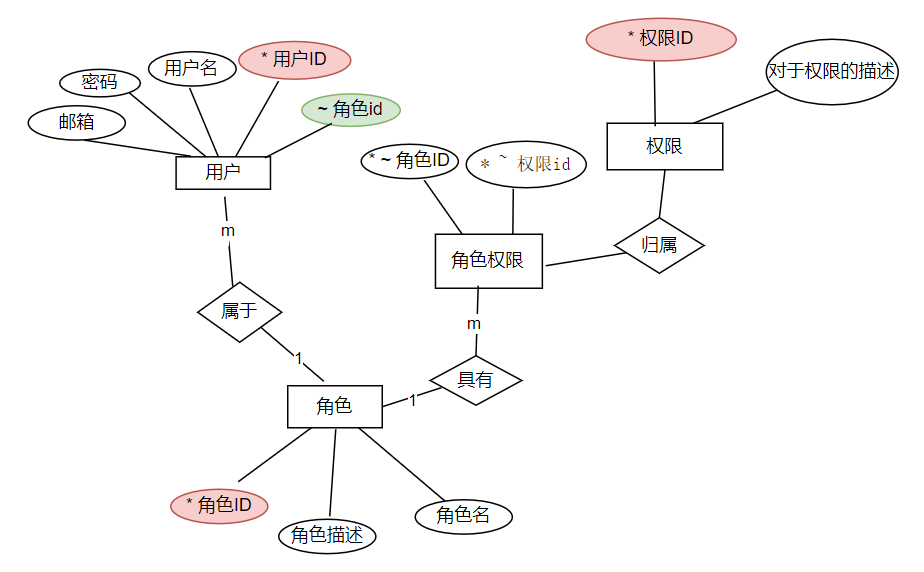


图 4 登录与权限模块E-R图

#### ****脱敏数据管理模块****

涉及脱敏数据存储、脱敏策略、脱敏任务。包含实体SensitiveData、MaskingPolicy、MaskingTask

敏感数据（SensitiveData）:(\_数据ID\_, 身份证号, 上传用户ID, 姓名, 性别, 手机号, 邮箱, 银行卡号, 上传时间)

外键：敏感数据（上传用户ID） → 用户（用户ID）

属性含义：

* 数据ID：唯一记录编号。
* 身份证号、姓名、性别等：用户上传的敏感原始信息。
* 上传时间：记录上传时间。
* 上传者ID：记录是谁上传的。

模块意义：

存储待脱敏和已脱敏的数据内容。

脱敏策略（MaskingPolicy）: （\_策略ID\_, 策略名称, 目标字段, 脱敏规则, 脱敏描述）

属性含义：

* 策略ID：唯一标识。
* 策略名称：如“手机号掩码策略”。
* 目标字段：指定应用该策略的字段。
* 脱敏规则：用于描述如何脱敏，如“保留前3后4位”。
* 描述：策略详细说明。

模块意义：

定义如何对敏感数据字段进行处理。

脱敏任务（MaskingTask）：（\_脱敏任务ID\_, 任务状态, 执行时间， 脱敏策略ID, 脱敏数据ID）

外键：脱敏任务（脱敏数据ID） → 敏感数据（数据ID）

脱敏任务（脱敏策略ID） → 脱敏策略（策略ID）

属性含义：

* 任务ID：唯一记录每一次脱敏操作。
* 状态：如“待处理”、“完成”、“失败”。
* 执行时间：任务实际运行时间。
* 脱敏数据ID/脱敏策略ID：指定作用的对象和策略。

模块意义：

记录每次脱敏操作的执行记录，便于追踪与统计。

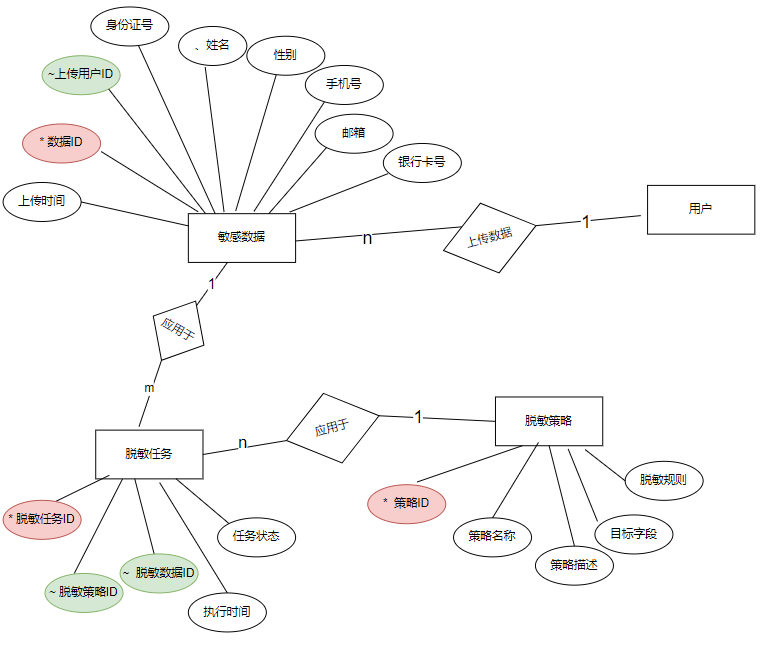


图 5 脱敏数据管理模块E-R图

#### 审计与日志模块

包括实体：AuditLog（审计日志）、OperationLog（操作日志）User（悬空，仅连接线）

操作日志（OperationLog）：(\_操作日志ID\_, 操作类型, 操作用户ID, 操作时间, 操作详细内容)。

外键：操作日志（操作用户ID） → 用户(用户ID)

属性含义：

* 日志ID：每条操作行为唯一标识。
* 操作类型：如“新增数据”、“发起任务”。
* 操作详细内容：记录操作的对象以及操作的内容。
* 操作时间：记录时间戳。
* 操作用户ID：操作者。

模块意义：

记录系统所有敏感操作的行为，支撑审计分析。

审计日志（AuditLog）：(\_审计日志ID ， 审计人员ID， 审核时间, 报告摘要，操作日志ID)

外键：审计日志（审计人员ID） → 用户(用户ID)

审计日志（操作日志ID） → 操作日志（操作日志ID）

属性含义：

* 审计日志ID：唯一标识一份审计报告。
* 审核时间：生成报告的时间。
* 报告摘要：总结日志分析结果。
* 审计人员ID：编写或导出报告的人员。
* 操作日志ID：引用被审计的具体操作日志

模块意义：

面向审计人员的总结与分析报告管理。

图示

AI 生成的内容可能不正确。

图 6 审计与日志模块E-R图

#### 总体E-R模型设计

将 3.4.1.1节、 3.4.1.2节和3.4.1.3节中的局部 E-R 图进行合并，消除冗余实体与冗余联系，得到如下的总体 E-R 图

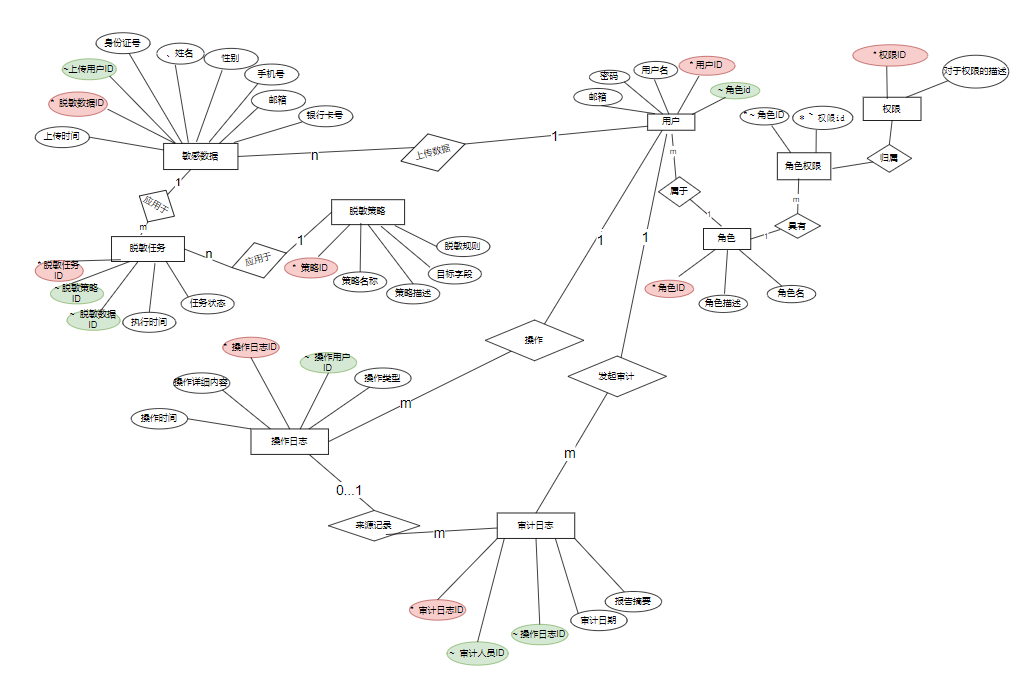


图 7 总体E-R图

### 逻辑结构设计

以关系模式形式（关系名 + 属性列表）列出如下：

Role(角色ID, 角色描述, 角色名称)

User(用户ID, 用户名, 密码, 邮箱, 角色ID)

外键：角色ID 引用 Role(角色ID)

Permission(权限ID, 权限描述)

Role\_Permission(角色ID, 权限ID)

外键：角色ID 引用 Role(角色ID)

权限ID 引用 Permission(权限ID)

SensitiveData(数据ID, 身份证号, 上传用户ID, 姓名, 性别, 手机号, 邮箱, 银行卡号, 上传时间)

外键：上传用户ID 引用 User(用户ID)

MaskingPolicy(策略ID, 策略名称, 目标字段, 脱敏规则, 脱敏描述)

MaskingTask(脱敏任务ID, 任务状态, 执行时间, 脱敏策略ID, 脱敏数据ID)

外键：脱敏策略ID 引用 MaskingPolicy(策略ID)

脱敏数据ID 引用 SensitiveData(数据ID)

OperationLog(操作日志ID, 操作类型, 操作用户ID, 操作时间, 操作详细内容)

外键：操作用户ID 引用 User(用户ID)

AuditLog(审计日志ID, 审计人员ID, 审核时间, 报告摘要, 操作日志ID)

外键：审计人员ID 引用 User(用户ID)

操作日志ID 引用 OperationLog(操作日志ID)

### 物理结构设计

表 4 用户表(User)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| user\_id | 用户ID | SERIAL | PK | 无 |
| username | 用户名 | VARCHAR(20) | NOT NULL, UNIQUE | 无 |
| password | 密码 | VARCHAR(20) | NOT NULL | 无 |
| email | 邮箱 | VARCHAR(20) | NOT NULL, CHECK (email LIKE '%@%'), UNIQUE | 无 |
| role\_id | 所属角色 |  | FOREIGN KEY REFERENCES Role(role\_id)，CHECK (role\_id BETWEEN 1 AND 4) | 4（数据使用人员） |

表 5 角色表(Role)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| role\_id | 角色ID | INT | PK ，CHECK (role\_id BETWEEN 1 AND 4) | 无 |
| role\_name | 角色名称 | VARCHAR(20) | NOT NULL, UNIQUE | 无 |
| role\_desc | 角色的职责说明 | VARCHAR(200) |  | 无 |

表 6 权限表(Permission)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| permission\_id | 权限编码 | VARCHAR(20) | PK | 无 |
| permission\_desc | 权限功能描述 | VARCHAR(500) |  | 无 |

表 7 角色权限关系表(Role\_Permission)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| role\_id | 角色ID | INT | FOREIGN KEY REFERENCES Role(role\_id), | 无 |
| permission\_id | 权限编码 | VARCHAR(20) | FOREIGN KEY REFERENCES Permission(permission\_id), | 无 |

其中(role\_id, permission\_id)为联合主键

表 8 敏感数据表(SensitiveData)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| data\_id | 数据ID | SERIAL | PK | 无 |
| id\_number | 身份证号 | VARCHAR(18) |  | 无 |
| name | 姓名 | VARCHAR(20) |  | 无 |
| gender | 性别 | VARCHAR(10) |  | 无 |
| phone\_number | 手机号 | VARCHAR(20) |  | 无 |
| email | 邮箱 | VARCHAR(100) | CHECK (email LIKE '%@%') | 无 |
| bank\_card | 银行卡号 | VARCHAR(30) |  | 无 |
| upload\_time | 上传时间 | DATATIME | NOT NULL | CURRENT\_TIMESTAMP |
| uploader\_id | 上传用户ID | INT | FOREIGN KEY REFERENCES User(user\_id), NOT NULL | 无 |

表 9 脱敏策略表(MaskingPolicy)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| policy\_id | 策略ID | SERIAL | PK | 无 |
| policy\_name | 策略名称 | VARCHAR(20) | NOT NULL, UNIQUE | 无 |
| target\_field | 目标字段 | VARCHAR(50) | NOT NULL | 无 |
| rule | 脱敏规则 | VARCHAR(200) | NOT NULL | 无 |
| policy\_desc | 策略详细说明 | VARCHAR(500) |  | 无 |

表 10 脱敏任务表(MaskingTask)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| task\_id | 脱敏任务ID | SERIAL | PK | 无 |
| task\_status | 任务状态 | VARCHAR(20) | NOT NULL, CHECK (task\_status IN ('待处理', '完成', '失败')) | “待处理” |
| execute\_time | 执行时间 | DATETIME |  | 无 |
| policy\_id | 脱敏策略ID | INT | FOREIGN KEY REFERENCES MaskingPolicy(policy\_id), NOT NULL | 无 |
| data\_id | 脱敏数据ID | INT | FOREIGN KEY REFERENCES SensitiveData(data\_id), NOT NULL | 无 |

表 11操作日志表(OperationLog)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| log\_id | 操作日志ID | SERIAL | PK | 无 |
| action\_type | 操作类型 | VARCHAR(50) | NOT NULL | 无 |
| user\_id | 操作用户ID | INT | FOREIGN KEY REFERENCES User(user\_id), NOT NULL | 无 |
| action\_time | 操作时间 | DATETIME | NOT NULL | CURRENT\_TIMESTAMP |
| action\_detail | 操作详细内容 | VARCHAR(500) |  | 无 |

表 12 审计日志表(AuditLog)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **字段名** | **字段描述** | **字段类型%** | **约束条件** | **默认值** |
| audit\_id | 审计日志ID | SERIAL | PK | 无 |
| auditor\_id | 审计人员ID | INT | PK, REFERENCES User(user\_id), NOT NULL | 无 |
| audit\_time | 审核时间 | DATETIME | NOT NULL | CURRENT\_TIMESTAMP |
| report\_summary | 报告摘要 | VARCHAR(500) |  | 无 |
| log\_id | 操作日志ID | INT | FOREIGN KEY REFERENCES OperationLog(log\_id), NOT NULL | 无 |

# 系统详细设计

## 推荐模块设计

（主要产出的图为流程图,流程图规范：

https://www.yuque.com/haoren250/mx7t4d/sg7ftwl3pzcgxs1k?singleDoc# 《流程图》）

### 界面原型设计

### 接口设计

https://www.yuque.com/haoren250/mx7t4d/ietmdypgemk0aqhk?singleDoc# 《接口设计》

## 订单模块设计

### 界面原型设计

### 接口设计

# 系统实现

## 推荐模块实现

核心代码以及界面截图。

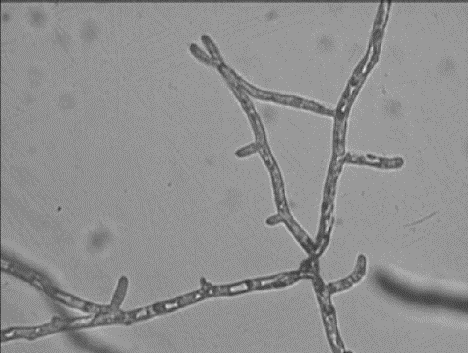


图1 霉菌菌株R1的菌丝（五号黑体）

Fig l Mycelium of mold strain R1 (五号Times New Roman)

## 订单模块实现

# 系统测试

提供测试用例

https://www.yuque.com/haoren250/mx7t4d/fgy64x6ilvzkttyy?singleDoc# 《测试用例》

# 总结与展望

总结一段(400字左右)

展望一段(200字左右)

# 参考文献

1. 作者.论文题名[J].期刊名，出版年，卷（期）：页码A-B. (五号宋体)
2. 蒋有绪，郭泉水，马娟，等.中国森林群落分类及其群落学特征[M].北京：科学出版社，1998：45-48.

不少于15篇文献（至少2篇英文文献、近三年文献不少于10篇）