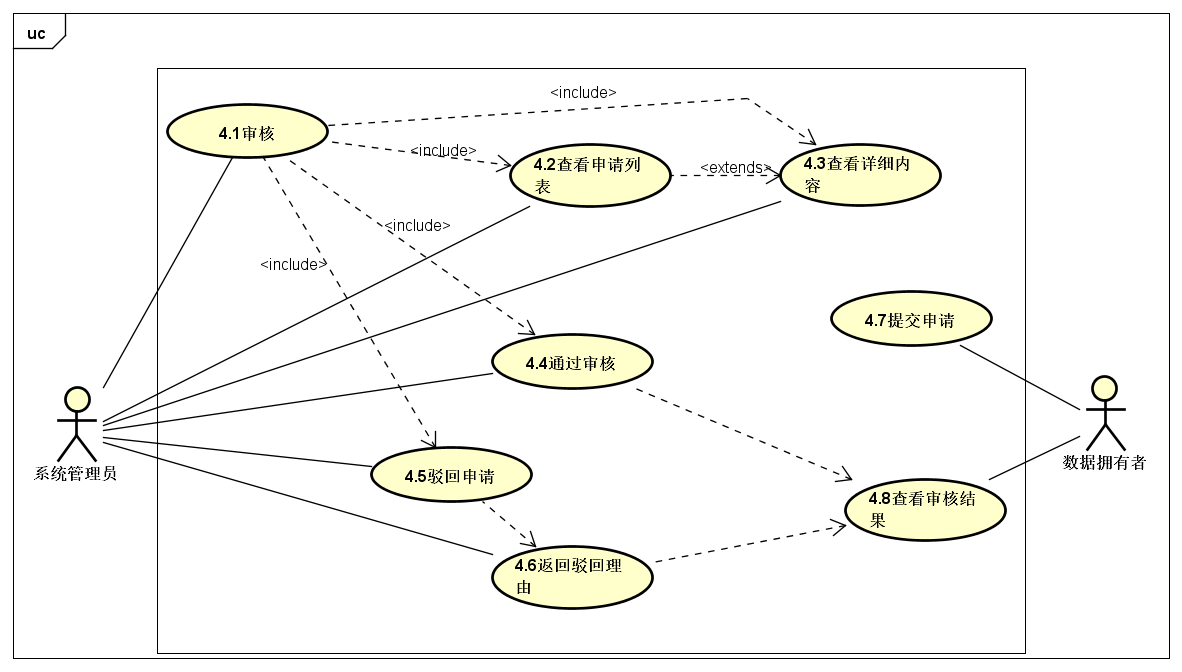
## 4.数据预处理

数据预处理主要有两个方面：1.上传预处理；2.标定完成后原始数据预处理。

上传预处理包含了数据分享用户上传数据的机器查重，人工审核，机器去冗余，数据有效化；标定完成后的原始数据的预处理主要是将标定完成的原始数据确定正确答案。

4.1 用例图



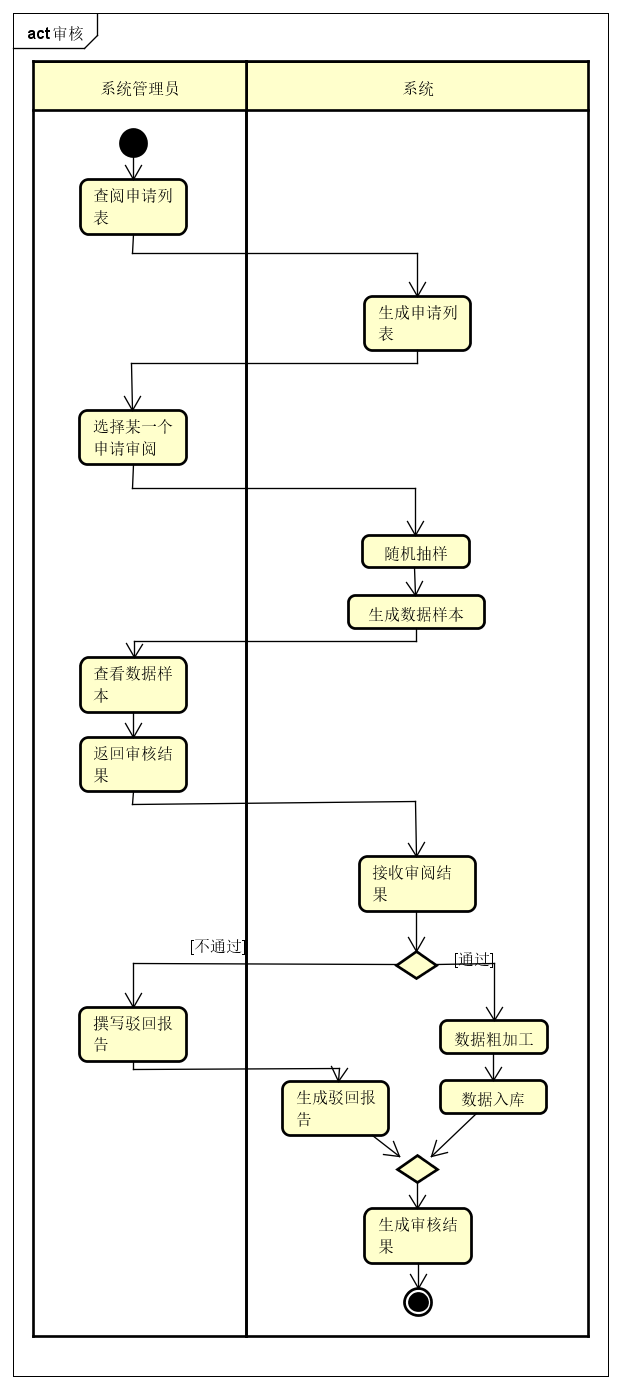
4.2 用例表

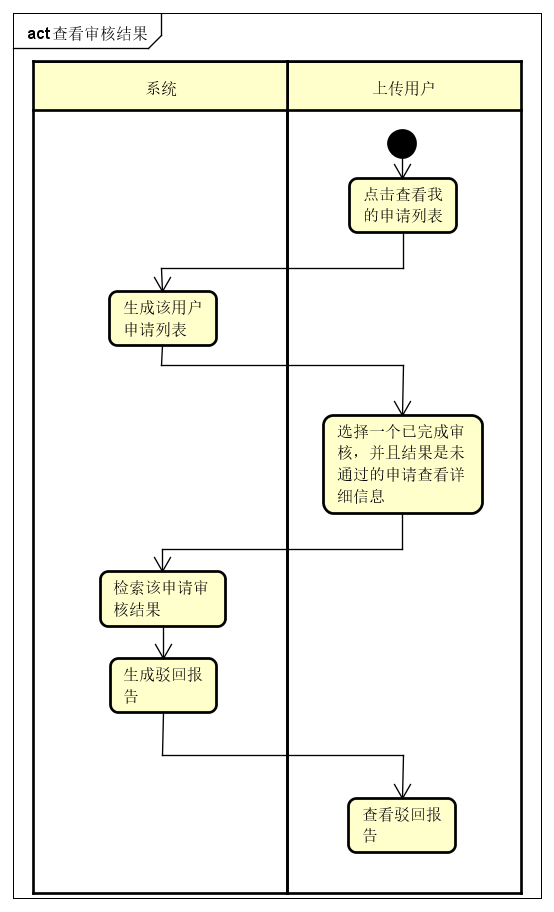
数据预处理部分涉及系统内部功能方面多，人机交互界面少，而且都逻辑简单。所以此处只举两个用例作为例子来展示。

|  |
| --- |
| 用例：审核 |
| ID：4.1 |
| 简单描述：系统管理员对上传数据进行审核 |
| 主参与者：系统管理员 |
| 副参与者：系统 |
| 前置条件：上传用户成功提交上传申请 |
| 主流：   1. 系统管理员以管理员身份登陆系统 2. 系统管理员查看申请列表 3. 系统管理员查看数据样本 4. 系统管理员返回审核结果 |
| 后置条件：系统管理员完成审核 |
| 附加流：审核结果为未通过时，系统管理员要返回驳回理由 |

|  |
| --- |
| 用例：查看审核结果 |
| ID：4.8 |
| 简单描述：上传用户查看审核状态 |
| 主参与者：上传用户 |
| 副参与者：系统 |
| 前置条件：上传用户成功提交过上传申请 |
| 主流：   1. 上传用户以用户身份登陆网页   2）上传用户查看申请列表 |
| 后置条件：上传用户完成查看审核状态 |
| 附加流：审核状态为未通过时，用户通过链接查看驳回报告 |

4.3 活动图





用户可以在登陆以后查看自己的申请列表。我们暂定申请列表每一个申请项里包含几个信息：

1.数据名称

2.数据类型

3.数据大小

4.上传日期

5.申请表

6.审核状态

审核状态是用户最关注的一项，审核状态有3个值：待审核，已通过，未通过。值为未通过的时候，用户可以通过链接查看驳回报告，以获知自己的数据未能通过审核的原因。