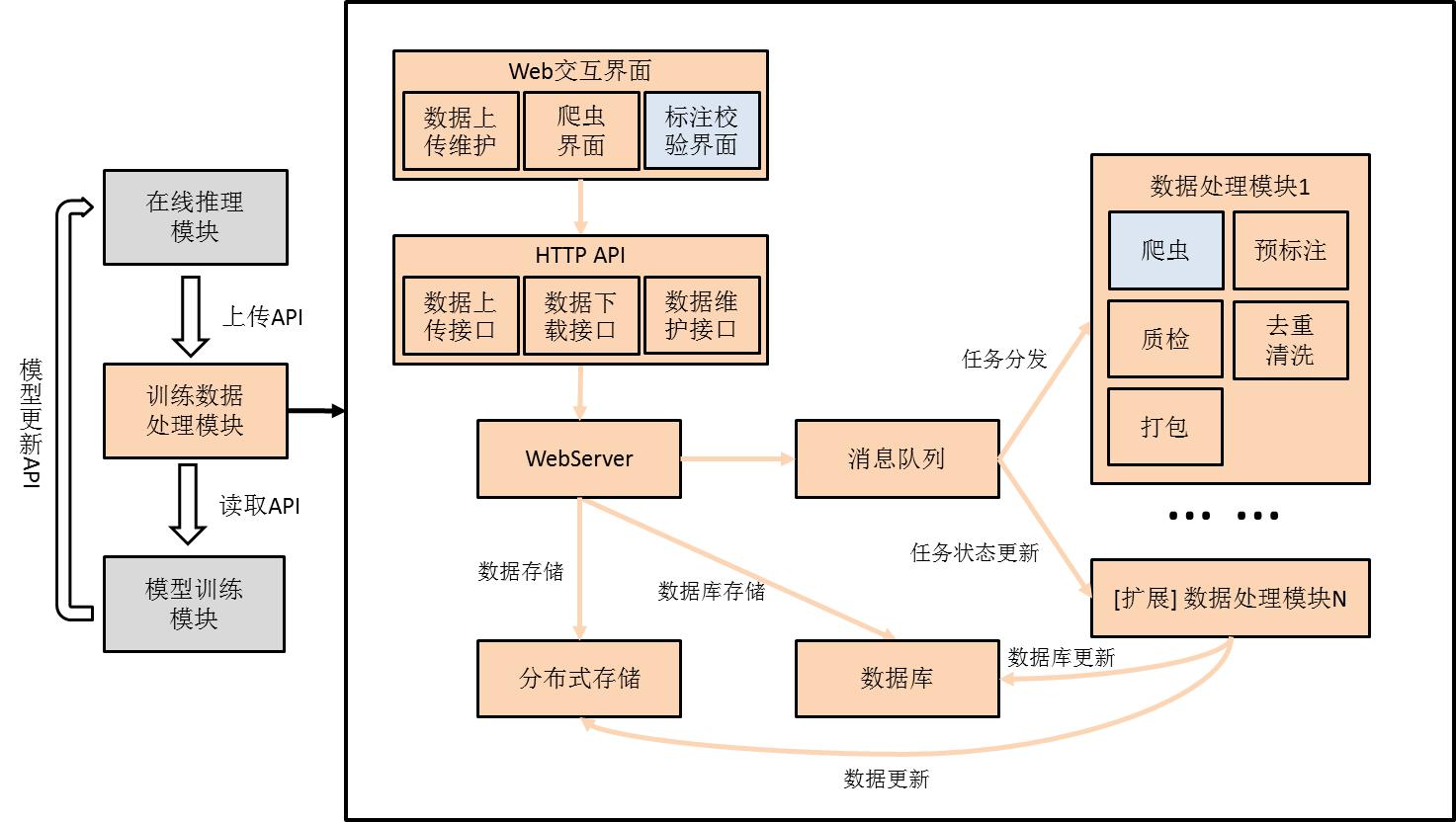
1. 训练数据处理模块



用户建立数据集，向数据集中上传图片，对该图片进行存储、标注、去重等处理；对外提供图片上传、下载接口。

1. 网站/接口：

a. 用户模块：~~管理员；~~普通用户（数据集和数据进行增删改）；标注用户（仅能标注图片）；

b. 数据集维护：增删改查；

c. 数据维护：增（上传、爬虫）、删、改（标注）、~~下载~~

2. 消息队列：

a. 将爬虫、预标注、质检等图像处理以任务形式分发至数据处理模块处理

3. 数据处理模块：

a. 进行图像处理，主要由研究院开发

b. 集群

4. 存储：

a. 存储图片（～1MB，10w～100w量级）及其标注文件(文本)，以及图片打包后文件

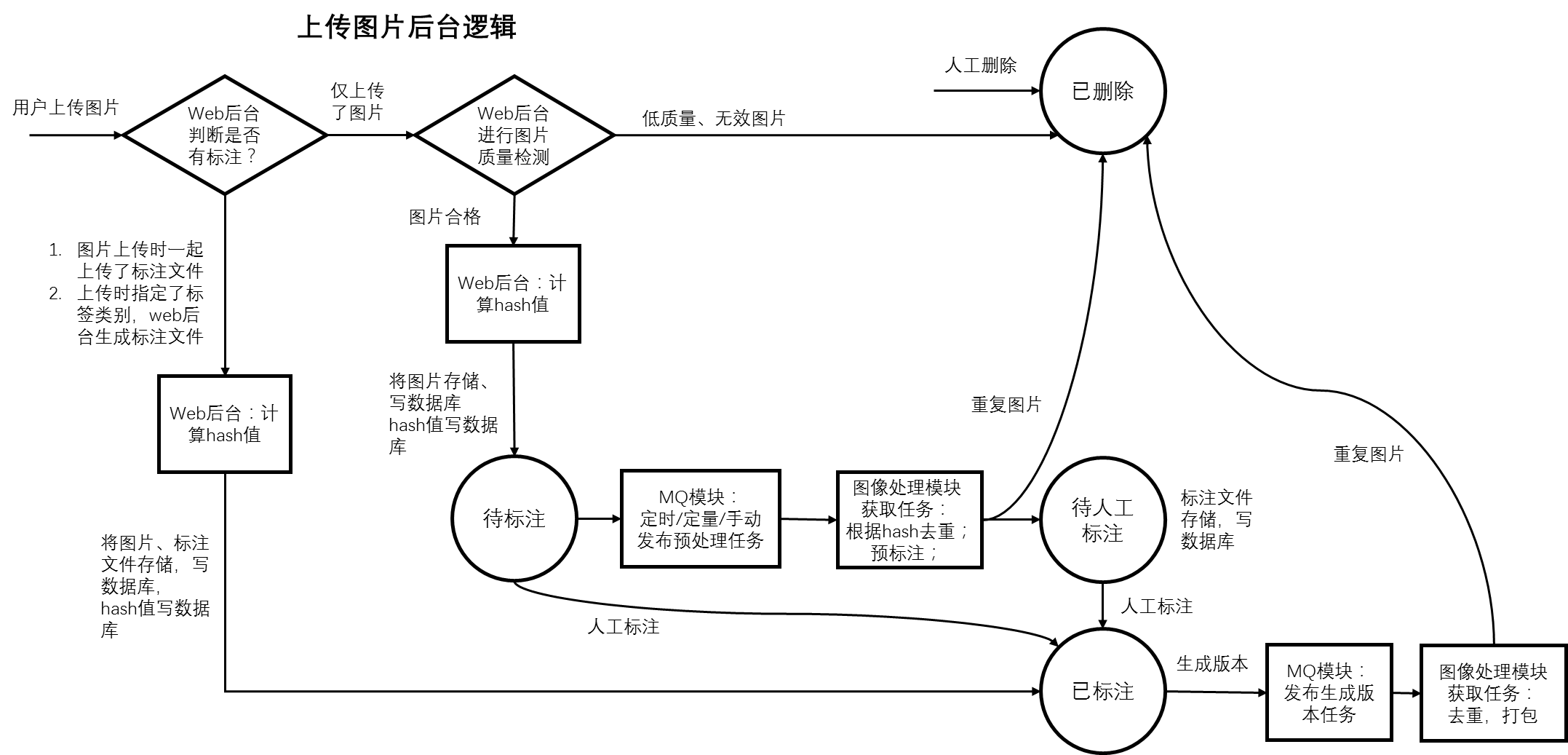
b. 分布式形式

5. 数据库：

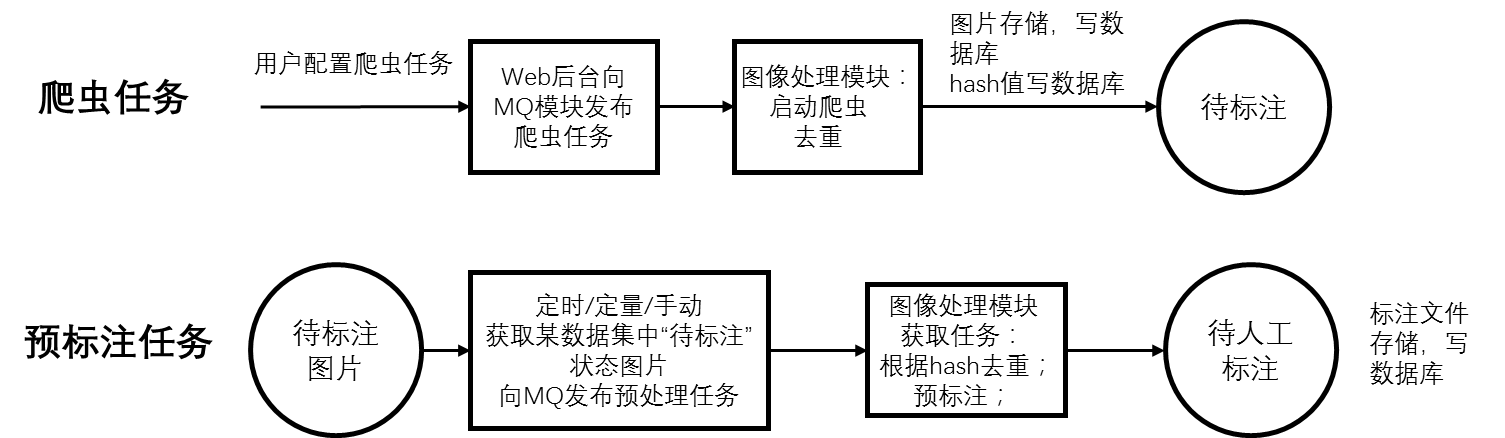
a. 存储数据集信息

b. 存储图片路径、标注路径等

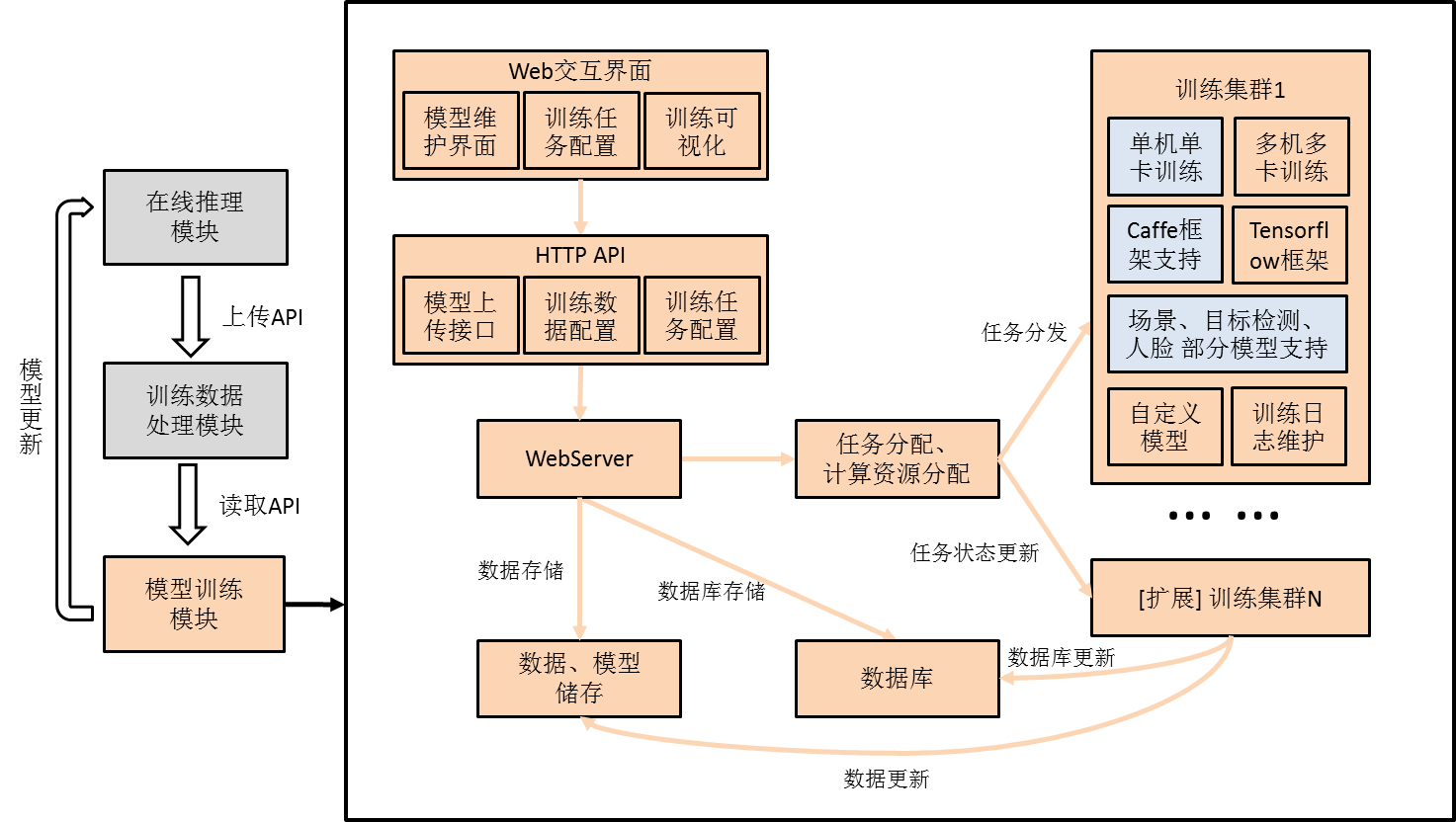
上传图片操作后台逻辑



MQ任务后台逻辑



二、模型训练模块



用户新建模型，指定使用的训练数据，配置训练参数。模型由训练集群训练，可视化训练过程、结果，提供模型下载；

1. 网站，接口：
2. 用户模块：管理员；普通用户；与训练数据处理模块同步
3. 模型配置：由用户上传模型配置文件
4. 训练数据配置：读取训练数据处理模块数据
5. 训练任务配置：任务增删查，任务参数配置
6. 任务分配、计算资源分配：
7. 主要由研究院完成
8. 训练集群：
9. 主要由研究院完成
10. 存储：
11. 用于储存训练数据（打包的），训练模型，日志
12. 分布式
13. 数据库：
14. 储存模型信息
15. 储存训练任务信息