

重庆邮电大学



知乎软件需求分析文档

课程名称： 技术文档编制（双语）

项目组号： 7

班级： 13002207 班

姓名： 朱烨彬、许世睿、唐欣轶

学号： 2022214283, 2022214309, 2022214306

指导老师： 蒋救

学期： 2024-2025 学年第 2 学期

2025 年 5 月 12 日

文档信息

文档标题	知乎软件需求分析文档
归档日期	2025. 5. 14
所有者	朱烨彬、许世睿、唐欣轶

修订历史

版本编号	版本日期	修订内容	备注/修订人
V0. 1		初始版本	
V0. 2	4. 24	1. 1	朱烨彬
V0. 3	4. 26	1. 2	唐欣轶
V0. 4	4. 27	1. 3	许世睿
V0. 5	5. 7	2. 1	朱烨彬、许世睿
V0. 6	5. 8	2. 2	唐欣轶
V0. 7	5. 10	3. 0	唐欣轶、朱烨彬
V0. 8	5. 12	3. 1	许世睿
V0. 9			
V1. 0			

目录

第 1 章 简介	1
1.1 目的	1
1.2 范围	1
1.3 定义、简写和缩略语	3
1.4 引用文件	4
第 2 章 总体描述	4
2.1 开发背景	4
2.2 系统目标	5
2.3 用户特点	5
2.4 系统功能结构图及其描述	7
第 3 章 具体需求	8
3.1 功能需求	8
3.1.1 内容生产	8
3.1.2 内容分发	8
3.1.3 社区互动	9
3.1.4 内容治理	9
3.1.5 商业化运营	10
3.1.6 用户成长	10
3.1.7 消息通知	11
3.1.8 数据管理	11
3.1.9 开放平台	12
3.2 性能需求	12
3.2.1 响应速度	12
3.2.2 并发能力	13
3.2.3 数据传输速度	13
3.2.4 数据处理速度	14
3.2.5 推荐算法性能	14
3.2.6 内容审核性能	15
3.2.7 搜索功能性能	15
3.3 数据需求	16
3.4 接口需求	19
3.4.1 用户认证接口	19
3.4.2 内容管理接口	20
3.4.3 互动管理接口	21
3.4.4 内容审核接口	22
3.4.5 推荐系统接口	22
3.4.6 推广管理接口	22
3.4.7 用户成长接口	23
3.4.8 数据统计接口	23
3.4.9 第三方接口	24
3.5 环境需求	24
3.5.1 硬件环境	24

3.5.2 软件环境	25
3.5.3 运行环境	26
第 4 章 用例模型	27
4.1 用例图	27
4.2 用例规约	27
4.3 用例活动图	32
第 5 章 验收准则	35
第 6 章 附录	35
6.1 大模型运用	35
6.1.1 识别模块的运用	35
6.1.2 大模型二次检测	35
6.1.3 语言润色	35
6.1.4 大模型的逆向分析进行图表解读	36
6.1.5 大模型应用的局限性	36
6.2 分工	36

第 1 章 简介

1.1 目的

本需求规格说明书旨在全面定义“知乎”知识分享社区的核心业务需求与系统功能边界，为以下目标提供基准：

- 1).指导开发团队构建符合用户预期的问答社区平台
- 2).建立可量化的系统性能验收标准（响应时间、并发承载等）
- 3).规范内容安全审核流程与用户行为管理体系
- 4).对齐产品、研发、测试三方对系统能力的共同认知

适用读者：

产品经理（需求决策者）、开发工程师（系统实现者）、质量保障团队（验收验证者）、运营管理人员（业务使用者）。

1.2 范围

知乎知识分享社区平台旨在构建高质量 UGC 内容生态,核心业务需求包括：

表 1.1 核心业务需求

功能模块	业务场景描述
1. 内容生产	支持多形式创作（问答/文章/视频/直播）、专业领域认证、草稿云同步与多端发布
2. 内容分发	基于用户画像的个性化推荐（热榜/关注流/圆桌）、知识图谱关联内容智能聚合
3. 社区互动	点赞/收藏/分享三级传播体系，专业领域圆桌讨论，实时盐值信用反馈机制

功能模块	业务场景描述
4. 内容治理	瓦力 AI 内容审核（7x24 小时违规检测）、用户举报分级处理、创作者信用分动态评估体系
5. 商业化运营	知+内容推广系统、付费咨询对接、品牌任务定制化投放及 ROI 数据分析看板
6. 用户成长	盐值成长体系（0-1000 分）、创作者等级系统（Lv1-Lv10）、专业领域影响力榜单
7. 消息通知	分级推送系统消息（违规提醒/收益结算）、个性化内容更新提醒（关注话题动态）
8. 数据管理	创作者数据中心（阅读量/互动率/收益）、社区健康度监控仪表盘（违规率/响应时效）
9. 开放平台	API 接口对接第三方应用（内容同步/数据分析）、开发者沙箱环境测试

系统用例图：

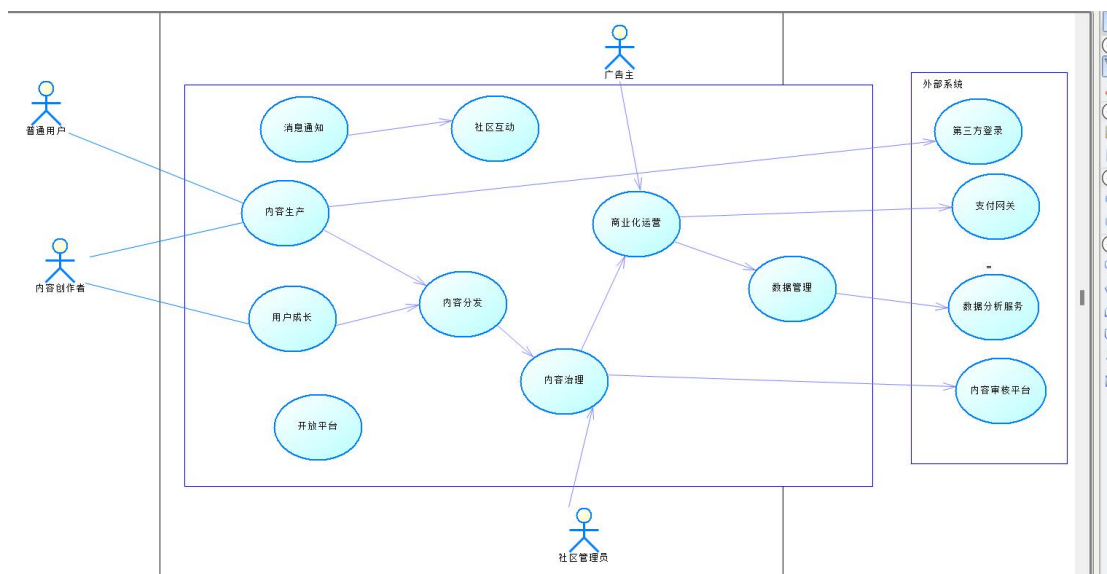


图 1.1 用例图

1.3 定义、简写和缩略语

表 1.2 术语

术语	定义
盐值体系	用户社区行为信用评分系统（0-1000 分），影响内容曝光权重和违规处理梯度
瓦力识别	知乎自研的 AI 内容审核引擎，涵盖文本/图片/视频的多模态违规检测
知+	商业化内容推广系统，支持精准投放效果追踪
LIVE	实时互动产品模块，包含直播问答、连麦辩论等场景
创作者中心	面向专业用户的内容管理后台，提供数据分析、收益结算、版权保护等工具

1.4 引用文件

表 1.3 引用文件

文档编号	文档名称
ZH-CM-2024	《知乎社区管理规范（2024 修订版）》
GB/T 35273-2020	《信息安全技术 个人信息安全规范》
ISO/IEC 25010	《系统和软件质量模型》
TR-ZH-API-003	《知乎开放平台接口规范 V3.2》

第 2 章 总体描述

2.1 开发背景

如今用户生成内容爆发式增长，知乎面临以下严峻挑战：

1) 信息过载导致优质内容淹没

用户日均提问量超 50 万，海量内容中高质量回答易被稀释，用户需投入更多时间筛选有效信息。

2) 社区氛围维护压力增大

日均举报量超 2 万+，需高效处理涉黄、暴恐、引战等违规内容，保障内容安全与社区理性讨论氛围。

3) 多端体验一致性需求

需适配 Web/Android/iOS/小程序等多终端，确保用户在不同场景下获得流畅体验。

4) 商业化与内容质量的平衡

知识付费、广告投放等商业化功能需融入社区生态，避免影响用户体验。

2.2 系统目标

表 2.1 系统目标

维度	指标要求	支撑技术/策略
响应速度	首屏加载≤1.2 秒	CDN 加速、客户端渲染优化（如 React Native 动态加载）
并发能力	支持 10 万	QPS 微服务架构（K8S 容器化）、Redis 集群缓存、数据库分片策略
内容安全	审核准确率≥99.9%	AI 模型（BERT+NLP）+人工复审、敏感词动态过滤、品牌词黑名单
推荐精度	CTR 提升 30%	深度学习推荐算法（DNN+强化学习）、用户画像实时更新、冷启动泛化策略
用户增长	创作者周活跃度提升 20%	流量曝光工具（如热榜加权）、盐值激励体系、创作数据分析中台

2.3 用户特点

知乎用户主要考虑四个群体用户，分别是新用户，创作者，机构用户，高知用户。

表 2.2 用户特点

角色	行为特征	关键需求
新用户	浏览为主，互动率<5%	个性化冷启动推荐（基于设备信息、初始搜索词生成临时兴趣标签）
创作者	周均发布≥3 篇长内容	流量分发分析（曝光→阅读→点赞漏斗）、内容版权保护（水印+反爬机制）
机构用户	需品牌词过滤/营销内容托管	定制化运营工具（舆情监控、竞品关键词屏蔽、广告投放 ROI 分析）

角色	行为特征	关键需求
----	------	------

高知用户	活跃于科技、教育等专业领域	深度内容消费（专栏/电子书）、行业圆桌讨论、学术资源整合
------	---------------	------------------------------

1) 新用户

以浏览为主，互动率低（小于 5%）。关键需求是个性化冷启动推荐，需要通过其注册信息及初始浏览行为，推送具有广泛吸引力和代表性的热门话题、优质回答和基础知识点等内容，帮助其快速融入知乎社区，发现个性化内容，降低获取有价值知识的门槛，提升对平台的认同感和归属感。

2) 创作者

平均每周发布至少 3 篇长内容，包括回答问题、撰写文章、开设专栏等。关键需求是流量曝光数据分析，期望平台提供全面、细致的流量曝光数据分析工具，如文章的阅读量、点赞数、评论数、分享数、粉丝增长趋势等数据的可视化展示和深度分析报告，以便了解作品的受欢迎程度和受众特征，有针对性地调整创作方向和策略，持续产出高质量作品，提升自身创作积极性和忠诚度。

3) 机构用户

代表企业、品牌、媒体、学术机构等官方身份进行知识传播和品牌推广。关键需求一是品牌词过滤功能，确保发布内容时突出自身品牌相关关键词和信息，维护品牌独特性和专业性形象；二是定制化运营工具，如批量上传文章、定时发布内容、管理营销活动、监测品牌形象等工具，以提高工作效率和管理效果，更好地实现品牌传播和商业价值最大化目标，同时保证内容符合知乎平台社区规范和用户体验要求。

4) 高知用户

通常具有较高的专业知识水平和学历背景，对内容的质量和深度有较高要求。他们关注的是一些专业领域的前沿知识、深度分析和学术讨论等内容，注重从知乎平台上获取能够拓展其专业知识和视野的优质内容，以满足自身在学习、研究或职业发展等方面的需求，同时也会积极参与到专业话题的讨论和交流中，分享自己的见解和经验，对平台知识的深度和专业性建设有一定贡献。

2.4 系统功能结构图及其描述

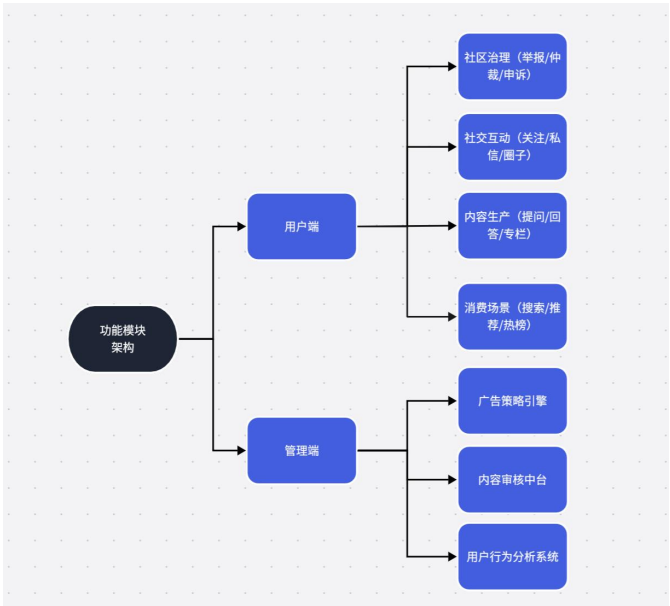


图 2.1 系统功能结构图

该系统功能结构图展示了知乎平台的主要功能模块架构，分为用户端和管理端两大核心部分，具体如下：

1) 用户端

① 内容生产：支持用户进行提问、回答以及创建专栏，涵盖内容创作的多种形式，为知识分享和交流提供基础平台。

② 社交互动：包括关注、私信和圈子功能，方便用户之间建立联系、进行私密交流以及参与不同主题的社区圈子互动，增强用户粘性和社区氛围。

③ 消费场景：提供搜索、推荐和热榜功能，帮助用户快速找到感兴趣的内容，发现热门话题和优质推荐，提升内容浏览体验。

④ 社区治理：设有举报、仲裁和申诉功能，保障社区内容的健康性和合法性，为用户提供更安全、有序的交流环境。

2) 管理端

① 广告策略引擎：负责广告投放计划的制定、执行与效果评估，旨在通过精准的广告投放策略提升广告收益，同时确保广告与用户体验的平衡。

② 内容审核中台：对用户生成的内容进行审核，通过制定审核规则、分配审核任务和处理审核结果等操作，确保内容安全合规，符合平台规范和法律法规

要求。

③ 用户行为分析系统：收集、分析用户行为数据，为平台运营决策提供数据支持，助力优化用户体验和提升平台的运营效率。

第 3 章 具体需求

3.1 功能需求

3.1.1 内容生产

1) 多形式创作

支持用户创建问答、文章、视频、直播等多种形式的內容。普通用户可进行基础内容创作，专业用户能创作专业内容，且专业内容需关联权威参考文献。创作时，用户能输入正文、添加话题标签、上传图片或视频等多媒体附件，提交至内容分发系统。若上传附件大小超限，系统提示用户压缩重传。

2) 专业领域认证

为专业领域用户提供账号提供认证服务，提升内容专业性和可信度，认证成功后，用户创作的专业内容在相应领域有更高权重和曝光机会，且专业领域内容需盐值 ≥ 500 才能评论。

3) 草稿云同步与多端发布

用户编辑内容时，草稿自动云同步，确保数据安全。用户可在不同设备上登录账号，继续编辑或发布草稿，实现多端无缝创作体验。

3.1.2 内容分发

1) 个性化推荐

基于用户画像（兴趣标签、浏览历史等），首页生成个性化内容流。推荐算法混合排序（60%兴趣+30%热点+10%探索），新用户冷启动阶段采用热度优先策略。若数据源异常，降级返回全站热榜；算法超时，启用缓存结果。同时，可为创作者等级 lv5+ 的内容加权 20%，提升优质内容曝光。

2) 多渠道分发

内容发布后，通过站内消息、邮件、短信等渠道，将内容推荐给可能感兴趣的用戶。用戶可在个人中心设置推荐内容的接收方式和频率，如每天接收 3 次邮件推荐、开启或关闭短信通知等。

3) 知识图谱关联

利用知识图谱技术，智能聚合关联内容。当用戶浏览某内容时，系统自动展示相关知识领域的其他优质内容，拓宽用戶视野，加深对知识的理解。

3.1.3 社区互动

1) 三级传播体系

构建点赞、收藏、分享的三级传播体系。用戶可对内容进行点赞、收藏、分享至站内或站外平台。敏感内容禁止分享至站外，若检测到刷量行为，临时封禁互动权限。分享时，用戶可添加自定义分享语，个性化表达对内容的看法。

2) 专业领域圆桌讨论

开设专业领域圆桌，用戶可发起或参与圆桌讨论，交流见解。圆桌讨论有明确主题和规则，主持人负责引导讨论方向，邀请专家参与，提升讨论质量和深度。用戶可关注圆桌话题，及时获取最新讨论动态。

3) 盐值信用反馈

根据用戶互动行为（点赞优质内容、分享有价值信息等）和违规行为（发布不良信息、恶意刷量等），动态计算盐值。盐值影响内容曝光权重和用戶权益，如盐值高可获得更多创作权限、专属标识等；盐值过低则限制部分功能。

3.1.4 内容治理

1) 瓦力 AI 审核

瓦力 AI 系统 7×24 小时对内容进行初筛，置信度≥90%自动裁决。初筛通过的内容正常展示；未通过的进入人工复审。若审核超时（>5 分钟），自动放行并标记；系统过载时，触发熔断机制，优先保障核心功能审核。

2) 用戶举报处理

用户可举报违规内容，系统接收举报后，快速定位内容并进行审核。根据举报内容的严重程度和类型，确定审核优先级。如涉及违法信息的举报，优先处理。同时，建立举报反馈机制，向举报用户反馈处理结果，对有效举报给予奖励。

3) 创作者信用评估

建立创作者信用动态评估体系，根据创作者的内容质量（如文章原创性、准确性等）、互动情况（如回复评论的及时性和质量等）、违规记录等多维度数据，定期评估信用分。信用分影响创作者的推荐权重、权益和收益，激励创作者持续产出优质内容。

3.1.5 商业化运营

1) 知 + 内容推广

广告主可创建推广任务，设置目标人群（如年龄、地域、兴趣等）、预算。系统匹配目标创作者，创作者创作符合要求的内容进行推广。推广内容需通过合规性预审，按 **cpm/cpc** 计费投放。若点击率异常，暂停投放并核查；余额不足，降权展示。品牌任务需签署电子合同，效果达标自动续投。

2) 付费咨询

创作者可开通付费咨询服务，设定咨询价格、服务范围和时间。用户发起付费咨询，系统匹配合适创作者。咨询过程中的聊天记录加密存储，保障隐私安全。咨询结束后，用户对创作者进行评价，评价结果影响创作者的付费咨询推荐权重和收益。

3) 品牌任务定制投放

品牌方可定制专属任务，明确任务要求（如产品推广、活动宣传等）、投放周期和预算。系统筛选符合要求的内容创作者，进行个性化投放。根据投放效果（如曝光量、互动量等），调整投放策略，提高品牌推广效果。定期生成投放效果报告，为品牌方提供数据支持。

3.1.6 用户成长

1) 盐值成长体系

构建盐值成长体系（0 - 1000 分），监测用户日常行为（创作、互动、遵守社区规范等），动态计算盐值增减。盐值达到一定标准，用户等级提升(lv1 - lv10)，解锁新权益。连续 30 天盐值 ≥ 800 授予“优秀答主”标识；重大违规直接降至lv0。若数据异常，冻结积分变动；申诉成功，补偿历史积分。

2) 创作者等级

创作者等级系统（lv1 - lv10），用户通过持续创作、提升内容质量和互动效果，累计经验升级。高等级创作者享受更多创作工具、流量扶持和收益分成。创作者可查看升级所需经验和当前经验值，明确成长方向。

3) 专业领域影响力榜单

定期发布专业领域影响力榜单，根据用户在专业领域的创作数量、质量、互动情况和粉丝数量等数据进行排名。上榜用户获得更多曝光机会和粉丝关注，激励创作者深耕专业领域，提升专业素养。

3.1.7 消息通知

1) 违规提醒

用户内容违规时，系统及时推送违规提醒消息，告知违规原因和处理结果。提醒通过站内信、短信或邮件发送，用户需在规定时间内处理。若用户对处理结果有异议，可在线申诉，系统重新审核并反馈结果。

2) 收益结算通知

对于创作者的收益（如付费咨询、内容推广等），系统每月定期推送收益结算通知。通知详细列出收益项目、金额和结算时间，确保创作者了解收益情况。创作者可查看历史收益记录，对账和分析收益来源。

3) 个性化内容更新提醒

根据用户关注的话题和创作者，系统推送个性化内容更新提醒。提醒用户关注话题有新内容发布或关注的创作者有新动态，提升用户活跃度。用户可在设置中自定义提醒内容和频率，如只接收关注话题的精华内容提醒。

3.1.8 数据管理

1) 创作者数据中心

为创作者提供数据中心，展示内容的阅读量、互动率、粉丝增长等数据。数据以图表形式呈现，创作者直观了解内容表现。创作者可按时间、内容类型等维度筛选和分析数据，优化创作策略。

2) 社区健康度监控

建立社区健康度监控仪表盘，展示违规率、内容审核时效、用户活跃度等指标。运营人员实时监控社区健康状况，指标异常时及时调整运营策略。定期生成社区健康度报告，为平台长期发展提供数据支持。

3.1.9 开放平台

1) API 接口对接

提供标准化 API 接口，允许第三方应用接入。第三方应用可通过接口获取知乎公开内容（如热门文章、话题列表等）、提交优质内容至知乎。知乎对第三方应用的接口调用进行流量控制和安全认证，防止恶意调用。

2) 开发者沙箱环境

为开发者提供沙箱环境，用于测试和调试 API 接口。沙箱环境模拟知乎真实环境，开发者在不影响生产环境的情况下进行开发。沙箱环境提供测试数据和工具，帮助开发者快速定位和解决问题。

3.2 性能需求

3.2.1 响应速度

1) 首屏加载

系统首屏加载时间应小于等于 1.2 秒。这是为了确保用户在打开知乎时能够快速获取到首页的推荐内容，避免因加载时间过长而导致用户流失。例如，当用户从手机应用图标点击进入知乎时，能够在短短 1.2 秒内看到首页展示的各类推荐内容，从而提高用户的使用体验和满意度。

2) 操作响应

用户进行内容发布、点赞、评论等各类操作后，系统应在 0.5 秒内给出响应。这意味着当用户点击“点赞”按钮时，页面能在 0.5 秒内完成点赞操作并及时更新点赞数等相关信息，让用户感受到操作的即时性和流畅性。

3) 内容详情页加载

内容详情页加载时间应小于等于 1 秒。当用户点击某个问题或文章的标题进入详情页面时，在 1 秒内完成页面的加载和展示，包括文字内容、图片、视频等各类元素，确保用户能够快速获取到详细信息，减少等待时间。

4) 搜索结果页面加载

搜索结果页面加载时间应小于等于 1.5 秒。考虑到用户在搜索时希望快速获取到相关的搜索结果，因此在用户输入关键词并提交搜索请求后，系统应在 1.5 秒内返回准确且全面的搜索结果页面，提升用户搜索体验。

3.2.2 并发能力

系统应能够支持同时在线用户数达到 10 万以上，并发请求量达到每秒 1000 次以上，且在高并发情况下，系统资源利用率（如 CPU、内存）不应超过 80%。以确保在高流量的访问高峰时段，如知乎的早晚高峰期，数万用户同时在线浏览、发布内容、互动等，系统依然可以稳定运行。系统能够在每秒处理 1000 次以上的并发请求，例如大量的用户同时提交评论、点赞等操作请求，而服务器的 CPU 和内存资源不会过度消耗，始终有 20% 以上的剩余资源来应对突发的流量增长，从而保障系统的可靠性。

3.2.3 数据传输速度

1) 多媒体内容上传

在用户网络条件允许的情况下，上传图片、视频等多媒体内容的速度不低于 1MB/秒。这可以让用户在发布包含图片或视频的内容时，如分享生活中的精彩

瞬间或专业知识的视频讲解，能够快速地将多媒体文件上传至知乎服务器，减少用户的等待时间，提高内容发布的效率。

2) 数据传输延迟

数据传输过程中的延迟应小于等于 0.2 秒。无论是用户请求数据还是服务器返回数据，低延迟的数据传输可以确保用户与知乎平台之间的交互更加实时和顺畅。比如，当用户在观看直播内容时，低延迟的数据传输能够保证直播画面和声音的同步，让用户的观看体验更加连贯。

3.2.4 数据处理速度

对于每日产生的大量数据，如用户行为数据、内容数据、互动数据等，系统应在 1 小时内完成数据的采集、清洗、存储和分析等处理过程。随着知乎平台的不断发展和用户数量的增加，每天都会产生海量的数据。及时的数据处理能够为个性化推荐、内容审核等功能提供实时的数据支持。例如，通过对用户行为数据的快速分析，系统可以在短时间内更新用户的兴趣标签，从而为其提供更精准的个性化推荐内容。

3.2.5 推荐算法性能

1) 推荐结果生成速度

个性化推荐系统的推荐结果应在用户打开首页或进入相关内容页面时，在 0.3 秒内生成并展示给用户。这意味着当用户每次打开知乎首页或者进入某个内容页面时，系统能够在极短的时间内根据用户的兴趣和行为，快速计算并推送出个性化的推荐内容列表，提高用户发现感兴趣内容的效率。

2) 推荐准确率

推荐结果的准确率应达到 80% 以上，且随着用户行为数据的积累和算法的优化，推荐准确率应不断提高。例如，通过对大量用户数据的学习和分析，推荐算法能够精准地把握用户的兴趣点，为用户推荐他们真正感兴趣的内容，如推荐

给科技爱好者最新的科技资讯和产品评测文章，从而提高用户对推荐内容的满意度和点击率。

3.2.6 内容审核性能

1) AI 初筛速度

瓦力 AI 内容审核系统的审核速度应满足实时审核的需求，对于新发布的内容，应在 0.1 秒内完成初筛。这可以迅速过滤掉明显违规的内容，如包含色情、暴力等违禁词的内容，保障社区内容的健康和安全，让优质内容能够更快地展示给用户。

2) 人工复审速度

对于低置信度内容的人工复审，应在 5 分钟内完成审核并给出结果。当 AI 初筛无法确定内容是否违规时，人工审核团队需要在规定时间内对内容进行仔细审查，并做出准确的判断。这既保证了审核的准确性，又不会让内容长时间处于待审核状态，影响用户的发布体验和内容的及时性。

3.2.7 搜索功能性能

1) 搜索结果返回速度

用户在使用搜索功能时，搜索结果应在 0.5 秒内返回。快速的搜索结果返回能够让用户在短时间内获取到他们想要查找的内容，提高搜索效率。例如，当用户搜索某个专业知识问题时，能够在 0.5 秒内看到相关的高质量回答和文章，帮助他们快速解决问题。

2) 搜索结果相关性和准确性

搜索结果的相关性和准确性应达到 85% 以上。这要求知乎的搜索算法能够精准理解用户搜索意图，返回与用户需求高度相关的搜索结果。例如，当用户搜索“人工智能的应用”时，搜索结果页面展示的大部分内容都与该主题紧密相关，而不是一些无关或关联度低的信息，从而提升用户的搜索满意度。

3.3 数据需求

知乎软件系统中的数据需求主要包括用户信息、内容信息、互动信息、审核信息、推广信息、成长信息等。以下是各主要数据元素/实体、属性及其相互关系的详细描述，并绘制了相应的 ER 图：

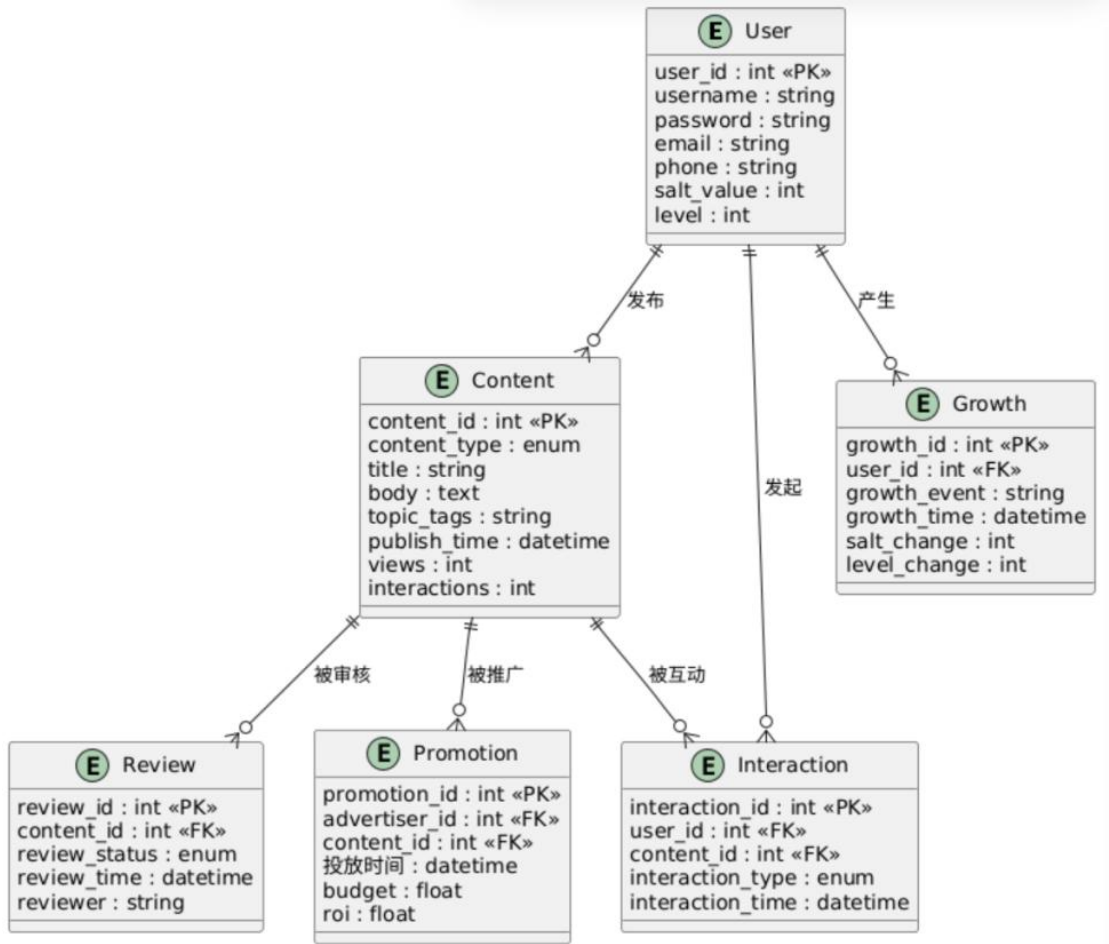


图 3.1 ER 图

数据元素/实体及其属性

用户（User）

用户 ID（user_id）：唯一标识用户的编号，整数类型，单位无，精确度为整数，范围为大于 0 的正整数，格式为自增整数。

用户名（username）：用户注册时选择的名称，字符串类型，单位无，精确度为无，范围为 2 - 20 个字符，格式为只能包含字母、数字、下划线。

密码（password）：用户设置的登录密码，字符串类型，单位无，精确度为无，范围为 8 - 20 个字符，格式为至少包含字母和数字。

邮箱（email）：用户的电子邮箱地址，字符串类型，单位如无，精确度为无，范围为符合邮箱格式，格式为如 xxx@xxx.com。

手机号（phone）：用户的手机号码，字符串类型，单位如无，精确度为 11 位，范围为 11 位手机号，格式为以 1 开头的 11 位数字。

盐值（salt_value）：用户社区行为信用评分，整数类型，单位为分，精确度为整数，范围为 0 - 1000 分，格式为整数。

等级（level）：用户在社区的成长等级，整数类型，单位如无，精确度为整数，范围为 1 - 10 级，格式为整数。

内容（Content）

内容 ID（content_id）：唯一标识内容的编号，整数类型，单位如无，精确度为整数，范围为大于 0 的正整数，格式为自增整数。

内容类型（content_type）：内容的类型，如问答、文章、视频、直播等，枚举类型，单位如无，精确度为无，范围为问答、文章、视频、直播，格式为枚举值。

标题（title）：内容的标题，字符串类型，单位如无，精确度为无，范围为 1 - 100 个字符，格式为字符串。

正文（body）：内容的主体部分，文本类型，单位如无，精确度为无，范围为 0 - 65535 个字符，格式为文本。

话题标签（topic_tags）：内容所属的话题标签，字符串类型，单位如无，精确度为无，范围为 0 - 10 个标签，格式为以逗号分隔的字符串。

发布时间（publish_time）：内容发布的日期和时间，日期时间类型，单位为秒，精确度为秒，范围为系统时间，格式为 yyyy-MM-dd HH:mm:ss。

浏览量（views）：内容的浏览次数，整数类型，单位为次，精确度为整数，范围为大于等于 0 的整数，格式为整数。

互动量（interactions）：内容的互动次数，包括点赞、收藏、分享等，整数类型，单位为次，精确度为整数，范围为大于等于 0 的整数，格式为整数。

互动（Interaction）

互动 ID (interaction_id) : 唯一标识互动的编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为自增整数。

用户 ID (user_id) : 参与互动的用户编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为整数。

内容 ID (content_id) : 被互动的内容编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为整数。

互动类型 (interaction_type) : 互动的类型, 如点赞、收藏、分享等, 枚举类型, 单位无, 精确度无, 范围为点赞、收藏、分享, 格式为枚举值。

互动时间 (interaction_time) : 互动发生的日期和时间, 日期时间类型, 单位为秒, 精确度为秒, 范围为系统时间, 格式为 yyyy-MM-dd HH:mm:ss。

审核 (Review)

审核 ID (review_id) : 唯一标识审核的编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为自增整数。

内容 ID (content_id) : 被审核的内容编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为整数。

审核状态 (review_status) : 审核的结果, 如通过、限流、删除等, 枚举类型, 单位无, 精确度无, 范围为通过、限流、删除, 格式为枚举值。

审核时间 (review_time) : 审核完成的日期和时间, 日期时间类型, 单位为秒, 精确度为秒, 范围为系统时间, 格式为 yyyy-MM-dd HH:mm:ss。

审核人 (reviewer) : 进行审核的管理员或系统名称, 字符串类型, 单位无, 精确度无, 范围为 1 - 50 个字符, 格式为字符串。

推广 (Promotion)

推广 ID (promotion_id) : 唯一标识推广的编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为自增整数。

广告主 ID (advertiser_id) : 推广所属的广告主编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为整数。

内容 ID (content_id) : 被推广的内容编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为整数。

投放时间 (投放时间) : 推广开始和结束的日期和时间, 日期时间类型, 单位为秒, 精确度为秒, 范围为系统时间, 格式为 yyyy-MM-dd HH:mm:ss。

预算 (budget) : 推广的预算金额, 浮点数类型, 单位为元, 精确度为小数点后两位, 范围为大于 0 的正数, 格式为浮点数。

ROI (roi) : 推广的投资回报率, 浮点数类型, 单位为百分比, 精确度为小数点后两位, 范围为大于等于 0 的正数, 格式为浮点数。

成长 (Growth)

成长 ID (growth_id) : 唯一标识成长记录的编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为自增整数。

用户 ID (user_id) : 用户的编号, 整数类型, 单位无, 精确度为整数, 范围为大于 0 的正整数, 格式为整数。

成长事件 (growth_event) : 用户成长的事件, 如盐值增加、等级提升等, 字符串类型, 单位无, 精确度无, 范围为 1 - 100 个字符, 格式为字符串。

成长时间 (growth_time) : 成长事件发生的日期和时间, 日期时间类型, 单位为秒, 精确度为秒, 范围为系统时间, 格式为 yyyy-MM-dd HH:mm:ss。

盐值变化 (salt_change) : 盐值的变化量, 整数类型, 单位为分, 精确度为整数, 范围为大于等于 0 的整数, 格式为整数。

等级变化 (level_change) : 等级的变化量, 整数类型, 单位为级, 精确度为整数, 范围为大于等于 0 的整数, 格式为整数。

3.4 接口需求

3.4.1 用户认证接口

1) 登录接口

用户通过提交用户名和密码进行登录, 系统验证成功后返回认证令牌和用户相关信息。例如, 用户在登录页面输入用户名和密码, 点击登录按钮后, 前端发

送登录请求，后端验证通过后返回包含 `token` 等信息的响应，前端将 `token` 存储起来，用于后续请求的身份验证。

2) 注册接口

新用户通过提交用户名、密码、邮箱和手机号进行注册，系统验证信息合法性和唯一性后完成注册。比如，用户在注册页面填写完整信息后，点击注册按钮，前端将数据发送到注册接口，后端验证通过后返回注册成功状态和新用户的 `user_id`。

3) 忘记密码接口

用户通过提供用户名和邮箱来重置密码，系统验证用户身份后允许设置新密码。例如，用户点击忘记密码链接，进入重置密码页面，输入用户名、邮箱和新密码，点击提交后，系统通过邮箱验证等方式确保是用户本人操作，然后返回密码重置成功状态。

3.4.2 内容管理接口

1) 内容发布接口

用户通过此接口发布不同类型的内容，提交相应的信息后，系统审核（包括自动审核和人工审核）通过后返回内容的唯一标识和发布成功状态。例如，创作者在内容编辑页面填写标题、正文、添加话题标签和多媒体附件后，点击发布按钮，前端将数据发送到此接口，系统进行审核，审核通过后返回内容的 `content_id` 和发布成功状态。

2) 内容查询接口

用户可以通过关键词、内容类型、话题标签等多种条件查询内容，系统根据条件返回匹配的内容列表。例如，在搜索页面，用户输入关键词，选择内容类型和话题标签等条件，点击搜索按钮后，系统调用此接口，返回包含符合条件内容的列表，展示在搜索结果页面。

3) 内容修改接口

创作者可以修改自己发布的内容，通过提交新的内容信息，系统验证用户身份和权限后进行修改。例如，创作者在内容编辑页面点击“编辑”按钮，修改标

题、正文等信息后，点击“保存”按钮，前端将修改后的数据发送到此接口，系统验证通过后返回修改成功状态。

4) 内容删除接口

创作者或管理员可以删除内容，系统验证用户身份和权限后删除对应的内容。例如，创作者在自己的内容管理页面点击“删除”按钮，系统弹出确认框，用户确认后，前端发送删除请求到此接口，系统验证通过后返回删除成功状态。

3.4.3 互动管理接口

1) 点赞接口

用户通过此接口对内容进行点赞操作，系统记录点赞行为并更新点赞总数。例如，用户在内容详情页面点击“点赞”按钮，前端发送点赞请求到此接口，系统验证用户身份和权限后返回点赞成功状态和更新后的点赞总数。

2) 收藏接口

用户通过此接口对内容进行收藏操作，系统记录收藏行为。例如，用户在内容详情页面点击“收藏”按钮，前端发送收藏请求到此接口，系统验证通过后返回收藏成功状态。

3) 分享接口

用户通过此接口将内容分享到不同的平台，系统记录分享行为。例如，用户在内容详情页面点击“分享”按钮，选择分享平台并填写自定义分享语，点击“确定”后，前端发送分享请求到此接口，系统验证通过后返回分享成功状态。

4) 评论接口

用户通过此接口对内容进行评论，系统记录评论内容并返回评论成功状态和评论的唯一标识。例如，用户在内容详情页面的评论区输入评论内容，点击“提交”按钮后，前端发送评论请求到此接口，系统验证通过后返回评论成功状态和 `comment_id`。

5) 互动数据查询接口

通过此接口查询内容的互动数据，包括点赞数、收藏数、分享数和评论数。例如，在内容详情页面加载时，系统调用此接口获取内容的互动数据，并展示在页面上，让用户了解内容的受欢迎程度。

3.4.4 内容审核接口

1) 提交内容审核接口

当有新内容发布时，系统通过此接口将内容提交给审核系统进行初筛。例如，创作者发布新内容后，系统自动调用此接口将内容发送给瓦力 AI 审核系统，返回审核任务 id 和提交成功状态。

2) 获取审核结果接口

通过审核任务 id 查询审核结果，了解内容是否通过审核以及审核的详细信息。例如，审核系统完成审核后，内容管理系统调用此接口获取审核结果，根据结果对内容进行相应的处理，如通过、限流或删除。

3) 查询审核记录接口

管理员可以通过此接口查询历史审核记录，方便进行数据统计和分析。例如，在审核管理页面，管理员可以输入条件查询审核记录，系统调用此接口返回符合条件的审核记录列表，展示在页面上。

3.4.5 推荐系统接口

个性化推荐接口：系统可以根据用户的需求和行为动态调整推荐策略，以提高推荐的准确性和满意度。例如，用户在设置页面选择推荐策略为“基于热门内容推荐”，系统调用此接口调整推荐策略，并返回调整成功状态。

3.4.6 推广管理接口

1) 创建推广任务接口

广告主通过此接口创建推广任务，系统验证广告主信息和内容合规性后返回推广任务 id 和创建成功状态。例如，广告主在推广管理页面填写推广任务信息，点击“创建”按钮后，前端发送请求到此接口，系统验证通过后返回 promotion_id 和创建成功状态。

2) 查询推广数据接口

通过此接口查询推广任务的执行情况和效果数据。例如，广告主在推广管理页面点击“查看数据”按钮，系统调用此接口获取推广数据，并展示在页面上，让广告主了解推广效果。

3) 统计推广效果接口

广告主可以通过此接口统计和分析一段时间内推广活动的整体效果。例如，广告主在推广分析页面选择时间范围，点击“统计”按钮后，系统调用此接口返回统计结果，展示在页面上，帮助广告主评估推广策略的有效性。

3.4.7 用户成长接口

1) 盐值查询接口

通过此接口查询用户的盐值和盐值排名，让用户了解自己在社区中的贡献和信用程度。例如，用户在个人中心页面点击“盐值”选项，系统调用此接口获取盐值和排名信息，并展示在页面上。

2) 等级查询接口

通过此接口查询用户的等级、当前经验值和下一级经验值，让用户了解自己的成长进度。例如，用户在个人中心页面查看自己的等级信息时，系统调用此接口获取相关数据并展示在页面上。

3) 成长记录查询接口

通过此接口查询用户的成长记录，了解用户在社区中的成长历程。例如，用户在个人中心页面查看成长记录时，系统调用此接口获取符合条件的记录，并展示在页面上。

3.4.8 数据统计接口

1) 用户行为数据统计接口

通过此接口统计用户在一段时间内的行为数据，帮助运营人员了解用户的行为习惯和兴趣爱好。例如，运营人员在数据分析页面选择时间范围和行为类型，点击“统计”按钮后，系统调用此接口获取统计数据，并展示在页面上。

2) 内容数据统计接口

通过此接口统计内容在一段时间内的数据表现，为创作者提供参考。例如，创作者在内容管理页面点击“数据统计”按钮，选择时间范围后，系统调用此接口获取内容的浏览量和互动量等数据，并展示在页面上。

3) 社区健康度统计接口

通过此接口统计社区在一段时间内的健康度指标，帮助运营人员了解社区的整体状况。例如，运营人员在社区管理页面点击“健康度统计”按钮，选择时间范围后，系统调用此接口获取违规率、内容审核时效和用户活跃度等数据，并展示在页面上。

3.4.9 第三方接口

1) 社交媒体分享接口

用户通过此接口将知乎上的内容分享到其他社交媒体平台。例如，用户在内容详情页面点击“分享到微信”按钮，前端发送请求到此接口，系统将内容分享到微信平台，并返回分享成功状态。

2) 支付接口

用户通过此接口进行付费咨询、购买会员等支付操作。例如，用户在付费咨询页面确认咨询订单后，选择支付方式并点击“支付”按钮，前端发送请求到此接口，系统完成支付操作并返回支付成功状态和交易 id。

3.5 环境需求

3.5.1 硬件环境

知乎软件系统对硬件环境的要求主要基于其高并发、大数据量的特点。服务器配置需要满足以下要求：

1) CPU：至少选用 8 核 CPU，以应对多任务处理和高并发请求。例如，在有大量用户同时在线进行内容浏览、发布和互动时，多核 CPU 可以保证每个请求都能得到及时处理。

2) 内存：配备 32GB 或以上大容量内存，确保系统在处理复杂数据运算和存储缓存数据时有充足的内存资源。比如，当进行个性化推荐计算或内容审核时，这些操作需要在内存中快速处理大量数据，大容量内存可以避免频繁的磁盘交换，提高系统性能。

3) 存储设备：使用 1TB SSD 硬盘作为主要存储设备，其高速读写能力能够加快数据的存储和读取速度。对于内容存储、用户数据等大量数据的保存以及系统日志等信息的记录，SSD 硬盘可以有效缩短数据访问时间，提升整体系统效率。

4) 网络带宽：确保服务器具备足够的网络带宽，以保障数据传输的速度和稳定性。建议网络带宽至少为 1Gbps，能够应对大量用户同时进行数据传输，如图片和视频的上传与下载等操作，避免因带宽不足导致用户等待时间过长。

3.5.2 软件环境

1) 操作系统

推荐使用稳定、安全的 Linux 系统，如 CentOS 或 Ubuntu。Linux 系统具有良好的多用户环境支持、高稳定性和强大的网络服务功能，能够有效管理硬件资源和调度多任务，为知乎软件系统提供可靠的运行基础。

2) 数据库管理系统

根据不同的数据特点和业务需求，可采用多种数据库管理系统。例如，MySQL 用于关系型数据存储，如用户信息、内容信息等；MongoDB 可用于处理非结构化数据，如日志文件等。数据库管理系统要具备高效的数据存储、查询、更新和管理能力，以支持系统的高并发读写操作。

3) 中间件

使用 Redis 作为缓存中间件，可用于存储热点数据、用户会话信息等，以减轻数据库压力，提高数据访问速度。其高效的键值对存储和快速读写能力，能够在短时间内响应大量数据请求，提升系统性能。

4) Web 服务器

选用 Nginx 或 Apache 作为 Web 服务器，负责处理用户请求、转发请求到后端应用服务器，并返回响应结果。Web 服务器需要具备良好的并发处理能力、稳定的性能和安全防护机制，确保能够高效地处理来自不同用户的大量并发

请求。

5) 开发语言

后端开发可采用 Python、Java 等编程语言。Python 具有丰富的开源库和简洁的语法，适合快速开发和部署；Java 则以其强大的跨平台能力和高性能著称，能够应对大规模系统的开发需求。前端框架可使用 React、Vue 等，以提高开发效率和用户体验，实现动态、交互式的用户界面。

3.5.3 运行环境

1) 云服务器环境

知乎软件系统应部署在云服务器环境中，以实现资源的弹性伸缩和自动扩展。云服务器环境可以根据实际用户量和业务量的变化，自动调整分配的计算、存储和网络资源，确保系统在高负载情况下也能稳定运行，避免因资源不足导致系统崩溃或性能下降。

2) 数据备份和恢复机制

建立完善的数据备份和恢复机制，定期对系统数据进行备份，并将备份数据存储在多个不同的地理位置。在发生数据丢失、硬件故障或自然灾害等意外情况时，能够及时恢复数据，确保系统的可用性和数据的完整性。同时，要定期进行备份数据的验证和恢复演练，以保证备份机制的有效性。

第 4 章 用例模型

4.1 用例图

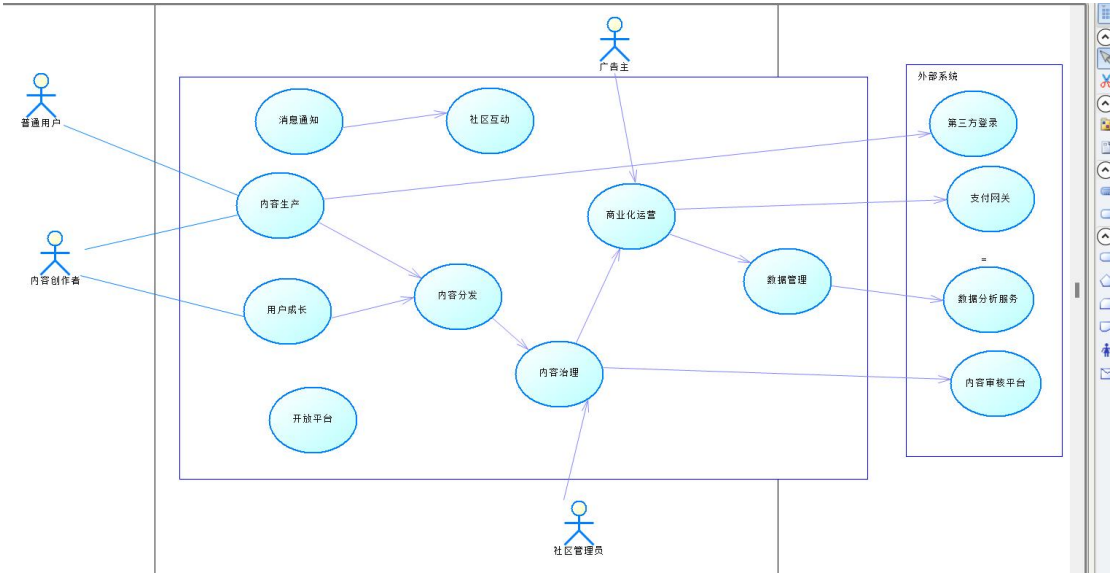


图 4.1 “知乎” 知识分享社区用例图

4.2 用例规约

表 4.1 多模态内容创作用例规约

ID	UC1.1
用例名	多模态内容创作
用例描述	支持用户创建问答/文章/视频/直播等内容
参与者	普通用户（基础内容）、内容创作者（专业内容）
前置条件	1. 用户已登录且通过实名认证 2. 盐值≥300（视频/直播需≥500）
后置条件	内容进入审核队列，草稿自动云同步
正常流程	1. 选择内容类型（问答/文章/视频/直播） 2. 输入正文并添加话题标签 3. 上传多媒体附件（图片/视频）

	4. 提交至内容分发系统
分支流程	专业领域内容需关联权威参考文献 直播内容需提前预约时间
异常流程	1. 内容包含违禁词→触发即时拦截（错误码 E1002） 2. 附件大小超限→提示压缩重传

表 4.2 个性化内容推荐用例规约

ID	UC1.2
用例名	个性化内容推荐
用例描述	根据用户画像生成首页内容流
参与者	推荐算法引擎
前置条件	1. 用户有至少 7 天行为数据 2. 内容已通过审核
后置条件	推荐结果曝光埋点记录
正常流程	1. 获取用户兴趣标签和历史行为 2. 计算知识图谱关联度 3. 混合排序（60%兴趣+30%热点+10%探索） 4. 返回推荐结果列表
分支流程	1.新用户冷启动阶段采用热度优先策略 2.创作者等级 Lv5+内容加权 20%
异常流程	1. 数据源异常→降级返回全站热榜 2. 算法超时→启用缓存结果

表 4.3 三级传播互动用例规约

ID	UC2.1
用例名	三级传播互动

用例描述	处理点赞/收藏/分享行为
参与者	普通用户
前置条件	1. 内容可见且未折叠 2. 用户当日未达到互动上限
后置条件	创作者收到实时通知
正常流程	1. 用户点击互动按钮 2. 系统校验权限（如收藏需 Lv2+） 3. 更新内容互动计数 4. 同步至创作者数据中心
分支流程	1.专业领域内容需盐值≥500 才能评论 2.敏感内容禁止分享至站外
异常流程	1. 检测刷量行为→临时封禁互动权限 2. 内容突然下架→撤回所有关联互动

表 4.4 AI+人工审核用例规约

ID	UC3.1
用例名	AI+人工审核
用例描述	7x24 小时内容合规审查
参与者	瓦力 AI 系统、社区管理员
前置条件	1. 内容已提交至审核队列 2. 审核资源池可用
后置条件	更新内容状态（通过/限流/删除）
正常流程	1. AI 初筛（置信度≥90%自动裁决） 2. 低置信度内容进入人工复审 3. 管理员做出最终判定 4. 同步处罚结果至用户成长系统

分支流程	1.争议内容发起社区陪审团投票 2.大 V 创作者内容优先审核
异常流程	1. 审核超时（>5 分钟）→自动放行并标记 2. 系统过载→触发熔断机制

表 4.5 知+内容推广用例规约

ID	UC4.1
用例名	知+内容推广
用例描述	广告主定制化内容投放
参与者	广告主、付费用户
前置条件	1. 广告主完成企业认证 2. 内容符合商业化规范
后置条件	生成 ROI 分析报告
正常流程	1. 创建推广任务（目标人群/预算） 2. 匹配目标创作者 3. 内容合规性预审 4. 按 CPM/CPC 计费投放
分支流程	1.品牌任务需签署电子合同 2.效果达标自动续投
异常流程	1. 点击率异常→暂停投放并核查 2. 余额不足→降权展示

表 4.6 盐值信用体系用例规约

ID	UC5.1
用例名	盐值信用体系
用例描述	动态计算用户社区贡献值

参与者	所有注册用户
前置条件	用户有至少 1 次有效互动
后置条件	更新等级权益
正常流程	<ol style="list-style-type: none">1. 监测用户行为（生产/互动/违规）2. 计算当日盐值增减3. 累计总分更新等级（Lv1-Lv10）4. 同步权益变化（如创作权限）
分支流程	<ol style="list-style-type: none">1.连续 30 天盐值≥800 授予"优秀答主"标识2.重大违规直接降至 Lv0
异常流程	<ol style="list-style-type: none">1. 数据异常→冻结积分变动2. 申诉成功→补偿历史积分

4.3 用例活动图

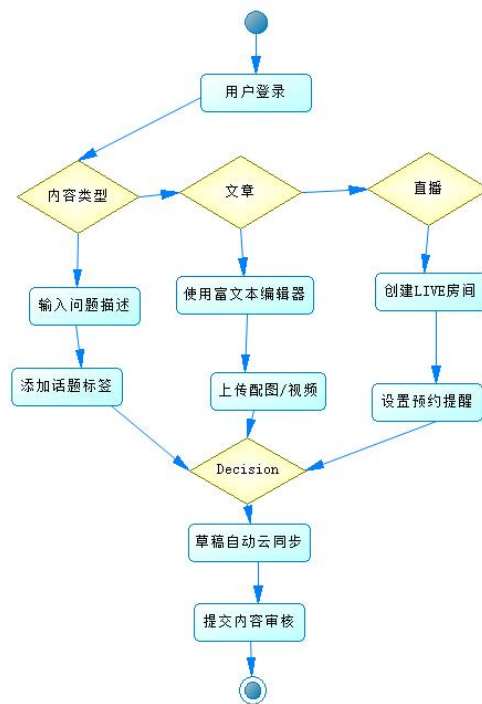


图 4.2 内容生产活动图

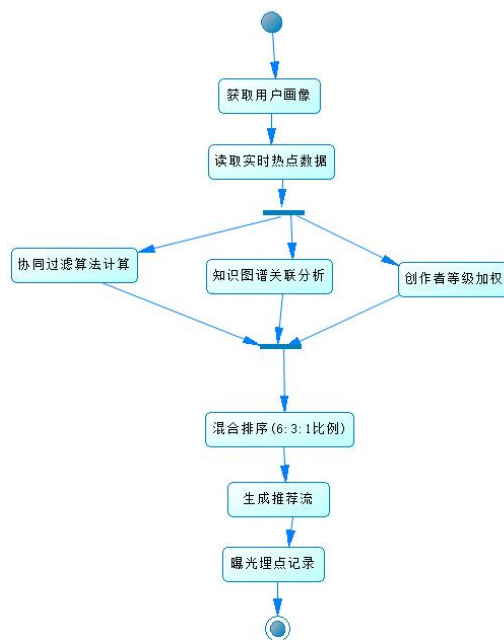


图 4.3 内容分发活动图

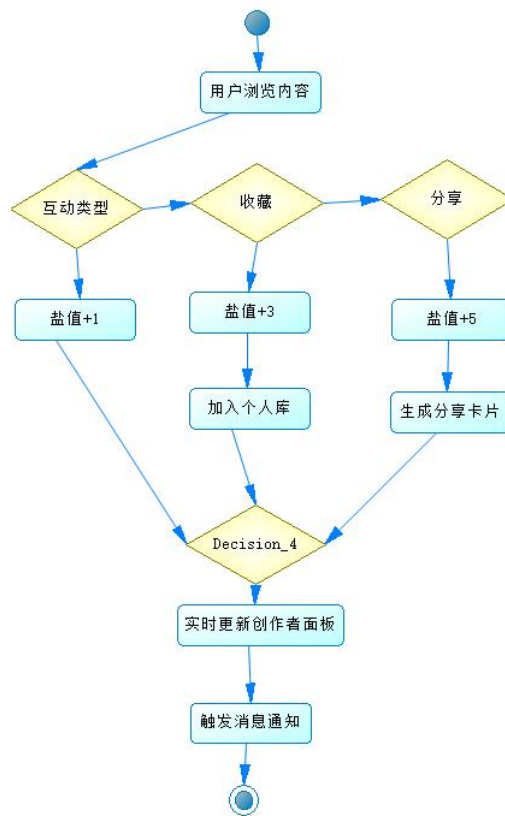


图 4.4 社区互动活动图

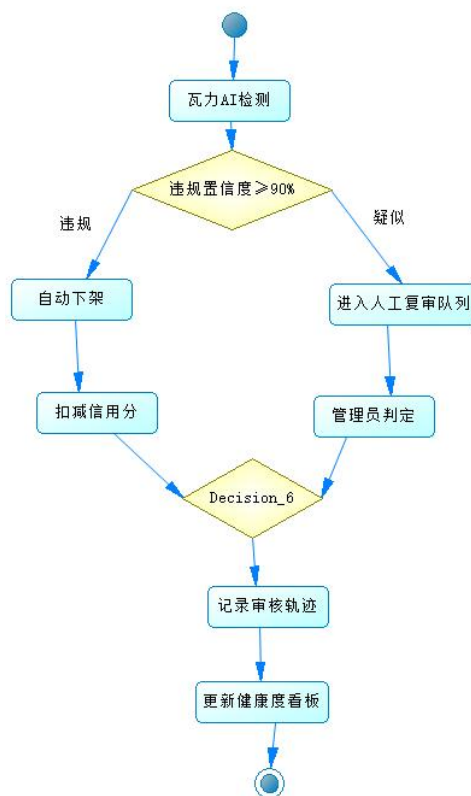


图 4.5 内容治理活动图

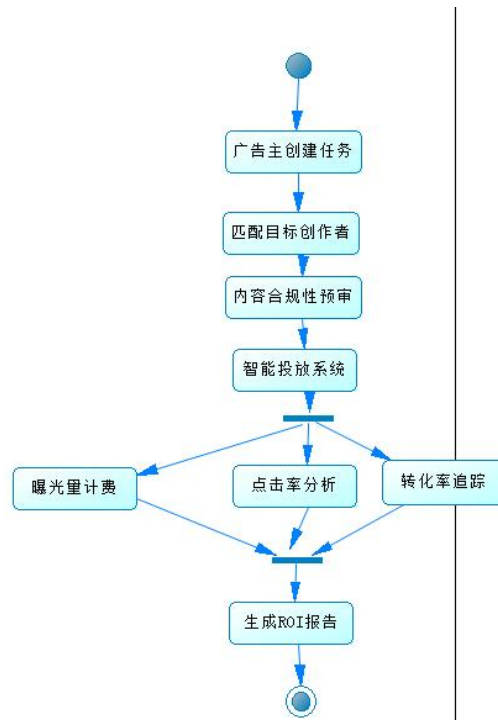


图 4.6 商业化运营活动图

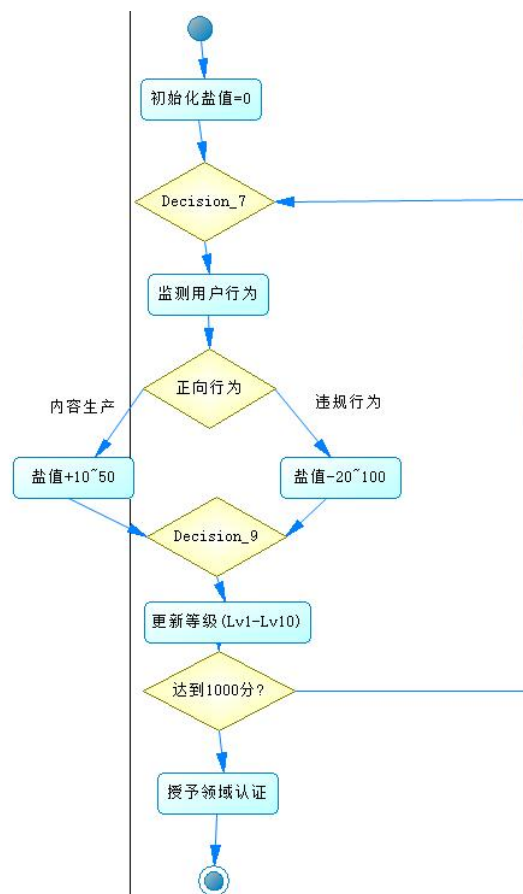


图 4.7 用户成长活动图

第 5 章 验收准则

本系统的验收准则为：

- 1.对本文档进行多人评审。
- 2.在项目计划的完成时间点，以评审会议的方式组织验收。
- 3.通过系统测试。

第 6 章 附录

6.1 大模型运用

6.1.1 识别模块的运用

本文档在识别功能模块时，通过上网搜索得到一些功能，再通过大模型给出测试建议，对不同功能选择相应需求类别。

6.1.2 大模型二次检测

在第 3 章各种需求中通过大模型进行检验区分是否正确，是否有添加的地方。如性能需求中，添加了“搜索功能性能”，从原本的“响应速度”的性能需求中单独提写出来。

6.1.3 语言润色

通过大模型，将文档语句的专业性进行润色。如原本语言“用户数据能够通过网络进行保存，从而保障数据安全。用户也可以根据此在不同地方，用账号登录后，能继续编辑。”通过大模型修改为“用户编辑内容时，草稿自动云同步，确保数据安全。用户可在不同设备上登录账号，继续编辑或发布草稿，实现多端无缝创作体验。”

6.1.4 大模型的逆向分析进行图表解读

利用大模型的识图分析功能，补充图片的文字说明，使其更加详细更加规整。
在 3.3 数据需求的数据元素/实体及其属性中进行运用，通过我们创建的表，将详细的文字描述自动生成出来，并通过简单调整进行运用

6.1.5 大模型应用的局限性

1) 领域知识偏差

大模型可能对知乎特有的业务逻辑理解不足，需人工复核关键用例。如盐值体系，需要人为参与进行分析，选择性摘要数据，并提供大模型相关数据进行预处理。

2) 数据风险

一方面测试中生成的模拟用户数据需经脱敏处理，避免真实用户信息泄露。另一方面需要重新查看数据是否是大模型的幻读。

3) 动态场景适配

针对快速迭代的功能模块，大模型的训练数据可能有所不足，确保建议的时效性。

6.2 分工

小组讨论确定后共同完成，以下不分先后，仅按学号排名，仅展示小组成员主要修改部分。

表 6.1 分工表

成员姓名	具体职责	贡献度评分
朱烨彬	第 1 章，第 3 章功能性需求，第 4 章	33.3%
许世睿	第 2 章，第 3 章性能需求，第 4 章	33.3%
唐欣轶	第 3 章接口需求，第 3 章环境需求，第 4 章	33.3%