<类chatshare跨模态平台>

前景文档

版本 <1.0>

修订历史记录

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **版本** | **说明** | **作者** |
| <日/月/年> | <x.x> | <详细信息> | <姓名> |
| 2023/6/23 | 1.0 | 界面原型迭代 | 蒋宸 |
|  |  |  |  |
|  |  |  |  |

目录

1. 简介 4

1.1 目的 4

1.2 范围 4

1.3 定义、首字母缩写词和缩略语 4

1.4 参考资料 4

2. 定位 5

2.1 商机 5

2.2 问题说明 5

2.3 产品定位说明 5

3. 涉众和用户说明 5

3.1 市场统计 5

3.2 涉众概要 6

3.3 用户概要 6

3.4 关键的涉众/用户需要 7

3.5 备选方案和竞争 7

3.5.1 <Runway ML> 7

3.5.2 <Playground AI> 7

4. 产品概述 7

4.1 产品总体效果 7

4.2 功能摘要 8

4.3 假设与依赖关系 8

5. 产品特性 8

5.1 <多模态整合> 8

5.2 <灵活部署> 8

5.3 <多端适配> 8

5.4 <自由定制> 8

5.5 <预设模块> 8

5.6 <功能新颖> 8

5.7 <多模型选择> 8

5.8 <智能聊天> 8

5.9 <参数控制> 8

5.10 <风格预设> 9

5.11 <训练加速> 9

5.12 <模板库> 9

5.13 <历史记录> 9

5.14 <用户分级> 9

5.15 <内容审核> 9

6. 约束 9

7. 质量属性 9

7.1 <易用性> 9

7.2 <可靠性> 10

7.3 <性能> 10

7.4 <可支持性> 10

7.5 <安全性> 10

8. 优先级 10

9. 其他产品需求 10

9.1 适用的标准 10

9.2 系统需求 11

9.3 环境需求 11

10. 文档需求 11

10.1 用户手册 11

10.2 联机帮助 11

10.3 安装指南、配置文件、自述文件 11

前景

# 简介

## 目的

此文档的目的是收集、分析和定义<<类chatshare跨模态平台>>的高层次需求和特性。它侧重于涉众和目标用户所需的功能以及这些需要存在的原因。<<类chatshare跨模态平台>>如何满足这些需要的详细情况记录在用例和补充规约中。

## 范围

<<类chatshare跨模态平台>>项目的定位，涉众和用户，概述，特性，约束，质量属性，优先级和其他产品需求，文档需求。

## 定义、首字母缩写词和缩略语

AIGC：人工智能生成内容。

跨模态：在不同模态（如视觉、听觉、文本等）之间进行信息转换或关联的过程。

多模态：多模态是指利用多种不同形式或感知渠道的信息进行表达、交流和理解的方式，通常包括视觉、听觉、文本、触觉等多种感官输入和输出方式。在计算机科学、人工智能和机器学习领域，多模态技术指的是通过整合来自不同模态的数据（如图像、文字、音频、视频等），从而增强模型的理解能力和推理能力。

云服务：通过网络按需提供可动态伸缩的廉价计算服务。

大模型微调：利用特定领域的数据集对已预训练的大模型进行进一步训练的过程。

## 参考资料

[1]《中国及全球人工智能生成内容(AIGC)市场规模预测》，中国产业研究院

# 定位

## 商机

1. 一站式AI创作平台：整合多种AI生成能力，满足创作者从文字到3D的全流程需求，提供其他单一功能产品无法提供的综合解决方案。
2. 降低技术门槛：通过简化界面和流程，让非技术用户也能使用高级AI生成能力，扩大潜在用户群体。
3. 垂直领域定制：可以针对特定行业开发专用模板和微调模型。

## 问题说明

|  |  |
| --- | --- |
| 问题是 | 当前AI生成工具碎片化严重，创作者需要在多个独立平台间切换才能完成跨模态内容创作，缺乏统一的创作环境和连贯的工作流 |
| 影响 | 数字内容创作者、市场营销人员、游戏开发者、教育工作者、中小企业和独立艺术家 |
| 问题的后果 | 创作效率低下，学习成本高，数据资产分散，难以实现多模态内容的协同创作，增加了项目管理和版本控制的复杂度 |
| 成功的解决方案 | 提供一站式跨模态AI创作平台，支持云端API和本地部署的灵活选择，实现不同模态生成内容间的无缝衔接，提供垂直行业定制模板和工作流 |

## 产品定位说明

|  |  |
| --- | --- |
| 针对于 | 数字内容创作者、市场营销人员、游戏开发者、教育工作者、中小企业和独立艺术家 |
| 谁 | 需要高效、灵活且经济地使用AI生成多种形式内容，但受限于现有工具的碎片化和高门槛 |
| 该（产品名） | AI跨模态内容创作平台 |
| 功能 | 统一平台集成文生文/图/3D/视频全流程创作，提供云端轻量化与本地高性能部署双重选项，独创跨模态内容联动工作流，企业级数据隐私和安全保障 |
| 不同于 | Midjourney、ChatGPT、Deepseek等单一功能平台 |
| 我们的产品 | 将分散的AI创作能力整合为有机整体，不仅提供工具集，更构建了完整的跨模态创作生态，支持从创意构思到成品输出的全流程协作 |

# 涉众和用户说明

## 市场统计

中国AIGC市场是全球AIGC市场的主要组成部分，中国AIGC市场受强大的市场需求及增强的AI技术所推动。数据显示，2022年中国人工智能AIGC核心产业规模11.5亿元，预计在2030年将达到4441亿元，未来中国有望发展为全球最大的人工智能AIGC市场。据预测，全球AIGC市场的市场规模将由2020年的13亿元增至2027年的1275亿元，复合年增长率达89.7%。[1]

初步调查身边同学（潜在用户群体之一）对跨模态平台的需求，包括以下三个问题：

1. 您使用生成式大模型辅助学习，工作的频率

|  |  |
| --- | --- |
| **选项** | **比例** |
| 极高 | 40% |
| 高 | 56.67% |
| 适中 | 3.33% |
| 低 | 0% |
| 从来不用 | 0% |

1. 您会同时使用多个模型吗

|  |  |
| --- | --- |
| **选项** | **比例** |
| 会 | 93.33% |
| 不会 | 6.67% |

1. 您认为对生成式模型按自己的要求进行微调有帮助吗

|  |  |
| --- | --- |
| **选项** | **比例** |
| 会 | 90% |
| 不会 | 10% |

由调查数据可知大学生群体经常使用生成式大模型，其中绝大多数都会使用多个模型，并且认为生成式模型如果能按自己的要求进行微调会有帮助，符合本项目的需求预期。但目前的调查仍然比较局限，有待完善，需要扩大调查范围，更全面地反应用户需求。

## 涉众概要

|  |  |  |
| --- | --- | --- |
| **名称** | **说明** | **角色** |
| 内容创作者 | 独立艺术家、自媒体人、设计师等创意工作者 | 提供真实使用反馈，参与UI/UX测试，验证创作流程合理性 |
| 企业客户 | 市场营销部门、产品设计团队等 | 确保企业级功能需求得到满足，验证数据安全和合规性 |
| 技术支持 | AI模型提供商 | 提供底层模型支持 |
| 监管机构 | 数据保护、版权管理等相关政府部门 | 制定合规标准，审核内容生成合法性 |
| 开发人员 | 软件工程原理与实践学生 | 开发该项目，完成课程大作业 |

## 用户概要

|  |  |  |
| --- | --- | --- |
| **名称** | **说明** | **涉众** |
| 自由创作者 | 使用基础生成功能完成个人创作 | 内容创作者 |
| 管理员 | 管理企业账户和项目协作 | 企业客户 |
| 技术人员 | 负责API对接和本地部署 | 企业客户，技术支持 |
| 审核员 | 内容合规性审查 | 监管机构 |

## 关键的涉众/用户需要

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **需要** | **优先级** | **关注的要点** | **目前的解决方案** | **提议的解决方案** | |
| 跨平台工作流断裂 | 高 | 需在不同工具间手动转移内容和数据 | 使用多个独立工具拼凑解决方案 | | 统一平台集成全流程创作工具链 |
| 数据隐私担忧 | 高 | 敏感商业数据需上传第三方云服务 | 限制使用或采用低效本地方案 | | 提供企业级本地化部署选项 |
| 成本不可控 | 中 | 各平台独立计费难以预估总成本 | 手工记录各平台使用量 | | 统一计费系统，提供成本预测工具 |
| 版权风险 | 中 | 生成内容可能侵权 | 人工审查或回避敏感内容 | | 内置版权检测和内容过滤系统 |
| 专业场景适配 | 高 | 通用模型难以适配特定专业场景 | 多次交互选取最优结果 | | 用户输入数据网站进行微调，对大语言模型实现基于ray或者deepspeed的分布式训练 |

## 备选方案和竞争

### <Runway ML>

提供文生视频、图生视频等多种AI创作工具，功能全面但价格较高。

### <Playground AI>

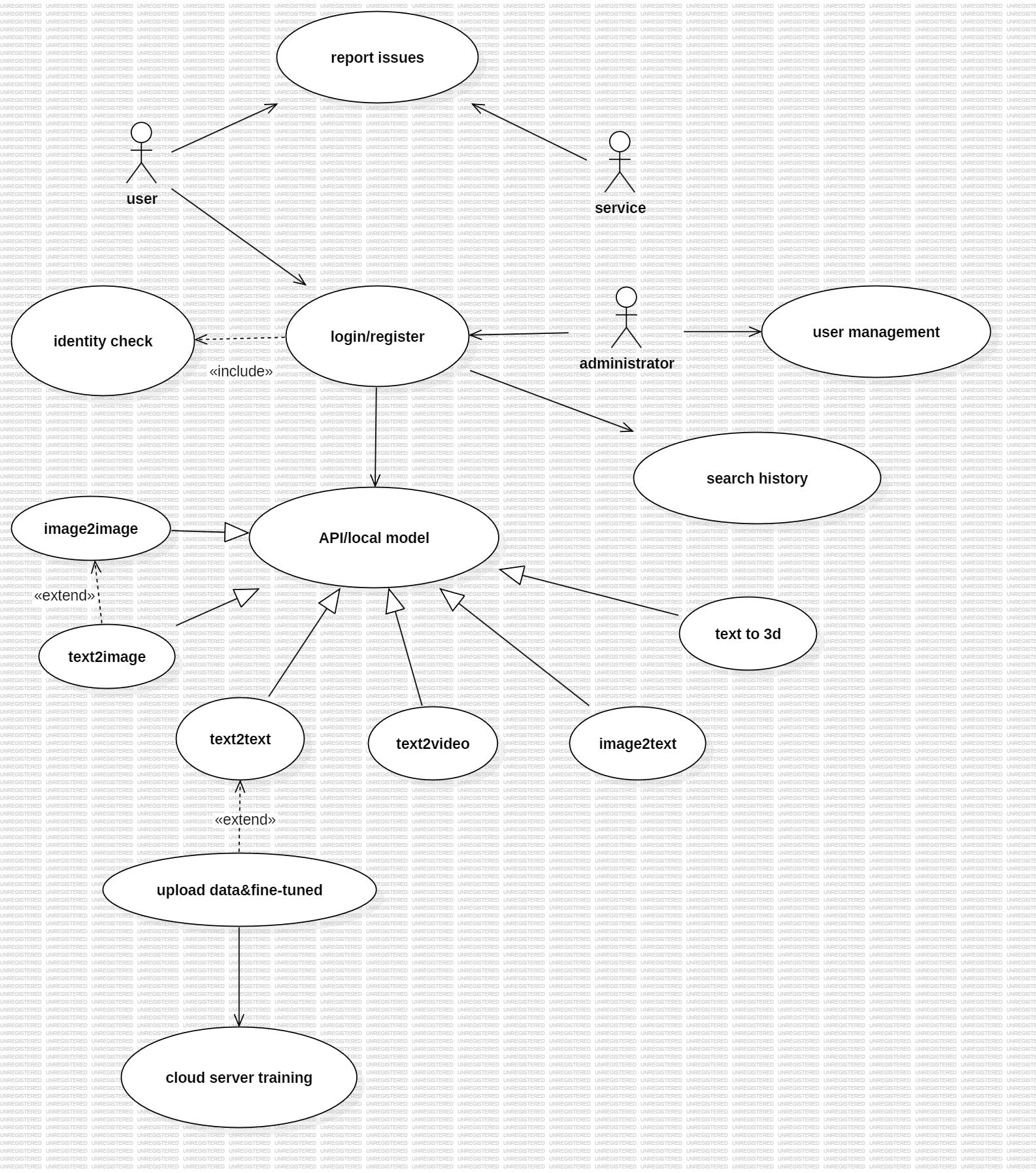
免费的文生图平台，支持多个开源模型，但功能单一。

# 产品概述

## 产品总体效果

该产品是一款跨模态AI内容生成平台，整合了文生文、文生图、图生图、文生3D、文生视频等多种AI生成能力，提供一站式的创作解决方案。该产品旨在解决当前AI工具碎片化、学习成本高、跨模态创作流程断裂等问题，适用于个人创作者、企业团队和教育机构。

该产品的用例图如下：



## 功能摘要

**客户支持系统**

|  |  |
| --- | --- |
| **客户利益** | **支持特性** |
| 减少工具切换，提升创作效率。 | 统一平台集成文生文、文生图、文生3D、文生视频等功能，支持跨模态内容联动。 |
| 灵活适应不同隐私和性能需求。 | 支持云端API调用和本地模型部署两种模式。 |
| 精准服务行业需要。 | 微调训练通用大模型适配特定专业场景。 |

## 假设与依赖关系

**关键假设**：

1.主流AI模型API（如OpenAI、Stability AI）将保持可用性和兼容性或开源环境良好

2.目标用户群体具备基本数字创作能力和需求

**技术依赖**：

必须项：第三方模型API的稳定访问或支持CUDA的GPU算力（针对本地部署方案）

可选项：Ray/DeepSpeed框架（分布式训练功能）、Blender/Polycam集成（3D内容后处理）

# 产品特性

## <多模态整合>

从文本生成图像、图像再生成文字/3D内容，支持自然语言与视觉数据的双向融合

## <灵活部署>

用户可选择在线调用或下载静态模型在本地运行，提升系统弹性和可控性

## <多端适配>

UI设计兼顾浏览器与移动设备，交互体验一致，适合多终端展示和使用

## <自由定制>

为个性化训练和私有数据场景提供接口与系统结构预设，满足专业场景输出要求

## <预设模块>

内置AI续写、角色扮演、高清放大、线稿填充等预设功能

## <功能新颖>

集成了文生视频，文生3D等功能，满足更多样化的创作需求

## <多模型选择>

支持切换不同生成模型，API调用多样化

## <智能聊天>

上下文感知的对话系统，支持多轮对话记忆

## <参数控制>

生图模块可调节分辨率、生成数量、采样器、随机种子等多种参数

## <风格预设>

生图模块内置油画、水彩、赛博朋克等典型艺术风格，支持自定义风格描述

## <训练加速>

基于Ray/Deepspeed的分布式微调，加速微调模型训练

## <模板库>

常用提示词模板一键保存/应用

## <历史记录>

保存对话记录历史，可以按时间或关键词筛选

## <用户分级>

用户分为不同等级，有不同的可用功能与任务优先级；支持管理员，客服等特殊用户

## <内容审核>

自动过滤违规生成内容

# 约束

1. 技术约束
   1. 前端：JavaScript+React
   2. 后端：Java（对话）/ Python（微调）
   3. 大模型API依赖
   4. 部署环境：云端服务器
2. 设计约束
   1. 兼容性：多端适配；适应用于客户端不同的分辨率/尺寸。
   2. 性能：数据量不少于10k，在1000并发的场景下，响应时间<3s
3. 安全/法律约束
   1. 数据隐私：用户输入需加密存储，符合GDPR/CCPA等法规
   2. 内容审核：生成内容需过滤违规信息
   3. API 许可：商用API需合规订阅
4. 社会/文化约束
   1. 多语言支持：需处理不同语言的Prompt输入
   2. 伦理风险：避免生成虚假信息、深度伪造内容

5.时间约束：大作业限期四周完成

# 质量属性

## <易用性>

1. 支持Prompt模板/历史记录复用
2. Web/APP界面一致性
3. 提供API文档及用户指南

## <可靠性>

1. 用户操作错误需友好提示，而非服务崩溃
2. 用户生成内容持久化存储，防丢失

## <性能>

1. 响应时间：1000并发用户下，文生文响应时间 <3秒
2. 存储数据：支持数据量不少于10k

## <可支持性>

1. 模块化设计
2. 代码规范
3. 文档说明

## <安全性>

1. 用户数据加密传输及存储
2. API调用需认证
3. 自动过滤违规生成内容

# 优先级

|  |  |
| --- | --- |
| **功能** | **优先级** |
| 多模态整合 | 高 |
| 本地部署 | 中 |
| 多端适配 | 低 |
| 微调模型 | 中 |
| 文生文 | 高 |
| 文生图 | 高 |
| 图生图 | 低 |
| 图生文 | 低 |
| 文生视频 | 中 |
| 文生3D | 中 |
| 历史记录 | 中 |
| 登录/注册 | 高 |

# 其他产品需求

## 适用的标准

**数据隐私**：GDPR（欧盟通用数据保护条例）

**AI伦理与安全**：ISO/IEC 42001（AI管理系统标准）

**网络安全**：ISO 27001（信息安全管理）、SOC 2 Type II（云服务安全审计）

**内容审核**：《互联网信息服务算法推荐管理规定》

**媒体格式标准**：

图片：JPEG/PNG/WebP

3D模型：GLTF/GLB

视频：MP4/H.264/WebM

**API协议**：RESTful API（OpenAPI 3.0规范）、WebSocket（实时通信）

**模型格式**：PyTorch（.pt）/ TensorFlow（.pb）、ONNX（跨框架推理）

## 系统需求

1. 网络平台：阿里云
2. 服务器配置：
   1. 操作系统：Ubuntu 22.04 64位
   2. CPU：8核vCPU
   3. 内存：30GB
   4. GPU：NVIDIA A10
   5. 公网带宽：5Mbps (峰值)

## 环境需求

1. 开发环境：
   1. Node.js:22+
   2. Java:23+
   3. Python:3.12+

（2）浏览器兼容性：Chrome/Firefox/Safari/Edge

# 文档需求

## 用户手册

1. 目的：说明操作流程和使用技巧
2. 内容：分模块说明文生文/图/3D/视频的操作流程，不同场景下的Prompt编写技巧，模型微调配置指 南，本地大模型部署步骤，手册索引，术语表等
3. 格式：Web/md
4. 打印：A4尺寸

## 联机帮助

暂无需求

## 安装指南、配置文件、自述文件

暂无需求