



# CONVEST

## 立项建议书

文件状态：	文件标识：	CONVEST
<input type="checkbox"/> 草稿	当前版本：	1.0
<input checked="" type="checkbox"/> 正式发布	作者：	李阳，应悦，章程路
<input type="checkbox"/> 正在修改	完成日期：	2020-09-26

版 本 历 史

版本/状态	作者	参与者	起止日期	备注
1.0	李阳	应悦、章程路	2020-09-21 至 2020-09-26	CONVEST: connection&investment

# 目录

0. 文档介绍 .....	4
0.1 文档目的 .....	4
0.2 文档范围 .....	4
0.3 读者对象 .....	4
0.4 参考文献 .....	4
0.5 术语与缩写解释 .....	4
1. 产品介绍 .....	5
1.1 产品定义 .....	5
1.2 产品开发背景 .....	5
1.3 产品主要功能和特色 .....	5
1.4 产品范围 .....	6
2. 市场概述 .....	6
2.1 客户需求 .....	6
功能性需求 .....	6
非功能性需求 .....	6
2.2 市场规模与发展趋势 .....	7
3. 产品发展目标 .....	8
3.1 定性目标 .....	8
3.2 定量目标 .....	8
4. 产品技术方案 .....	8
4.1 产品体系结构 .....	8
4.2 关键技术 .....	9
5. 产品优缺点分述 .....	10
6. Make-or-Buy 决策 .....	12
7. 项目计划 .....	12
7.1 项目团队 .....	12
7.2 软硬件资源估计 .....	12
7.3 成本估计 .....	13
7.4 进度表 .....	13
8. 市场营销计划 .....	13
8.1 产品赢利模式和销售目标 .....	13
8.2 促销和渗透方式 .....	14
8.3 销售方式和渠道 .....	15
9. 成本效益分析 .....	16
10. 总结 .....	16

## 0. 文档介绍

### 0.1 文档目的

本立项建议书用以对项目整体开发提出概括性指导意见。通过阐述拟建项目的功能特色、市场定位、及技术方案等内容，便于项目组全体成员即时参考，以供调整校正，减少不必要的重复劳动。同时，将产品优缺点与效益分析列出，为后续迭代补充提供便利。

### 0.2 文档范围

本文档包括产品介绍、市场概述、产品发展目标、技术方案、产品优缺点分析、项目计划及营销计划等。

### 0.3 读者对象

项目组开发人员；运维管理人员；产品客户；投资方等。

### 0.4 参考文献

- [1]郝晓玲，韩松乔，韩冬梅 《信息系统开发——方法、案例与实验》 清华大学出版社 2012.3
- [2]菜鸟教程 《HTML 教程》<https://www.runoob.com/html/html-tutorial.html>
- [3]高清阔，张永淼《浅析信息系统开发的项目风险管理》中小企业管理与科技(下旬刊). 2020 年 07 期

### 0.5 术语与缩写解释

缩写、术语	解 释
SPP	精简并行过程, Simplified Parallel Process
PIM	立项管理, Project Initialization Management
HTML	超文本标记语言, HyperText Markup Language
SQL	结构化查询语言, Structured Query Language

表 1: 术语与缩写解释

# 1. 产品介绍

## 1.1 产品定义

CONVEST 是一款针对股票投资散户与机构，提供 A 股上市公司信息辅助用户进行投资决策的网站。通过知识图谱技术实现企业高管人际关系网络的搭建，搜集汇总企业相关资讯，调用已有企业信息数据库的接口，网站可以实现高管人脉查询、企业资讯获取、企业基本信息概览等功能，满足用户投资前查询企业相关信息的需求。

## 1.2 产品开发背景

随着经济发展，散户和机构股票投资需求日渐增长，我们发现市场上缺少一款专门针对投资用户的企业信息查询产品。虽然目前市场上的企业信息查询产品种类繁多，但提供的信息太过冗杂，删选投资需要的信息需要花费较多时间，而且可视化较差不够直观；专业股票投资软件操作复杂，存在一定技术门槛。目前 A 股企业信息数据是公开可得的，网站搭建技术也十分成熟，所以我们认为开发一款面向散户和机构、提供 A 股企业信息的网站是切实可行并且有一定市场的。

## 1.3 产品主要功能和特色

### 1. 基础功能：

- a. 用户登录部分：
  - i. 用户注册 包括角色选择 机构用户/散户
  - ii. 用户登录、设置密码、修改密码、退出
- b. 附加功能部分：
  - i. 将部分特色功能隐藏，只有付费成为 VIP 用户才能解锁隐藏信息
  - ii. 一键导出企业信息报告，方便用户离线阅读

### 2. 特色功能：

- a. 企业高管信息查询部分：
  - i. 信息查询：用户在输入框内输入待查高管，结果通过关系网图的形式展现
  - ii. 信息追踪：对于图中可能潜在的信息，跳转到百度搜索界面实现一键查询
  - iii. 热门推荐：提供热词榜，引导用户查询
- b. 投资信息部分：
  - i. 信息查询：输入待查企业，展示企业相关信息
  - ii. 可视化看板：当前股票信息使用可视化方式呈现
- c. 投资组合信息订阅部分：

用户可以输入自己希望追踪的投资组合，系统会通过邮件方式定期发送投资报告

## 1.4 产品范围

### 1. 适用领域

本产品涵盖了股票投资需要查询的企业信息，适于机构或投资散户查询 A 股企业辅助投资。由于本网站提供的信息针对性较强，而且仅包含 A 股企业，审计、咨询等其他行业。

### 2. 产品内容

本产品包含全部 A 股上市企业，不包含其他公司企业；提供的数据包括企业高管关系、企业基本信息、近期企业资讯，但仅进行信息的展示与查询，不提供投资建议。

## 2. 市场概述

### 2.1 客户需求

本产品面向一切有企业人事资讯查询分析需求的投资机构或个人。对于已有明确投资意向的机构而言，掌握特定投资组合的公司背景、发展方向及核心董事会的人脉关系、社会信誉等信息是必不可少的投资背调；对于手持相对较小量级的富余资金，希望投身股票市场获得盈利的个人而言，其他网站纷繁复杂的公司简介与布满投资术语的年报可读性不佳，他们更需要简洁易懂的可视化图表，以增强对各投资行业龙头企业的对比了解，有助于作出更明智的选择。

### 功能性需求

1. 用户的注册与登录。
2. 订阅指定企业的最新资讯，接收新消息提醒。
3. 搜索、查询指定企业的人事关系网，并可以下载对应的资讯简报。
4. 对想要进一步了解的企业，一键跳转至其官网（或百度百科）。
5. 看到清晰明了的企业数据分析可视化结果。

### 非功能性需求

1. 界面简洁清晰，便于使用者快速上手并掌握核心用法。

2. 本系统需要有较好的可维护性，可靠性，可理解性，易于扩充性。方便用户理解和操作，降低相关使用人员的时间成本。
3. 本系统应具有较高的安全性。涉及到全体登录用户的个人信息与账户密码，应对此类信息进行严格保密，防止因个人信息泄露而出现安全事故。
4. 要考虑到使用对象的不同需求：不同订阅需求的个性化设置、可视化的多风格协调等。
5. 要考虑运行条件的变化：包括技术条件（网络条件、硬件条件和软件系统平台条件等）的变化和应用方式的变化，如在具体应用中界面的变化、功能的剪裁、不同用户的职责分配和组合等。

## 2.2 市场规模与发展趋势

产品名	图标	公司	上线时间	优点	缺点
天眼查		北京天眼查科技有限公司 (北京金堤科技有限公司子公司)	2014/10/30	1. 产品定位在商业安全工具。2. 官方备案企业征信机构，可靠性强。	1. 更新时效慢，无即时查询功能；2. 页面体验有待加强。
企查查		苏州朗动网络科技有限公司	2014/3/12	1. 和其他同类产品相比，会员定价较低；2. 页面体验更好。	1. 功能导航栏过于纷繁杂乱，不够清晰直观。
启信宝		苏州贝尔塔数据技术有限公司 (上海合合信息科技发展有限公司全资子公司)	2014/5/28	1. 侧重在AI商业调查工具；2. 数据从宏观微观两个层面进行展示	1. 会员定价策略不够清晰，难以吸引用户买单。

表 2：竞品分析

截至 2019 年，天眼查的活跃用户规模为 275.2 万，占总活跃用户的比重为 31.6%；企查查的活跃用户规模为 245.1 万，占总活跃用户的比重为 28.1%；启信宝的活跃用户规模为 125.2 万，占总活跃用户的比重为 14.4%。三家企业为行业前三，在总活跃用户中合计占比超过 74%。

期初，市场上建立在 web 网页端的企业信息咨询平台大多追求“广而全”的目标，“抢人大战”之后，各平台逐步进行差异化发展，也各自有擅长的领域。例如：天眼查的核心亮点是在关系发现，这就有一点大数据挖掘的意思，它利用去重名技术解决了做关系关联时遇到的重名问题。但这三款市场主流产品的共同不足是，在高管知识图谱、数据可视化和简报订阅追踪方面只做到了及格水平。

以上表中三巨头中的“企查查”为例，其“关系图谱”是静态的图片，可以对不同元素间距进行拖拽，但无法点击公司名或人名等直接跳转至相关详情页面。我们认为，随着资讯的爆炸式增长，咨询搜索者在很大程度上不会止步于看到简单陈列出的数据与图片，而更需要的是数据良好的可视化呈现和可更改的动态图表。所以基于这两点，本项目团队将从这几部分进行精进，填补市场上的空白。

### 3. 产品发展目标

#### 3.1 定性目标

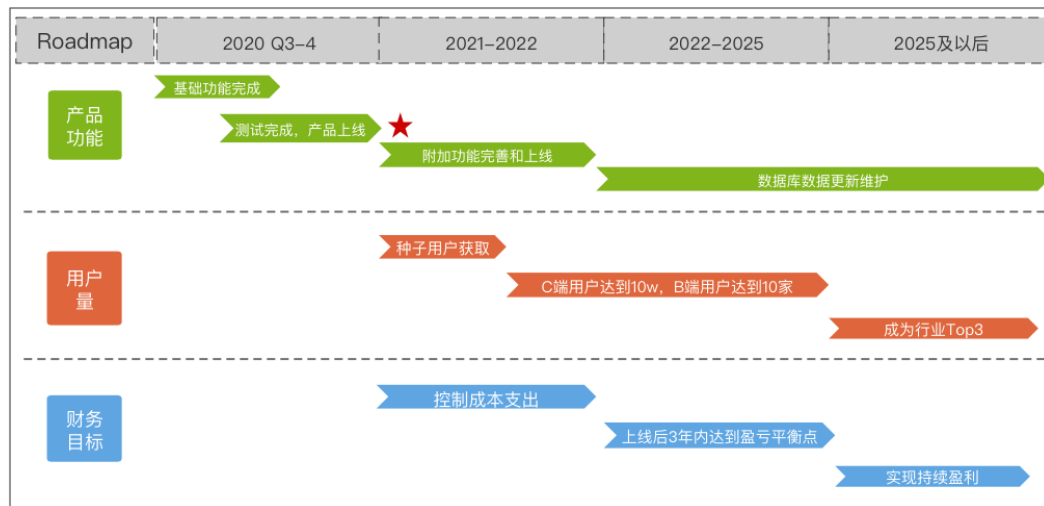


图 1：产品发展目标规划

#### 3.2 定量目标

1. 产品功能目标
  - a. 2020 年 11 月底前完成完整产品开发，并经过完整测试进行产品上线
  - b. 2021 年-2022 年完成附加功能的完善，并及时更新迭代
  - c. 后续进行产品定期数据维护
2. 产品用户量目标：
  - a. 2021 年产品上线后，进行种子用户获取
  - b. 2022 年-2025 年，预期 C 端用户突破 10w，B 端用户达到 10 家
  - c. 2025 年后有望成为行业 top3 的企业信息查询软件
3. 产品财务目标：
  - a. 2021 年-2022 年控制成本支出，预计 2022 年-2025 年间实现盈亏平衡
  - b. 2025 年后实现持续盈利

### 4. 产品技术方案

#### 4.1 产品体系结构



绘制技术架构图，如下图所示：

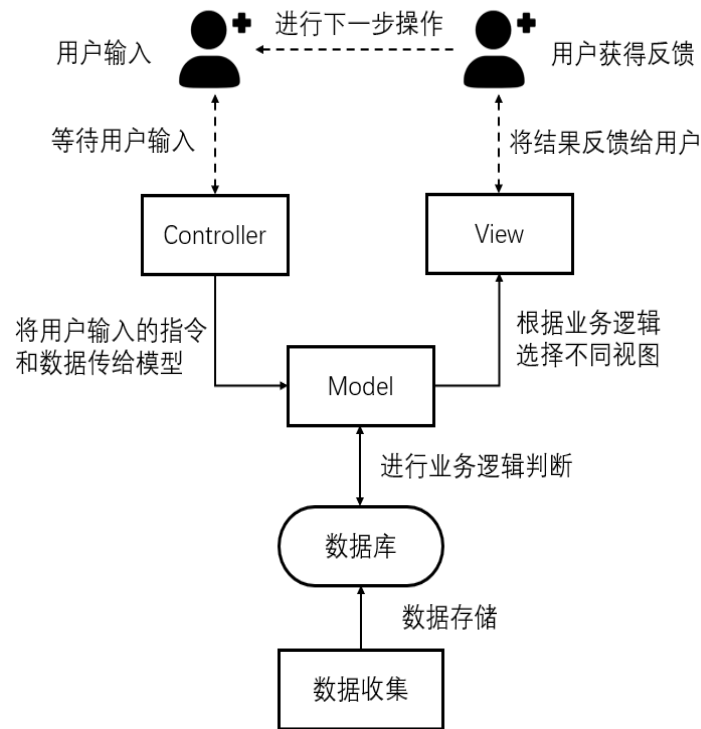


图 2：技术架构图

本网站使用 MVC 架构，把软件系统分为三个基本部分：模型（Model）、视图（View）和控制器（Controller）。

M 是 model 层，用来固化数据，封装与业务逻辑相关的数据和对数据的处理方法。M 层分为 Bean 和 DAO，Bean 保存数据的实体，并提供接口；Dao 为数据库提供了抽象接口的对象，在不暴露底层持久化方案实现细节的前提下提供了各种数据访问操作。这里我们使用 eclipse 软件 java 语言进行编写。

C 是 conrollor 层，控制器起到不同层面间的组织作用，用于控制应用程序的流程。控制器层里常常放一些业务逻辑方面的代码，一些业务方面的逻辑事件封装成一个个类中的方法，且方法中必调用 M 层中的 DAO 对象，来完成业务逻辑方面的操作。这里使用 eclipse 软件 jsp 服务器脚本语言来实现。

V 是 view 层，是应用程序中处理数据显示的部分。通常视图是依据模型数据创建的。这里使用 vscode 软件 Html + CSS 页面布局。

数据库使用 Mysql，将收集的高管数据信息和企业相关信息存储在数据库中。

技术细节参见 4.2 关键技术

## 4.2 关键技术

本产品需要用到的技术能力主要分成四个部分：

### 1. 高管简历特征抽取

- 概述：该部分的技术实现将分为三个阶段，数据爬虫、简历特征抽取、数据清洗

- b. **数据爬虫**: 使用 spyder 爬虫架构对股城网的高管简历数据进行爬取
  - c. **特征抽取**: 使用标注精灵进行人工特征标注, 并搭建 Bi-LSTM-CRF 深度学习模型, 训练得到能够自动抽取高管简历数据的模型
  - d. **数据清洗**: 对抽取到的特征信息进行清洗和整理, 输出结构化的完整高管特征数据库
2. 知识图谱构建和可视化
- a. **概述**: 该部分的技术实现将分为两个阶段, Neo4j 图数据库搭建和 web 接口调用
  - b. **Neo4j 图数据库搭建**: 将前一个环节中得到的特征数据结构化地导入 Neo4j 的数据库, 构建现有所有上市公司人脉网络图谱
  - c. **web 接口调用**: 将搭建完成的 Neo4j 数据库展示和查询接口调用到 web 前端界面中, 实现前端界面输入关键信息, 后端可以自动生成查询语句进行查询操作, 并使得查询到的知识图谱动态展示在 web 页面中
3. Web 前端界面开发
- a. **概述**: 该部分的技术实现将分为三个阶段, 知识图谱的调用和查询操作、固定的投资组合动态追踪和定期邮件推送、股票信息的查询和展示操作
  - b. **知识图谱相关操作**: 此处主要通过 Neo4j 的接口调用来实现, 可以实现用户输入关键词, 例如: 高管名称、股票代码、股票名称、公司名称或简写、学校名称等字段, 后台可以搜索到和这些关键词相关的所有图数据库中的信息, 并用关系图的方式进行呈现。此处涉及到的技术难点主要在于, 需要把用户输入的语句的逻辑用某种策略变成 Neo4j 的查询代码, 此处的实现难度仍需后续评估。
  - c. **动态追踪操作**: 对于已经有明确投资意向的用户来说, 对于特定股票的实时信息获取将会是一个很重要的功能, 因此我们计划实现投资报告的自动生成和邮箱推送, 这里的技术实现暂未确定具体的实现方式, 待后续开发阶段根据实际需求选择合适的开发方式。
  - d. **股票信息查询和展示操作**: 除了对于企业高管关系的查询之外, 有很多基础的股票信息对于投资者形成投资决策来说也是至关重要的, 因此这里计划使用 E-chart 工具将股票信息进行可视化呈现, 实现方式主要是使用 E-chart 模型并调用数据库中的数据进行组合, 对关键信息进行可视化呈现
4. Mysql 数据库开发
- a. **概述**: 此处主要是通过爬虫得到关键数据, 结构化存储在数据库中, 并在前端页面需要用到对应的数据时进行动态调取和呈现, 技术实现分为两个步骤: 数据库设计和数据库构建
  - b. **数据库设计**: 从实际需求出发, 分析对用户需求的辨伪存真, 挑选出真实有价值的信息, 对数据进行结构化整合, 构建完整的数据库模型
  - c. **数据库构建**: 根据设计的数据库模型, 使用 Navicat 进行实际数据库的实现

## 5. 产品优缺点分述

### 1. 产品功能: 产品功能部分主要分成以下三个板块:

- a. 企业高管信息可视化查询看板

- i. 信息输入: Neo4j 查询需要依靠特有编程语言, 需要实现用户在输入框内输入自然语言就可以在后台转化为查询语言实现数据库的查询
    - ii. 信息展示: 引用 Neo4j 接口, 实现查询出的图数据库结果展示
    - iii. 信息追踪: 对于图中可能潜在的信息, 接入百度接口, 实现一键查询跳转
    - iv. 热门推荐: 对于用户查询较多的信息, 提供热词榜, 引导用户查询
  - b. 投资信息可视化数据库
    - i. 信息输入: 简简单单一个搜索框
    - ii. 可视化看板: 当前股票信息使用可视化方式呈现
  - c. 投资组合信息订阅
    - i. 输入: 用户可以输入自己希望追踪的投资组合
    - ii. 输出: 系统会通过邮件方式定期发送投资报告
2. **产品质量:** 产品在开发资源允许的情况下将尽量考虑高并发、响应速度等对于用户体验较为关键的因素
3. **产品价格:**
- a. **盈利模式:** 产品的盈利模式将主要通过以下两个途径进行实现
    - i. 产品基础服务费: 付费会员账号
    - ii. 增值服务费: 定制化、个性化服务付费
  - b. **产品定价:** 由于产品的先入优势并不明显, 因此在初期, 产品将以较低的价格打入市场, 吸引足够的用户量, 等到用户量达到一定的规模之后参照实际收益数据和市场需求情况进行价格调整
4. **产品优点:**
- a. **完美解决市场需求:** 中国市场的投资者缺少一款实时的、精准的解决投资相关需求的产品辅助投资者进行决策, convest 的上线将很好的弥补市场的空白
  - b. **用户友好的交互体验:** 基于对用户的深入洞察, 产品设计将在细节处体现对用户思维和操作习惯的思考, 打造一款用户体验超预期的优质软件
  - c. **精准解决用户需求:** 经过前期市场调研和用户洞察, 我们发现用户存在以下痛点:
    - i. 散户缺少客观全面的渠道收集信息
    - ii. 机构用户缺少智能化分析数据的能力针对这两大痛点, 我们分别设计了不同的产品功能满足不同的诉求:
    - i. 针对散户: 我们将信息进行全面的收集和整理, 并进行可视化的呈现, 能够在帮助用户最大程度上节省时间的同时, 输出对于用户而言最有价值的数据
    - ii. 针对机构用户: 我们引入了深度学习模型对于原本复杂、分析难度大的文本数据进行特征抽取和可视化呈现, 使得深入分析和洞察从遥不可及变得轻松便捷
5. **产品缺点:** 由于产品仍然处于研发初期, 且受限于开发资源的限制, 产品仍会存在以下几点暂时的缺陷
- a. **知识图谱无法实现动态更新:** 由于模型训练需要租用服务器进行, 因此产品中的高管关系图谱部分无法实现从爬取-模型抽取-呈现这整条链条的动态更新

- b. **包含信息尚不全面：**受限于信息渠道的限制，无法获取企业内部数据，因此本产品所使用的数据均来自于公开数据的整合和呈现，不涉及到企业私密数据，因此对于完整的投资决策来说，只能作为辅助参考

## 6. Make-or-Buy 决策

1. 采购部分：
  - a. 涉及板块：上线服务器租用
  - b. 涉及风险：预计将会带来一定的资金成本，此处需要评估详细的 ROI 以确定最合适的服务器租用方案
2. 外包开发部分：
  - a. 本产品不涉及到外包开发，将全部由开发团队自行完成
3. 自主研发部分：
  - a. 涉及板块：核心算法、前端界面、后端数据库
  - b. 涉及风险：可能面临开发资源不足等问题，将采用敏捷开发模型，先行开发解决更多用户需求和满足更多市场趋势的功能板块，完成 MVP 后再迭代更多功能进行完善

## 7. 项目计划

### 7.1 项目团队

角色	知识技能要求	建议人选、人数	工作时间
项目经理	统筹安排，对整体开发进程把控	1	11
需求开发人员	调查分析能力，总结沟通能力，上传下达	3	7
系统设计人员	系统的构建能力，数据库的搭建能力	3	22
编程人员	编程能力，熟练掌握编程技术	3	22
测试人员	网站系统水平测试，发现问题的能力	3	12
质量保证人员	程序与网站的维护和日常监测	3	7
配置管理人员	对相关软硬件和系统配置的了解	3	7
服务与维护人员	对系统运行的了解以及一些后勤工作	3	7

表 3：项目团队规划

### 7.2 软硬件资源估计

资源名称	级别	详细配置	获取方式	费用
阿里云服务器	关键	用于网站上线	需要购买	50 元
Navicat	关键	用于数据库搭建	已经存在	无
Neo4j	关键	用于图数据库搭建	已经存在	无
j2ee	关键	用于前端网站开发	已经存在	无

表 4：软硬件资源估计

### 7.3 成本估计

条款	成本（人民币）	备注
人力资源	27000 元	成本=人月(9)*月薪(3000)
软硬件资源	55 元	主要用于购买阿里云服务器辅助网站上线
营销成本	10000 元/月	主要用于 BD 和广告投放
运维成本	5000 元/月	主要用于后续长期维护人员的工资

表 5：成本估计

### 7.4 进度表

		2020年9月			2020年10月				2020年11月				2020年12月		
		第二周	第三周	第四周	第一周	第二周	第三周	第四周	第一周	第二周	第三周	第四周	第一周	第二周	第三周
项目实施进度计划	网站功能构思														
	网站数据收集														
	数据库建设														
	网站框架搭建														
	网站前端设计														
	网站功能实现														
	网站测试 上线														
	网站运营推广														
	后期维护与展示														

图 3：项目计划表

## 8. 市场营销计划

### 8.1 产品赢利模式和销售目标

#### 赢利模式：

1. 对于 C 端用户：
  - a. 通过广告投放实现浏览量变现；
  - b. 开启会员服务。非会员：浏览基础信息，限制查看次数，可通过分享转发等形式换取更多的查看机会，下载简报收费 15 元/份；普通会员：开放全部功

能，可查看可下载，但限制次数，咨询简报免费额度 2 份/天；VIP 会员：提供个性化订阅组合，咨询简报免费额度 20 份/天。

2. 依靠 B 端赢利。依照客户需求将平台数据进行分类整合，随后售卖给投资机构、企业等；或者将数据进行打包整理，向有需求的金融机构收取接口调用费。

### 销售目标：

1. 在用户数量方面：短期通过密集营销和多层级广告投放实现快速的用户增长，通过三天免费试用 VIP 会员服务等手段，初期锚定种子用户，三年内将平台日活提升至 10000，实现 10 万用户基数和至少 10 家稳定合作的 B 端客户；长期希望在保持稳定的过百万日活的基础上，增设更多的个性化服务和更精准的订阅组合，扩大会员渗透率至 50%以上。
2. 在赢利金额方面：三年内实现盈亏平衡；长期维持稳定的 B 端赢利模式，实现融资，同时制定更精细化的付费会员分级制，实现月订阅收入突破百万。

## 8.2 促销和渗透方式

1. 促销渗透渠道和促销方式：
  - a. to B 渠道：主要为审计机构和投资机构等有公司研究和投资需求的机构，通过 BD 团队触达大客户，并且通过通用型方案售卖+根据不同公司要求进行部分产品功能个性化改造的方式，向大客户售卖整体解决方案，并且根据具体谈判情况调整收费标准
  - b. to C 渠道：对于个人投资者来说，建立全渠道营销体系，通过各种渠道触达用户后，通过个性化的营销服务提升留存率，并为产品迭代提供重要用户反馈参考，具体的营销体系见下图：

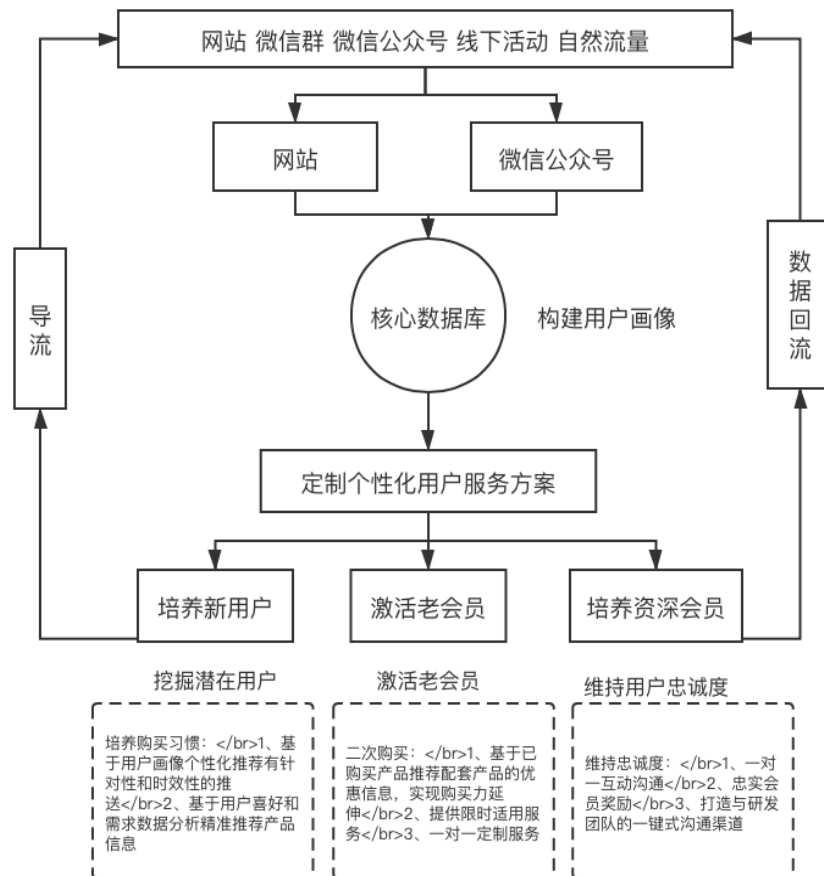


图 4：营销体系

### 8.3 销售方式和渠道

- 项目初期：**项目初期将主要依赖代理商的模式，进行快速的规模拓展。由于项目本身的先入优势不强，市场上已经出现了类似的产品，因此在短期内吸引足够多的用户使用本产品将会是防止竞争对手降维打击的重要战略，因此项目初期计划雇佣大量的代理商，进行短期内的规模拓展，以快速构建项目的护城河。
- 项目中期：**项目中期将构建自己的直销体系，由于代理商模式存在一定的不稳定性，且由于需要层层瓜分利益，长期来看对于公司盈利不利，因此项目中期将搭建起自己的销售团队进行进一步的规模拓展。
- 项目后期：**项目后期已经具有足够的用户量且产品已经具有一定规模，将采用联盟的方式和产业链上下游机构进行广泛的合作或联盟，实现优势互补。例如：可以和上游的数据收集商合作进一步拓展更具竞争壁垒的数据来源，也可以和下游的投资机构、投资者服务机构、审计机构合作，在售卖更多元化的服务组合的同时拓展更广的客户来源。

## 9. 成本效益分析

### 成本：

成本包括研发成本、运维成本和营销成本。其中已知和较可控的为研发成本和日常运维支出，就营销成本来看，初期需要较大投入以打响品牌、顺利进入市场；后续可能知名大平台也会跟进这部分的增值服务，需要进行下一步的产品迭代升级，届时需要进行下一轮的成本评估和营销方案设计。

### 效益：

1. 对于项目投资者来说，后期实现赢利后有股份分红；对于使用该平台的机构和公众来说，可以付出较少的成本快速得到所需的可视化数据。
2. 利于创建更规范的企业查询平台行业，填补了市场上这一部分的空缺；通过即时更新追踪企业动态的方式，增大企业的公众曝光度，间接性地对企业进行了社会督促，促使其进行公开透明的合法活动。

## 10. 总结

从市场分析层面，该项目瞄准的市场尚属于蓝海市场，具有广阔的市场前景；

从产品策划层面，产品能够切实解决用户痛点，具有较高的用户价值；

从技术壁垒和可行性层面，项目技术壁垒较高，且项目团队具有攻克核心技术的能力，能够在产品完成后形成很强的护城河；

从成本收益层面，项目投入少，但是预计收益高，具有很高的市场价值；

从团队构成层面，团队成员背景多元，具备丰富的算法模型实践经历、产品策划经验、web 和信息系统开发经验、商业实践经验等；

综上所述，项目具有较高的可行性和商业价值，值得投资和推广。