

2021-7-9

需求规格说明书

——KnowABall 足球垂直搜索引擎

《项目实训》课程 G07 小组

组长：张琦

组员：康大凯、聂俊哲、潘凯航、张溢弛

修改历史

修订日期	版本号	修改人	修改内容	审核人
2021-7-11	1.0	张琦、康大凯、聂俊哲、潘凯航、张溢弛	初稿	张琦
2021-7-12	2.0	张琦、康大凯、聂俊哲、潘凯航、张溢弛	终稿	张琦

目录

1	引言	5
1.1	需求规格说明书编写目的	5
1.2	软件项目背景	5
1.3	目标读者和阅读建议	5
1.3.1	软件需求规格说明书的组成部分	5
1.3.2	软件需求规格说明书面向的读者对象	6
1.4	参考文献与资料	7
2	总体描述	7
2.1	产品基本信息	7
2.2	产品前景	8
2.3	用户类及其特征	8
2.4	产品功能概述	9
2.4.1	搜索词条模块	9
2.4.2	搜索结果模块	9
2.4.3	历史搜索模块	9
2.4.4	内容展示模块	9
2.4.5	知识图谱模块	10
2.5	设计和实现上的约束	10
2.5.1	数据存储	10
2.5.2	网络服务吞吐	10
2.5.3	数据安全	10
2.6	假设和依赖	11
2.7	开发、测试和运行环境	11
2.8	项目文档	12
2.8.1	描述类文档	12
2.8.2	过程类文档	12
2.8.3	参考类文档	12
2.9	产品开发项目设计人员	12
3	系统需求分析概述	13
3.1	系统总用例图	13

3.2	需求优先级分析	14
3.3	系统模块架构图	15
4	功能需求	16
4.1	关键词搜索	16
4.2	关键词过滤	18
4.3	关键词交集	20
4.4	关键词并集	22
4.5	拼音搜索	24
4.6	搜索结果展示	25
4.7	结果详情展示	27
4.8	结果相关度展示	28
4.9	结果点击量展示	30
4.10	相关词条展示	32
4.11	优先补全	33
4.12	搜索提示	35
4.13	知识图谱问答	36
5	数据流图	38
5.1	搜索引擎子系统数据流图	38
5.2	网站维护子系统数据流图	38
6	DM 图	39
7	ER 图	40
8	外部接口需求	41
8.1	服务器接口	41
8.2	客户端接口	41
8.3	通信接口	42
9	非功能性需求	42
9.1	性能需求	42
9.2	数据传输与并发要求	42
9.3	数据管理要求	43
9.4	权限与安全需求	43
9.5	输入要求	44
9.6	可视化需求	44

9.7	防护性需求	44
9.8	软件质量属性	45
9.9	其他需求	45
10	数据字典	45
10.1	数据元素定义表	45
10.2	数据精确表	49
11	业务规则与业务算法	50
11.1	业务规则	48
11.2	业务算法	49
12	界面原型	49
13	上下文图	50
14	附录	51
14.1	术语表	51
14.2	待确定问题清单	53

1 引言

1.1 需求规格说明书编写目的

本软件需求规格说明书对《项目实训》的课程项目《足球领域垂直搜索引擎》的实现做了充分用户需求分析，针对足球爱好者等用户群体，清晰地描述了本搜索引擎的各项需求，包括功能性需求和非功能性需求。

本软件需求规格说明书旨在明确我们需要开发的足球搜索引擎应该具有的各项基本要求，并且尽可能完整地描述足球搜索引擎预期的各类行为和用户体验，使得项目经理和软件开发人员以及系统分析师可以清楚地了解不同用户的不同需求，为后续的系统设计和软件开发打下坚实的文档基础。

1.2 软件项目背景

本项目是浙江大学 2021 学年暑假短学期的《项目实训》课程的课程项目，目标为实现一个足球搜索引擎。

当今是计算机网络技术全面深入运用的时代，现代信息技术在各行各业都起到了重要作用。搭建足球搜索引擎，是为了解决在足球这个垂直领域下的用户需求痛点，可以方便足球爱好者和普通用户针对足球这个特定领域进行搜索，我们计划为用户提供足球运动员、球队和足球比赛、足球咨询等多维度的搜索服务。

同时我们对结构化和非结构化的数据都做了良好的处理和可视化，大大提升了用户的使用体验和人机交互的友好性。

1.3 目标读者和阅读建议

1.3.1 软件需求规格说明书的组成部分

本软件需求规格说明书主要由如下几个部分构成：总体描述、系统需求分析概述、功能需求、数据流图、外部接口需求、非功能性需求、数据字典、业务规则和业务算法以及附录这些部分。

- 总体描述部分对足球搜索引擎的基本信息、前景和用户类特征、设计和实现上的各项约束、假设和依赖、以及项目组的开发测试和运行环境等做了详细的介绍。

- 系统需求分析概述这一部分对系统的主要用例进行了需求优先级的分析并给出了整个系统的模块架构图，功能需求部分根据不同的用户类型对需求做了详细的分析，并给出了用例的时序图，是本软件需求规格说明书的主要部分。
- 数据库 ER 图给出了最基本的数据库中各项数据表的设计
- 数据流图部分对每个子系统的逻辑流动方向做了图形化的表达。
- 外部接口需求部分对硬件接口、软件接口和通信接口等等进行了描述。
- 非功能性的需求部分对该系统的需求进行了详细的描述。
- 数据字典部分对于系统的数据流定义、数据元素定义和数据精度做了详细的描述。
- 业务规则与业务算法部分详细列举了系统的操作规则和算法。

1.3.2 软件需求规格说明书面向的读者对象

- 项目经理：项目经理可以根据本文档了解预期产品的功能，并据此进行了系统设计和项目管理
- 设计人员：系统的设计人员可以根据本文档对需求进行分析，并设计出系统，包含系统架构的设计和数据库表的设计
- 程序开发人员：程序开发人员可以根据该需求规格说明书配合《系统总体设计计划》，了解系统功能，并进行系统的开发和《用户手册》的编写
- 销售人员：销售热源可以根据本文档了解目标产品的各项功能和性能
- 测试人员：测试人员可以根据本文档编写测试用例，并对软件产品进行功能性的测试和非功能性的测试
- 用户：用户可以根据本文档了解本产品的功能和性质，并和分析人员进行需求的讨论和协商
- 其他人员：其他人员比如部门领导、公司领导、政府机关领导等领导可以根据本文档了解产品的功能和性能

在阅读本软件需求规格说明书的时候，首先要了解产品的功能概貌，然后可以根据自身的需要对每一功能进行实当的了解

1.4 参考文献与资料

- 《软件设计文档国家标准》
- 《软件工程项目开发文档范例》
- 《软件需求》刘伟琴、刘洪涛译
- 《“软件需求工程”课程案例 足球搜索引擎 20200921》刑卫
- 《软件工程——实践者的研究方法》
- [G07]“足球搜索引擎”项目可行性报告
- [G07]“足球搜索引擎”项目章程
- [G07]“足球搜索引擎”项目计划
- [G07]“足球搜索引擎”需求工程计划
- [G07]“足球搜索引擎”前景与范围
- [G07]“足球搜索引擎”质量保证计划

2 总体描述

2.1 产品基本信息

产品名称

KnowABall 足球垂直搜索引擎

任务提出者

浙江大学项目实训课程组——任课教师：邵健；助教：陈泽群

开发者

浙江大学项目实训 G07 小组

用户

有检索足球数据信息需求的人

功能

根据关键字检索足球球队，足球球员以及与之相关的足球比赛数据和足球新闻资讯。而对于搜索将支持高级搜索、模糊搜索、拼音搜索等多项搜索功能。

性能

能够支撑至少 1000 名用户的使用请求。

完成期限

7 月 18 日前完成初步开发，7 月 23 日前完成最终系统。

2.2 产品前景

在足球运动的关注度日益升高的今天，一直缺少一个专门的搜索引擎将足球相关的各项数据聚合在一起，而传统搜索引擎检索足球信息时得到的结果杂乱，有较多的无关结果，且检索数据未分类不利于用户对目标信息进行查看。因此 KnowABall 作为针对足球领域的垂直搜索引擎，可以检索到专业化分类的搜索结果，而不会与用户的预期不符。

我们旨在通过收集足球领域不同来源的权威数据，通过多样化的搜索方式，对各种足球数据进行分类、聚合分析和整理，力求为用户提供完整直观的数据，改善用户搜索体验，使得对足球相关球员、球队和比赛数据的获取变得唾手可得。

2.3 用户类及其特征

我们将产品使用方定义为一个用户类，即对足球数据有检索需求的人。用户类的具体特征如下表所示：

使用者	主要特征	备注
用户	可以通过关键词进行搜索	包括高级搜索、模糊搜索等多种搜索方式
	可以查看搜索结果	搜索结果将按照相关度排序
	可以对搜索结果进行筛选	筛选可以根据结果类型进行
	可以查看相关球员、球队和新闻资讯	系统根据相关度进行推荐
	可以查看详细的新闻资讯信息	与搜索目标相关的新闻咨询
	可以查看详细的球员信息	球员的数据、俱乐部、位置等信息
	可以查看详细的球队信息	球队的数据、所属联赛、球员等信息
	可以查看足球相关的知识图谱	整合全部球员、球队等信息的

		知识图谱
	可以查看热搜	
	可以进行足球相关的智能对话	输入一些问题来获取自动问答

表 2-3-1 用户类及其特征

2.4 产品功能概述

2.4.1 搜索词条模块

该模块的基本功能为：用户输入关键字，从而对球员数据、球队数据和相关的新闻资讯进行搜索。搜索将支持高级搜索、关键词过滤、拼音搜索、模糊搜索、高亮显示、搜索历史、词条补全等功能。

而所谓高级搜索，也就是用户根据搜索符号规则组织搜索输入语句，从而对搜索范围进行限定，来获取更加精准的搜索结果。

2.4.2 搜索结果模块

用户输入关键字并点击搜索按钮后，系统会将搜索结果呈现给用户。搜索结果将以卡片列表的形式进行呈现，首先呈现相关的球员或球队词条，然后呈现相关的新闻列表。无论是哪种类型，用户都可以点击进去查看详细的内容。而搜索内容将按照搜索的相关度和时间进行排序。此外，也支持用户对想要查看的搜索结果类型进行筛选。

除此之外，系统将会根据用户的搜索内容，对用户进行相关人物和球队的个性化推荐，展示一些同类的词条方便用户跳转。而系统同样也会展现一些搜索热词为用户提供搜索参考。

2.4.3 历史搜索模块

系统将会对用户的搜索历史进行记录，搜索框默认显示用户最近两次的搜索词条，而在自动补全功能中，也将优先显示用户的历史搜索记录。

2.4.4 内容展示模块

用户对搜索结果进行点击后，将进行具体内容的展示。对于球员，将展示球员的年龄、国家、球场位置、当前身价、个人荣誉和总进球助攻数据等信息，然后会有一张列表来记录各个赛季该球员所效力的俱乐部和在该赛季的具体数据统计。同时会用图表的方式来展示球员的各项能力值。

对于球队，将展示该球队的所属联赛、联赛排名、主场等基本信息，并用一个列表来展现该球队中的所有教练、球员以及历年来联赛中的得失球、积分排名等数据。此外对于新闻资讯，除了展示基本的内容之外，还将贴上一些相关标签来指示新闻相关的人物和球队。当然球员和球队信息界面也会进行相关的推荐。

2.4.5 知识图谱模块

知识图谱模块包含两个部分，首先是利用知识图谱技术，通过一些关键词将各个球员和球队关联起来，构造出一个图谱。此外，还可以基于图表进行基于足球知识的智能自动问答。

2.5 设计和实现上的约束

系统的设计、编码以及维护将遵照文档严格进行。

在具体设计和实现上，按照以下约束进行。

2.5.1 数据存储

项目产品使用 MYSQL 数据库系统作为引擎，按照数据产生、转换和存储的策略，通过将数据导入数据库的方式进行数据的存储操作。此外，elastic search 作为搜索引擎，在对数据格式进行一定的处理转换后，也将起到数据存储的作用。

2.5.2 网络服务吞吐

根据项目要求，本项目将可以提供对外服务的能力，确保至少 1000 名用户可以同时使用本系统。

2.5.3 数据安全

将从完整性、保密性和可用性三个层面来保护用户的数据安全。

完整性保证数据在未授权的情况下不被修改，而且在一整个传输和存储的过程中不会被篡改、盗用和丢失。对此，将在一些成熟框架的支持下（如 `spring data` 等），通过一定的加密手段，综合运用一系列方案和技术实现上述要求。

保密性要求对数据进行加密，只有授权者才能使用，这一特性要求加密技术必须自动、实时、精确、可靠。

而可用性要求做到避免因为系统数据泄露而使得合法使用者无法接触可用数据，通过对使用者身份进行适当验证（比如说人机验证），为合法使用者提供更加安全便捷的使用。

2.6 假设和依赖

假设与依赖编号	假设与依赖描述
AS-1	平台可以从懂球帝、德国转会市场等网站爬取到充足的数据
AS-2	平台可以得到足够的服务器支持
AS-3	平台的更新信息能够得到实时刷新
AS-4	平台用户掌握了电脑的基本操作
AS-5	平台多轮的迭代开发过程有相应的人员负责维护与更新
AS-6	平台的开发人员具有足够的开发与学习能力
AS-7	不考虑一些不可控的突发意外

表 2-6-1 假设和依赖

2.7 开发、测试和运行环境

开发环境

Windows 10 操作系统

Ubuntu Server 20.04 系统

React+MySQL+Springboot+Mybatis+Elastic Search+jena+fuseki

测试环境

联网的 PC、linux 服务器、一些开源的开发和测试工具。

运行环境

联网的 PC、浏览器。

2.8 项目文档

所有项目文档将分为三类：关于系统功能和使用的描述类文档、在系统开发过程中产出的过程类文档、为开发提供方向和依据的参考类文档。这些文档将为整个开发过程以及后续的系统使用提供依据。

2.8.1 描述类文档

描述类文档将为用户对 KnowABall 足球垂直搜索引擎的使用提供帮助，包括展现整个系统的全部功能以及各个功能点的具体使用方式和注意点，这类文档的典型代表为用户手册。

2.8.2 过程类文档

过程类文档是对整个项目开发过程的里程记录，使得项目组人员更好地了解当前开发进展，这类文档的典型代表为测试报告。

2.8.3 参考类文档

系统在开发过程中，由于采取团队协作的迭代式开发，因此需要提前对一些系统功能和系统接口进行约定，这时候参考类文档也就发挥了重要作用，对一些系统的设计进行约定之后，将其作为参考类文档存储记录下来。这类文档的典型代码为需求规格说明书、设计概要说明书等等。

2.9 产品开发项目设计人员

编号	角色	角色职责	具体负责人
1	项目经理	在预算范围内领导小组成员	张琦

		完成所有的工作	
2	产品经理	负责进行市场调研,了解和设计具体的用户需求,并领导制定本项目的需求规格说明书	聂俊哲
3	设计总监	建立系统框架,负责系统的总体架构和具体模块的设计,参加技术评审	康大凯
4	测试团队经理	负责领导测试团队完成整个项目的测试	张溢弛
5	美术监督	设计网站原型,并负责维持网站的美术质量	潘凯航
6	质量经理	带领软件质量监督巡视组成员制定质量保证计划等文档,并对实际的文档和代码的质量进行监督评审	张琦
7	开发团队	负责完成搜索引擎的具体开发和部署工作	全体成员
8	测试团队	负责完成整个系统的测试和维护	全体成员
9	软件质量监督师	负责实时对质量经理以及项目经理提供的项目进度和项目实际开发时的差异提出报告并指出改进的办法。	全体成员

表 2-9-1 产品开发项目设计人员

3 系统需求分析概述

3.1 系统总用例图

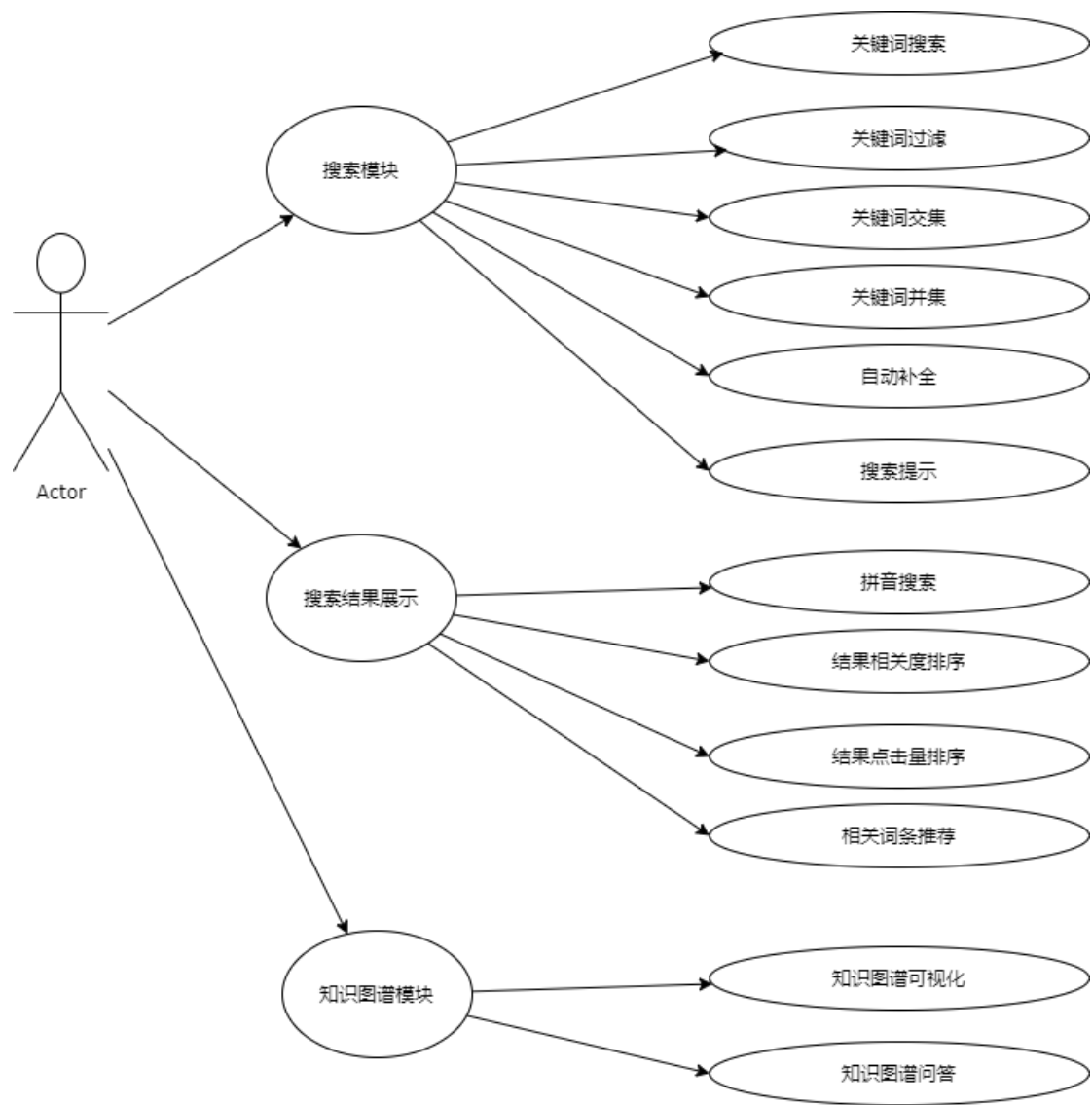


图 3-1-1 系统用例总图

3.2 需求优先级分析

本需求优先级分析模块依据风险-价值四象限矩阵进行评级，即按照下表所示四象限对需求进行优先级分类。

	高价值	低价值
高风险	首先处理	避免
低风险	其次处理	最后处理

表 3-2-1 风险-价值四象限矩阵

根据基本的用户需求分析，我们可以把定义的用户需求按照优先级作如下排序：

用例编号	用例名称	风险	价值	优先级
SE-01	关键词搜索	低	高	高
SE-02	关键词过滤	低	高	高
SE-03	关键词交集	低	高	高
SE-04	关键词并集	低	高	高
SE-05	拼音搜索	低	高	高
SE-06	搜索结果展示	低	高	高
SE-07	结果相关度排序	低	高	中
SE-08	结果点击量排序	低	高	中
SE-09	相关词条推荐	低	高	中
SE-10	优先补全	低	高	高
SE-11	搜索提示	低	高	中
SE-12	知识图谱问答	低	高	中

表 3-2-2 用例优先级表

3.3 系统模块架构图

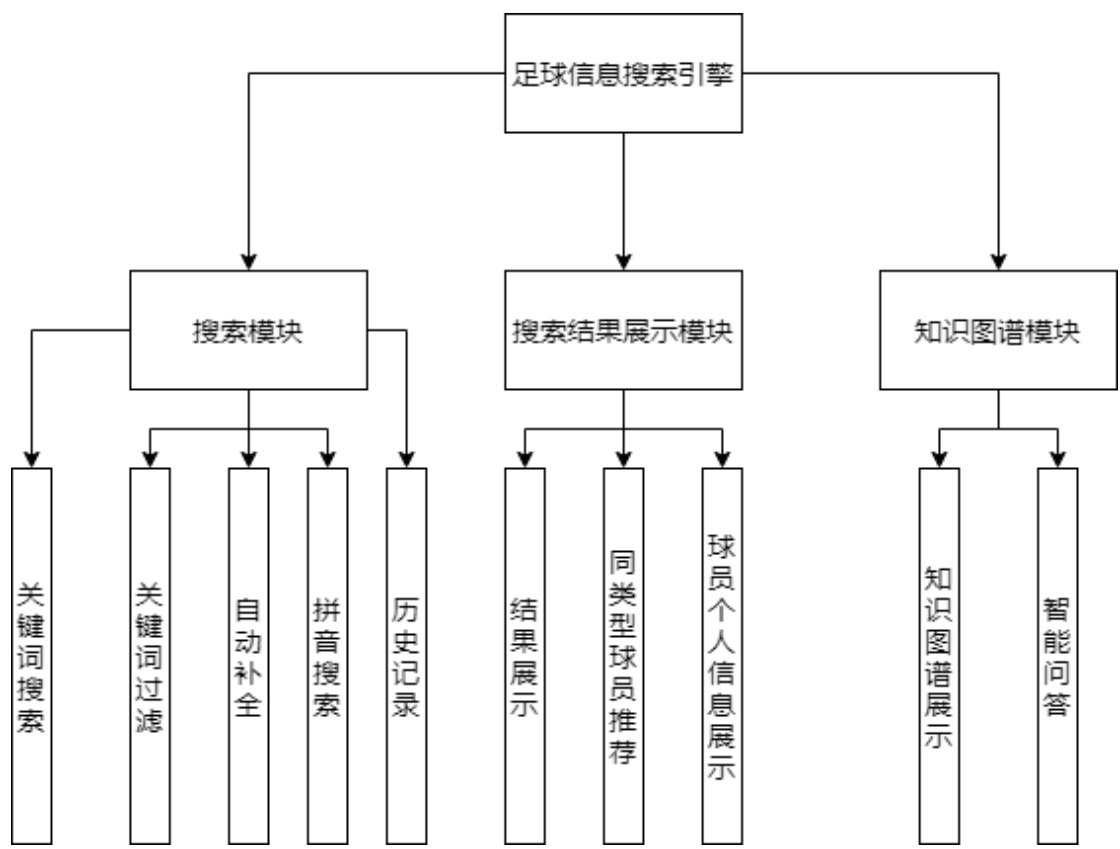


图 3-3-1 系统模块架构图

4 功能需求

4.1 关键词搜索

用例编号	SE-01	用例名称	关键词
创建人	张溢弛	最后修改人	张溢弛
创建日期	2021.07.11	最后修改日期	2021.07.11
角色	普通用户	需求来源	需求访谈和小组会议
主要参与者	普通用户		
描述	用户在搜索主页界面中的文本框中输入关键词或语句后，点击搜索按钮可以进行搜索		
前置条件	1.系统可以正常运行 2.用户拥有一定的电脑操作技术，会打字		
触发器	用户需要进入主页		

后置条件	用户搜索记录写入后端日志中
输入信息	用户希望搜索的关键词
主干过程	1.用户进入搜索引擎主页 2.用户输入关键词 3.用户点击搜索按钮 4.搜索成功后跳转至结果显示界面
分支过程	暂无
异常	1.用户输入危险信息（发生于主干过程2） 处理：系统提示“查询错误，请重新输入”，系统结束用例 2.搜索后页面出现异常（发生于主干过程4） 处理：根据网站的联系方式联系系统管理员，系统结束用例
假设	没有任何异常发生
输入	用户想要搜索的关键字和句子
输出	搜索结果显示页面
包括用例	关键词过滤，关键词交集，关键词并集
优先级	高
使用频率	高
备注	中

表 4-1-1 关键词搜索

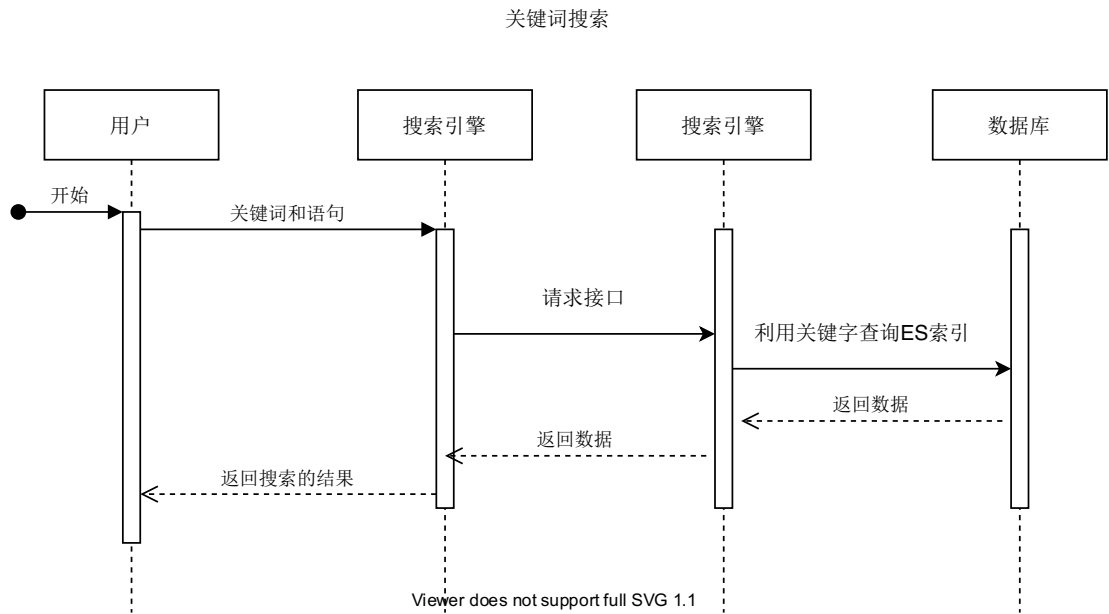


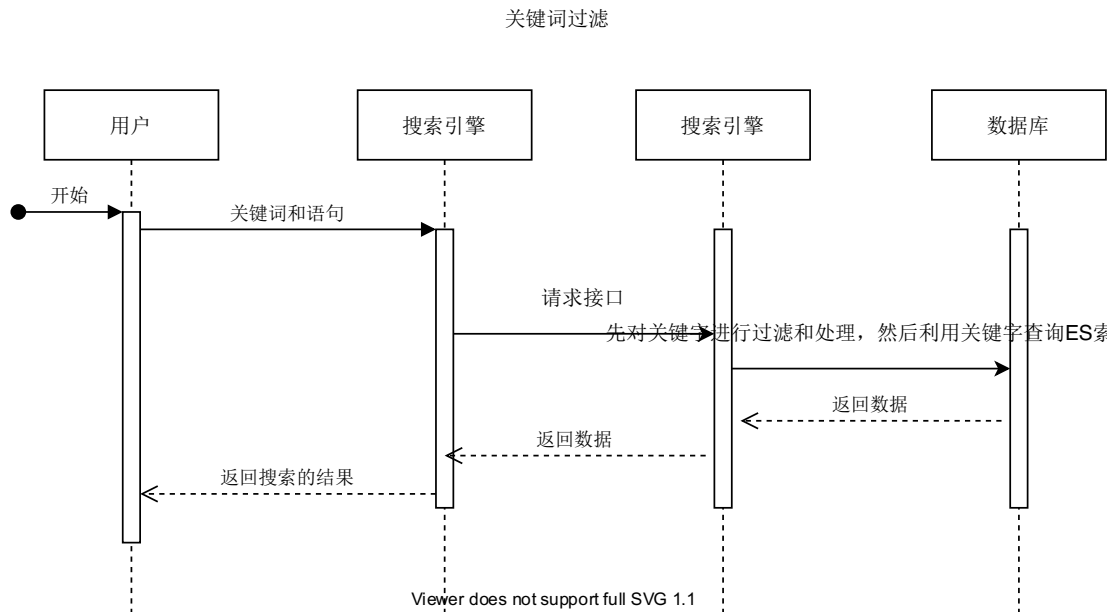
图 4-1-1 关键词搜索时序图

4.2 关键词过滤

用例编号	SE-02	用例名称	关键词过滤
创建人	张溢弛	最后修改人	张溢弛
创建日期	2021.07.11	最后修改日期	2021.07.11
角色	普通用户	需求来源	需求访谈和小组会议
主要参与者	普通用户		
描述	系统会对用户输入的关键词进行一定的过滤，过滤掉非法标点和词语，对输入的内容进行分词处理。		
前置条件	1.系统可以正常运行 2..用户拥有一定的电脑操作技术，会打字		
触发器	用户需要进入主页		
后置条件	用户搜索记录写入后端日志中		

输入信息	用户希望搜索的关键词
主干过程	1. 用户进入搜索引擎主页 2. 用户输入关键词 3. 用户点击搜索按钮 4. 搜索成功后跳转至结果显示界面
分支过程	暂无
异常	1. 用户输入危险信息（发生于主干过程2） 处理：系统提示“查询错误，请重新输入”，系统结束用例 2. 搜索后页面出现异常（发生于主干过程4） 处理：根据网站的联系方式联系系统管理员，系统结束用例
假设	没有任何异常发生
输入	用户想要搜索的关键字和句子
输出	搜索结果显示页面
包括用例	关键词交集，关键词并集
优先级	高
使用频率	高
备注	中

表 4-2-1 关键词过滤



4.3 关键词交集

用例编号	SE-03	用例名称	关键词交集
创建人	张溢弛	最后修改人	张溢弛
创建日期	2021.07.11	最后修改日期	2021.07.11
角色	普通用户	需求来源	需求访谈和小组会议
主要参与者	普通用户		
描述	1. 用户进入搜索引擎主页 2. 用户输入关键词 3. 用户点击搜索按钮 4. 搜索成功后跳转至结果显示界面		
前置条件	1.系统可以正常运行 2.用户拥有一定的电脑操作技术，会打字		
触发器	用户需要进入主页		
后置条件	用户搜索记录写入后端日志中		

输入信息	用户希望搜索的关键词
主干过程	1. 用户进入搜索引擎主页 2. 用户输入关键词 3. 用户点击搜索按钮 4. 搜索成功后跳转至结果显示界面
分支过程	暂无
异常	1. 用户输入危险信息（发生于主干过程2） 处理：系统提示“查询错误，请重新输入”，系统结束用例 2. 搜索后页面出现异常（发生于主干过程4） 处理：根据网站的联系方式联系系统管理员，系统结束用例
假设	没有任何异常发生
输入	用户想要搜索的关键字和句子
输出	搜索结果显示页面
包括用例	关键词过滤
优先级	高
使用频率	高
备注	中

表 4-3-1 关键词交集

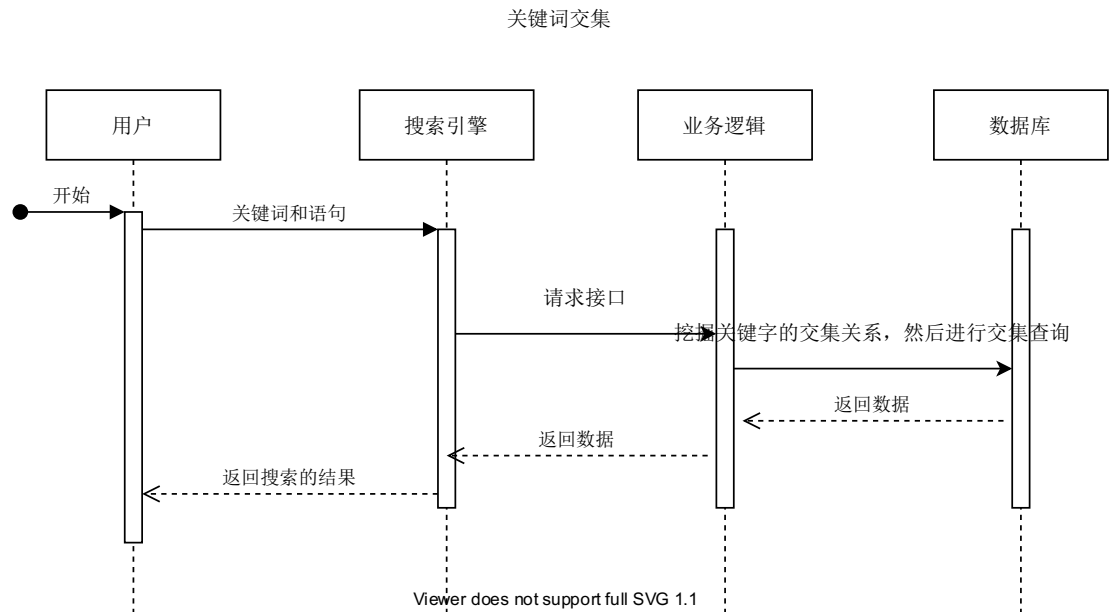


图 4-3-1 关键词交集时序图

4.4 关键词并集

用例编号	SE-04	用例名称	关键词并集
创建人	康大凯	最后修改人	康大凯
创建日期	2021/07/12	最后修改日期	2021/07/12
角色	用户	需求来源	用户代表
主要参与者	用户		
描述	用户在搜索界面的文本框中输入含有 的语句，点击搜索按钮进行搜索		
前置条件	1. 系统正常运行 2. 用户可正常打开相关页面		
触发器	用户需要搜索信息		
后置条件	用户搜索记录写入网站日志		
输入信息	含有 的搜索内容		
主干过程	1. 用户进入搜索页面 2. 用户输入关键词与符号，含 3. 用户点击搜索按钮 4. 搜索成功后跳转至结果显示界面		

分支过程	无
异常	1. 搜索后页面出现异常（发生于主干过程 4） 处理：根据网站的联系方式联系系统管理员 系统结束用例 2. 出现网络问题（发生于主干过程 1、4） 处理：检查网络，重新尝试搜索 系统结束用例
假设	无异常发生
输入	搜索内容的语句
输出	搜索结果页面
包括用例	搜索过滤、搜索交集
优先级	中
使用频率	低
备注	无

表 4-4-1 关键词并集

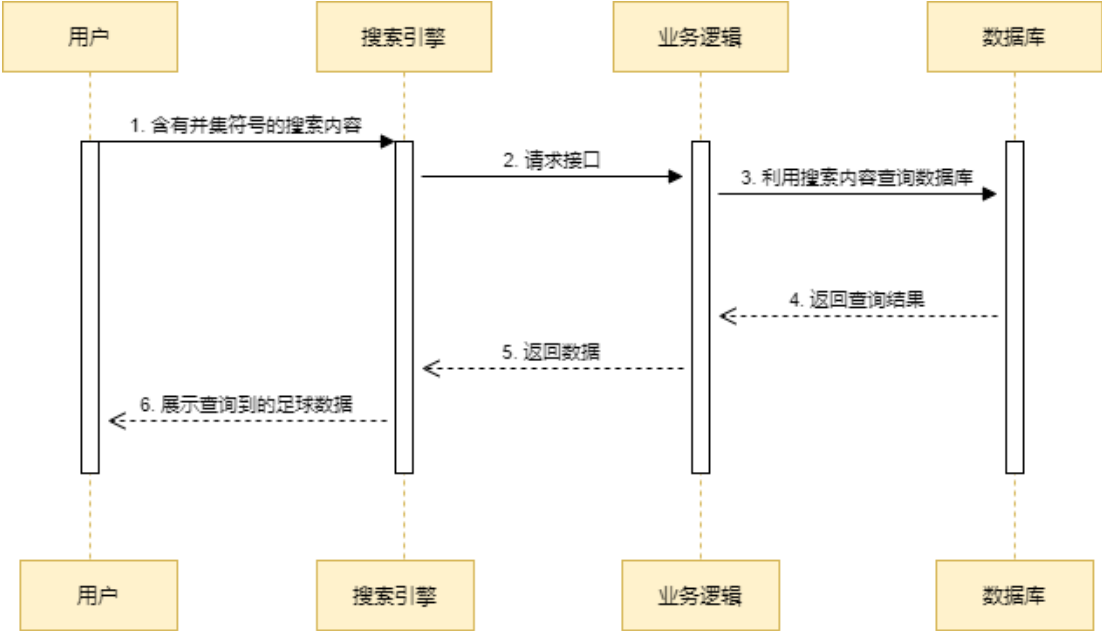


图 4-4-1 关键词并集时序图

4.5 拼音搜索

用例编号	SE-05	用例名称	拼音搜索
创建人	康大凯	最后修改人	康大凯
创建日期	2021/07/12	最后修改日期	2021/07/12
角色	用户	需求来源	用户代表
主要参与者	用户		
描述	用户在搜索界面的文本框中输入拼音形式的搜索内容后，点击搜索按钮进行搜索		
前置条件	1. 系统正常运行 2. 用户可正常打开相关页面		
触发器	用户需要搜索信息		
后置条件	用户搜索记录写入网站日志		
输入信息	拼音形式的搜索内容		
主干过程	1. 用户进入搜索页面 2. 用户输入拼音形式的搜索内容 3. 用户点击搜索按钮 4. 搜索成功后跳转至结果显示界面		
分支过程	无		
异常	1. 搜索后页面出现异常（发生于主干过程 4） 处理：根据网站的联系方式联系系统管理员 系统结束用例 2. 出现网络问题（发生于主干过程 1、4） 处理：检查网络，重新尝试搜索 系统结束用例		
假设	无异常发生		
输入	拼音形式的搜索内容		
输出	搜索结果页面		
包括用例	搜索过滤、搜索交集		
优先级	中		

使用频率	低
备注	无

表 4-5-1 拼音搜索

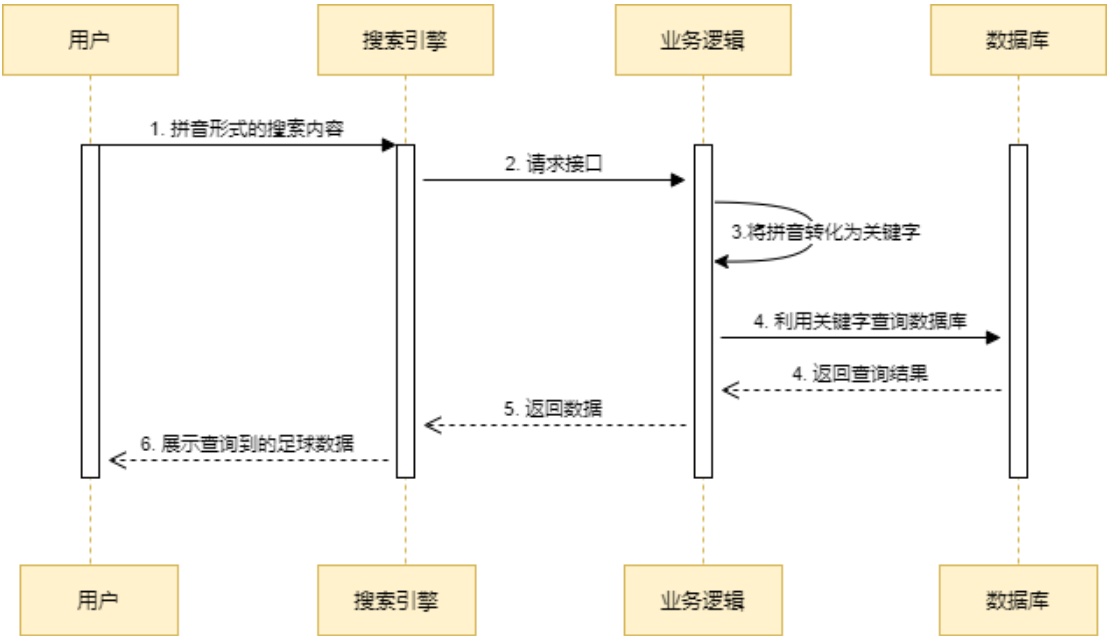


图 4-5-1 拼音搜索时序图

4.6 搜索结果展示

用例编号	SE-06	用例名称	搜索结果展示
创建人	潘凯航	最后修改人	潘凯航
创建日期	2021/07/11	最后修改日期	2021/07/11
角色	用户	需求来源	用户代表
主要参与者	用户		
描述	用户完成搜索，点击搜索按钮后，系统可以显示搜索结果		
前置条件	1. 系统正常运行 2. 用户可以正常打开相关界面		
触发器	用户需要进入搜索页面		
后置条件	无		
输入信息	搜索内容		

主干过程	1. 用户完成搜索输入 2. 用户点击搜索按钮 3. 结果页面按照默认相关度进行排序，显示搜索结果卡片列表 (并在右侧显示相关词条推荐)
分支过程	无
异常	1. 搜索后页面出现异常（发生于主干过程3） 处理：首先可能存在网络问题（网络连接状态不佳或者当前网络吞吐量过大），建议稍后重试，实在不行请联系系统管理员
假设	无异常发生
输入	关键词、语句
输出	搜索结果显示界面
包括用例	无
优先级	高
使用频率	高

表 4-6-1 搜索结果展示

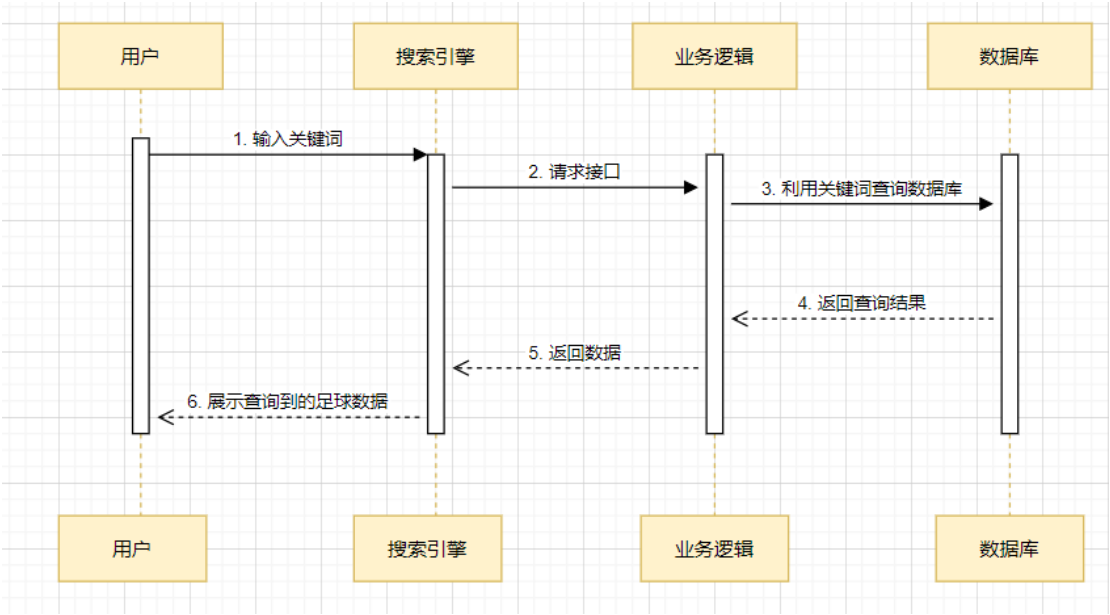


图 4-6-1 搜索结果展示时序图

4.7 结果详情展示

用例编号	SE-07	用例名称	结果详情展示
创建人	潘凯航	最后修改人	潘凯航
创建日期	2021/07/11	最后修改日期	2021/07/11
角色	用户	需求来源	用户代表
主要参与者	用户		
描述	用户对某搜索结果进行点击,进入球队或球员的详细界面,查看相关的一些信息		
前置条件	1. 系统正常运行 2. 用户可以正常打开相关界面 3. 搜索结果正常		
触发器	用户需要在搜索后进入某球队和球员的词条界面		
后置条件	无		
输入信息	对某球员或球队的词条进行点击		
主干过程	1. 用户完成搜索 2. 用户点击某结果词条 3. 展示该球队或球员的详细数据信息		
分支过程	无		
异常	1. 无法显示该球队或球员的详细数据（发生于主干过程3） 处理：可能该球员球队的数据发生了更新或丢失，也可能由于网络原因导致，建议稍后重试，实在不行请联系系统管理员		
假设	无异常发生		
输入	鼠标的点击事件		
输出	结果详情显示界面		
包括用例	无		
优先级	高		
使用频率	高		
备注	无		

表 4-7-1 结果详情展示

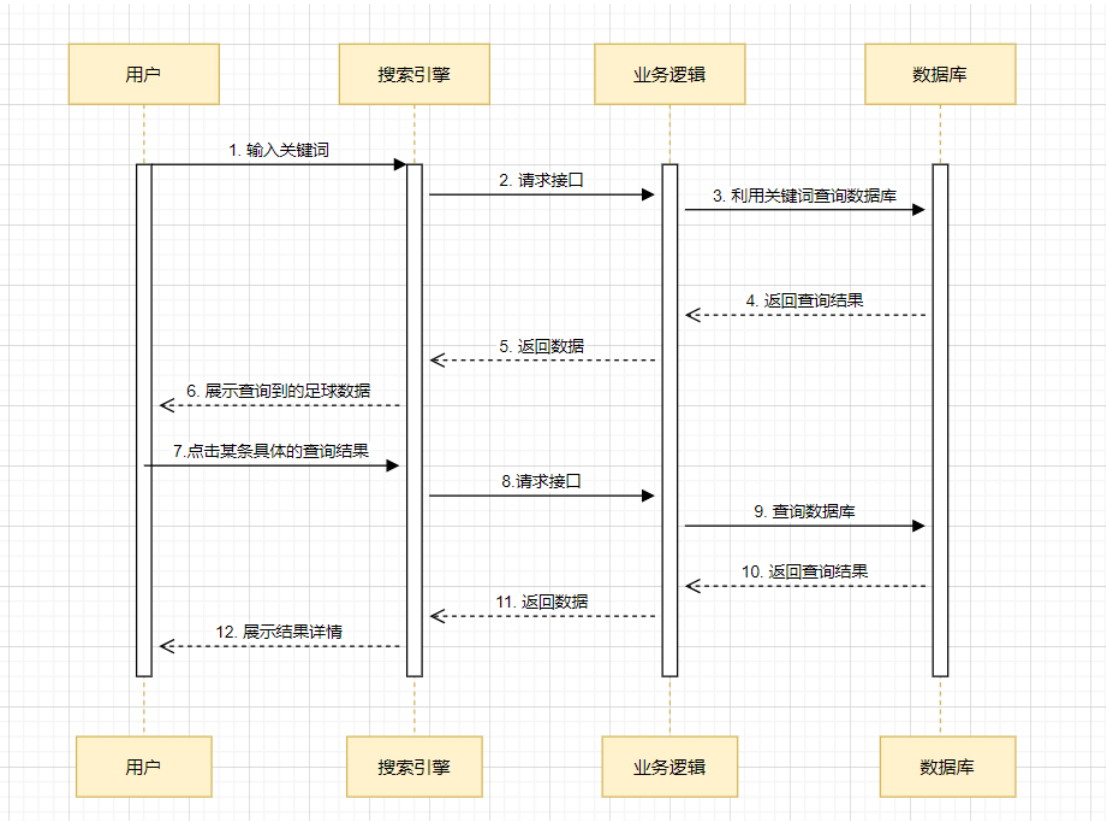


图 4-7-1 结果详情展示时序图

4.8 结果相关度展示

用例编号	SE-08	用例名称	结果相关度排序
创建人	潘凯航	最后修改人	潘凯航
创建日期	2021/07/11	最后修改日期	2021/07/11
角色	用户	需求来源	用户代表
主要参与者	用户		
描述	系统默认按照相关度排序显示搜索结果；用户可以点击相关度按钮，将搜索结果按照相关度排序		
前置条件	1. 系统正常运行 2. 用户可以正常打开相关界面 3. 搜索结果正常		

触发器	用户完成搜索
后置条件	无
输入信息	无
主干过程	1. 用户完成搜索输入 2. 用户点击搜索按钮 3. 用户点击按照相关度排序 4. 结果页面显示排序后的搜索结果卡片列表
分支过程	无
异常	1. 搜索后页面出现异常（发生于主干过程 4） 处理：根据网站的联系方式联系系统管理员
假设	无异常发生
输入	无
输出	按照相关度排序的搜索结果页面
包括用例	无
优先级	中
使用频率	高
备注	无

表 4-8-1 结果相关度展示

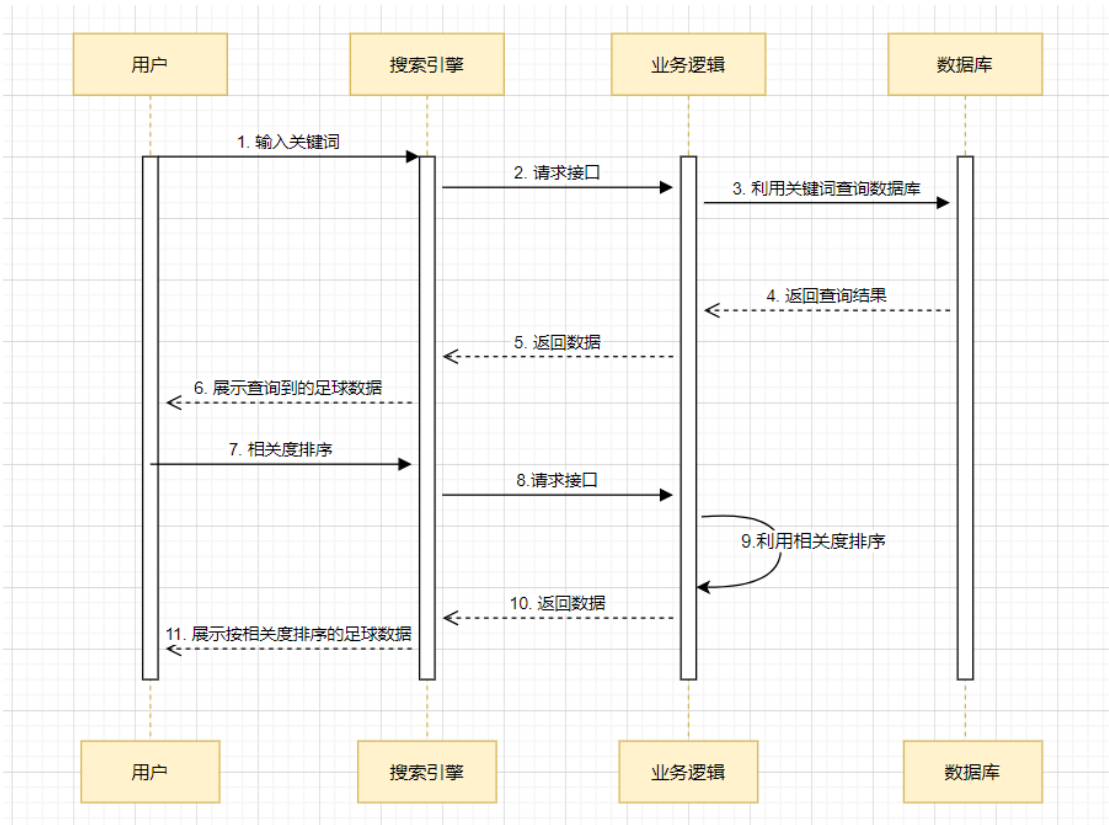


图 4-8-1 结果相关度展示

4.9 结果点击量展示

用例编号	SE-09	用例名称	结果点击量排序
创建人	聂俊哲	最后修改人	聂俊哲
创建日期	2021/07/12	最后修改日期	2021/07/12
角色	用户	需求来源	需求访谈和小组会议
主要参与者	用户		
描述	用户可以点击点击量按钮，将搜索结果按照点击量排序		
前置条件	1. 系统正常运行 2. 用户可正常打开相关界面 3. 搜索结果正常显示		
触发器	用户完成搜索		

后置条件	无
输入信息	无
主干过程	1. 用户完成搜索 2. 用户点击搜索按钮 3. 用户点击按照点击量排序 4. 结果页面显示排序后的搜索结果卡片列表
分支过程	无
异常	1. 搜索后页面出现异常（发生于主干过程 4） 处理：根据网站的联系方式联系系统管理员 系统结束用例
假设	无异常发生
输入	关键词、语句
输出	搜索结果显示页面
包括用例	关键词过滤、关键词交集
优先级	中
使用频率	低
备注	无

表 4-9-1 结果点击量展示

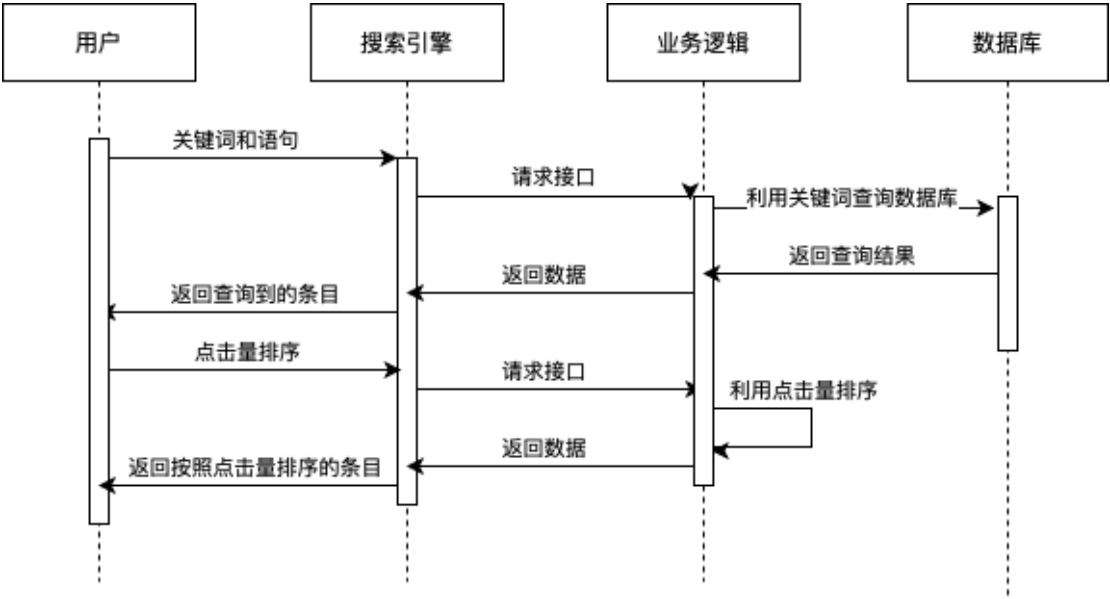


图 4-9-1 结果点击量展示时序图

4.10 相关词条展示

用例编号	SE-10	用例名称	相关词条推荐
创建人	聂俊哲	最后修改人	聂俊哲
创建日期	2021/07/12	最后修改日期	2021/07/12
角色	用户	需求来源	需求访谈和小组会议
主要参与者	用户		
描述	用户完成搜索后，系统自动进行相关词条搜索，并在页面上展示		
前置条件	1. 系统正常运行 2. 用户可正常打开相关界面		
触发器	用户需要完成搜索		
后置条件	无		
输入信息	搜索内容		
主干过程	1. 用户进入网站搜索页 2. 用户输入关键词 3. 用户点击搜索按钮 4. 搜索成功后跳转至结果显示界面		
分支过程	无		
异常	1. 用户输入信息不合法(发生于主干过程 2) 处理:系统提示“查询错误，请重新输入” 系统结束用例 2. 搜索后页面出现异常(发生于主干过程 4) 处理:根据网站的联系方式联系系统管理员 系统结束用例		
假设	无异常发生		

输入	关键词、语句
输出	显示同类词条推荐页面
包括用例	关键词过滤、关键词并集、关键词交集
优先级	高
使用频率	高
备注	无

表 4-10-1 相关词条展示

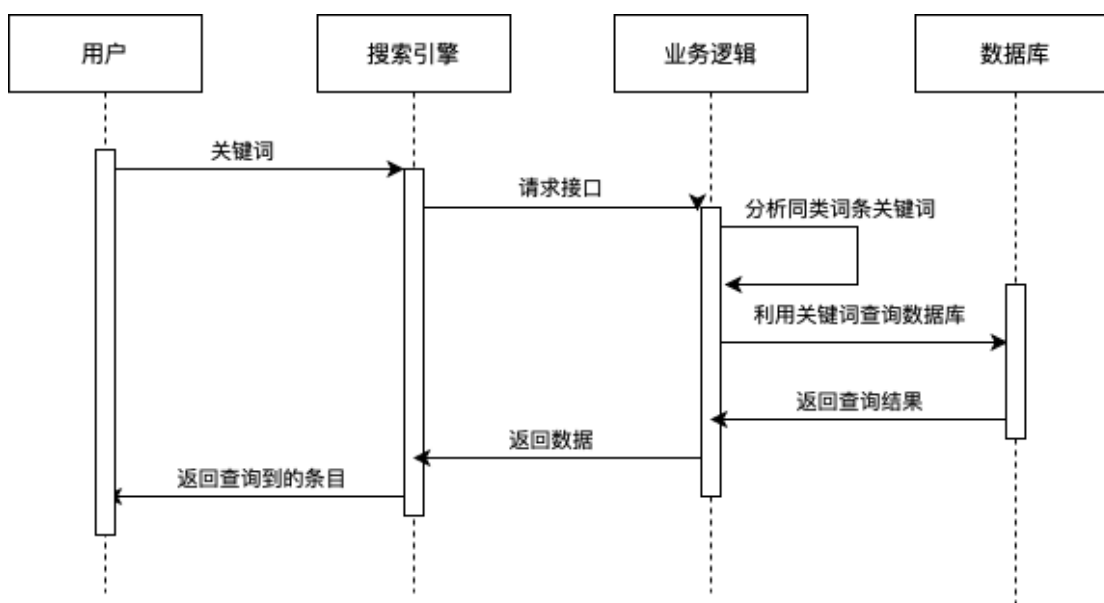


图 4-10-1 相关词条展示时序图

4.11 优先补全

用例编号	SE-11	用例名称	优先补全
创建人	张琦	最后修改人	张琦
创建日期	2021/07/12	最后修改日期	2021/07/12
角色	用户	需求来源	需求访谈和小组会议
主要参与者	用户		
描述	用户在输入部分搜索内容后，会根据输入内容优先补全。		
前置条件	1. 系统正常运行		

	2. 用户可正常打开相关界面
触发器	用户在搜索框中输入内容
后置条件	无
输入信息	部分搜索内容
主干过程	1. 用户进入网站搜索页 2. 用户输入部分关键词 3. 搜索框显示优先补全内容
分支过程	无
异常	无
假设	无异常发生
输入	关键词、语句
输出	显示同类词条推荐页面
包括用例	关键词过滤、关键词并集、关键词交集
优先级	高
使用频率	高
备注	无

表 4-11-1 优先补全

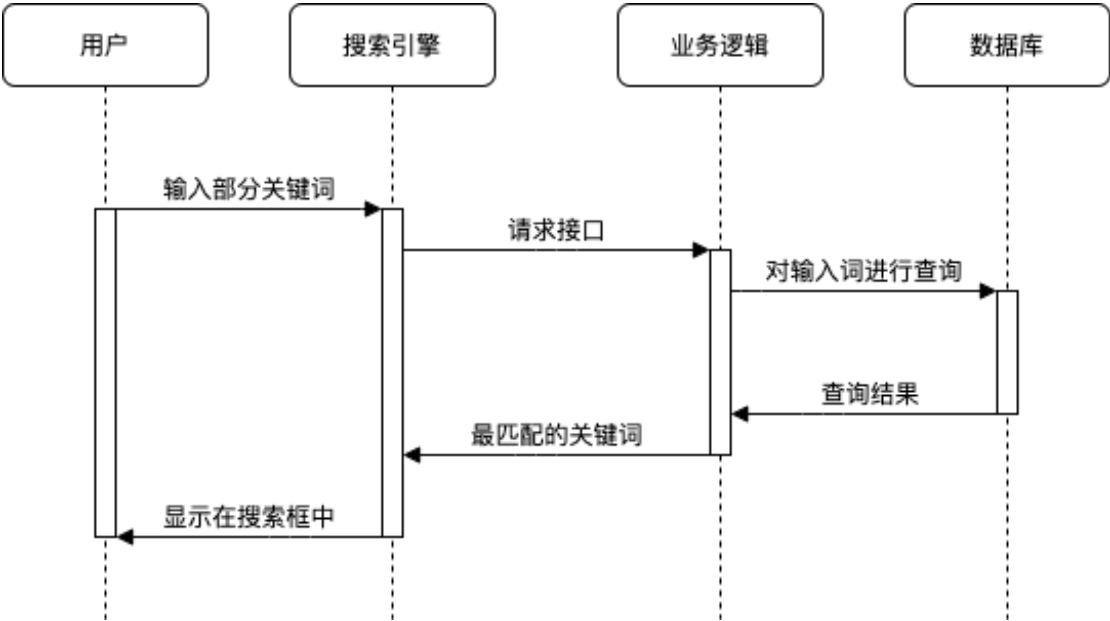


图 4-11-1 优先补全时序图

4.12 搜索提示

用例编号	SE-12	用例名称	搜索提示
创建人	张琦	最后修改人	张琦
创建日期	2021/07/12	最后修改日期	2021/07/12
角色	用户	需求来源	需求访谈和小组会议
主要参与者	用户		
描述	用户在输入部分搜索内容后，会根据输入内容优先补全。		
前置条件	1. 系统正常运行 2. 用户可正常打开相关界面		
触发器	用户在搜索框中输入内容		
后置条件	无		
输入信息	搜索内容		
主干过程	1. 用户进入网站搜索页 2. 用户输入部分关键词 3. 搜索框下方显示搜索提示		
分支过程	用户点击搜索提示直接跳转到搜索对应内容界面		
异常	无		
假设	无异常发生		
输入	关键词、语句		
输出	显示同类词条推荐页面		
包括用例	关键词过滤、关键词并集、关键词交集		
优先级	高		
使用频率	高		
备注	无		

表 4-11-1 优先补全

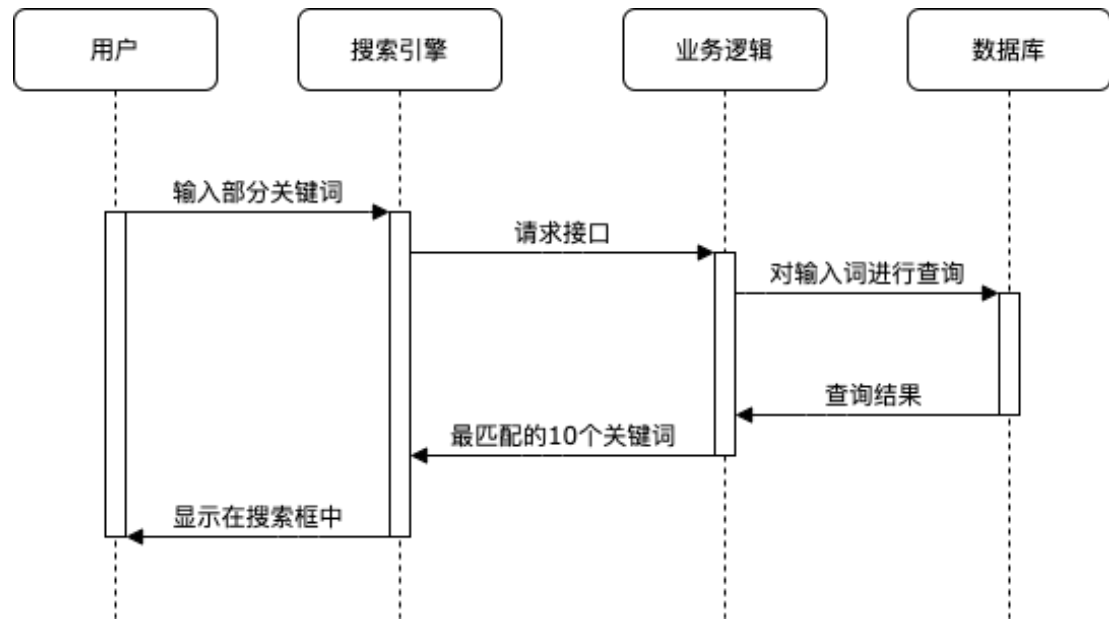


图 4-12-1 搜索提示时序图

4.13 知识图谱问答

用例编号	SE-13	用例名称	知识图谱问答
创建人	张溢弛	最后修改人	张溢弛
创建日期	2021.07.11	最后修改日期	2021.07.11
角色	普通用户	需求来源	需求访谈和小组会议
主要参与者	普通用户		
描述	该功能允许用户使用搜索引擎额外提供的模块进行简单的对话和问答。		
前置条件	1.系统可以正常运行 2.用户拥有一定的电脑操作技术，会打字		
触发器	用户需要进入智能问答页面		
后置条件	用户搜索记录写入后端日志中		

输入信息	用户希望搜索引擎回答的问题
主干过程	1.用户在问答功能对应的输入框中输入想问的问题 2.用户点击问答按钮 3.后端根据建立的足球知识图谱来推理出输入的问题对应的答案，并返回给网站前端 4.前端的答案栏显示出一个简短的问题答案。
分支过程	无
异常	1.用户输入的内容在知识图谱中没有出现 2.后端的知识图谱建立失败，或者推理过程中出现了比较大的误差
假设	无异常发生
输入	用户的问题
输出	系统智能生成的答案
包括用例	无
优先级	中
使用频率	高
备注	中

表 4-13-1 知识图谱问答

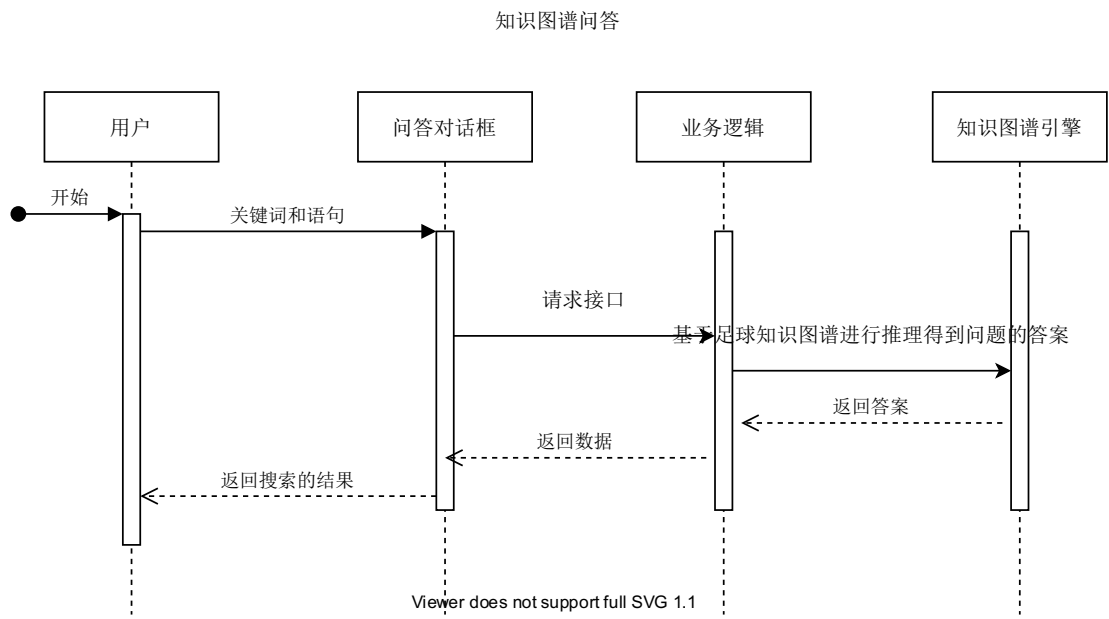


图 4-13-1 知识图谱问答时序图

5 数据流图

5.1 搜索引擎子系统数据流图

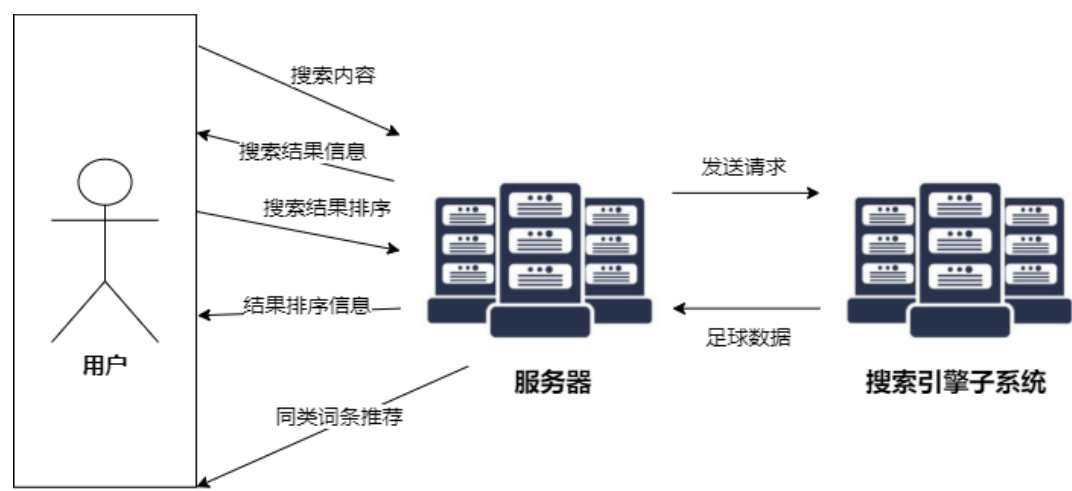


图 5-1-1 搜索引擎子系统数据流图

5.2 网站维护子系统数据流图

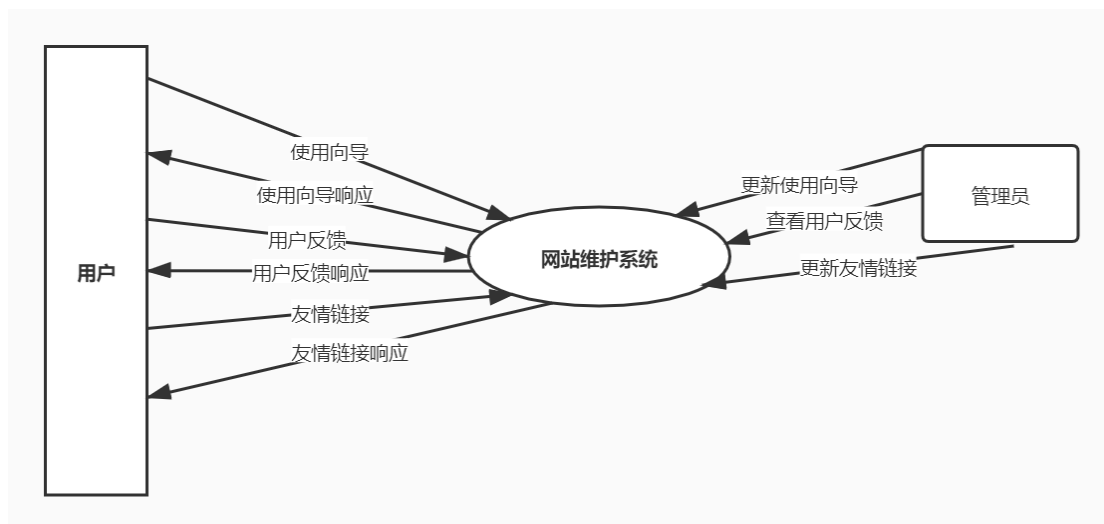


图 5-2-1 网站维护子系统数据流图

6 DM 图



图 6.1 来自搜索信息的对话图

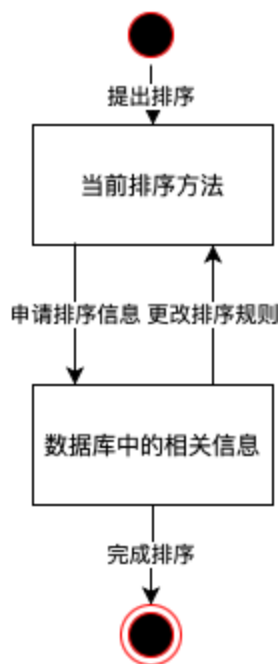


图 6.2 来自排序搜索结果的对话图

7 ER 图

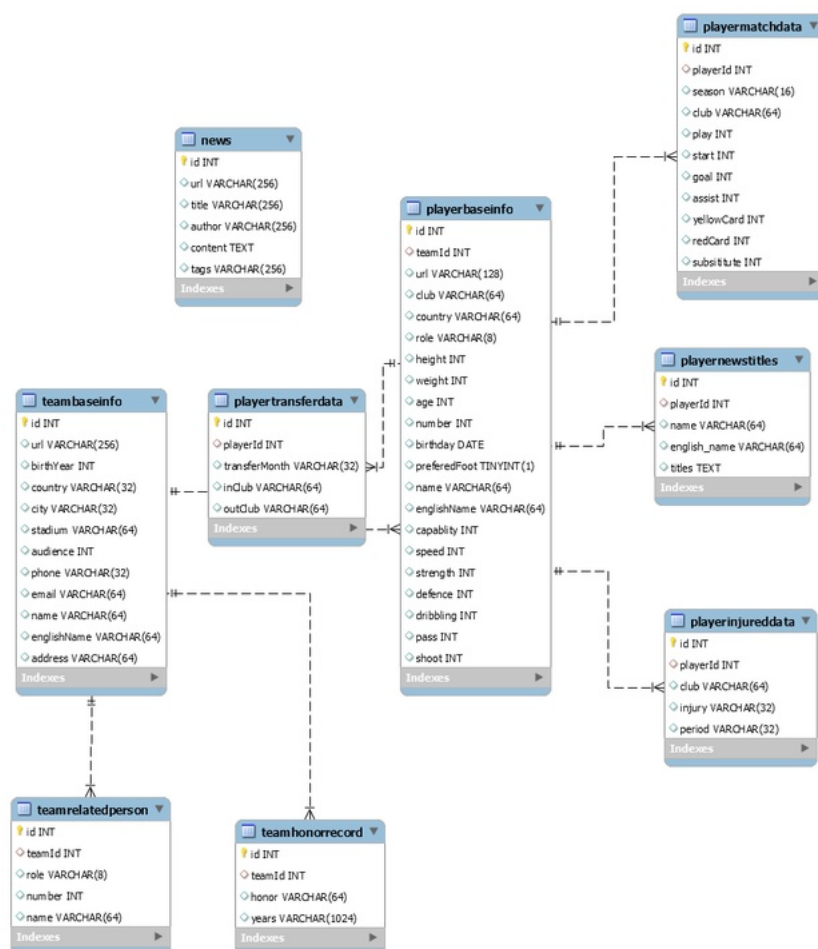


图 7-1 ER 图

8 外部接口需求

8.1 服务器接口

- 服务器硬件接口需求

项目	信息
服务器设备	阿里云服务器
处理器	Intel i3 芯片及以上
内存	4G 以上
存储	1T 及以上
网卡	速率 10Mbps 以上
备份	RAID 5

表格 8-1-1 服务器硬件接口需求

- 服务器软件接口需求

项目	信息
操作系统	Linux Ubuntu 18.04
服务器	Nginx
数据库	MySQL

表格 8-1-2 服务器软件接口需求

8.2 客户端接口

- 客户端硬件需求

项目	信息
处理器	Intel i5 及以上或者 AMD
内存	2G 以上

存储	任意主流的存储介质
网卡	速率 10Mbps 以上，支持校园网 ZJUWLAN
备份	RAID 5

表格 8-2-1 客户端硬件需求

● 客户端软件需求

项目	信息
操作系统	Windows 10、Linux Ubuntu 18.04
浏览器	主流浏览器，对于 IE 需要版本大于 IE6

表格 8-2-2 客户端软件需求

8.3 通信接口

本系统使用的通信网络环境采用公共网络或者浙江大学校园网 ZJUWLAN，有条件的时候可以扩展到。

9 非功能性需求

9.1 性能需求

- 系统应保证运行稳定，避免出现崩溃；
- 主流浏览器均能正常访问本系统；
- 系统应能保证至少 100 人的并发访问；
- 当用户登录以及进行任何操作时，系统应该能及时进行反应，反应的时间在 1s 以内；
- 系统应该能及时检测出各种非正常情况，如与设备的通信中断断开，无法连接数据库服务器等情况，避免用户长时间等待；
- 每个页面一般情况下应在 1s 内加载完毕，高峰期应在 5s 内加载完毕；
- 系统保证在一周内不超过一次维护与重启。

9.2 数据传输与并发要求

- 用户输入账号密码点击登录后，对登录的相应时间不能超过 1 秒，在此时间内将登录结果显示在屏幕上；
- 系统能支持 10 名用户同时搜索球员和查看球员信息
- 系统能支持 30 名用户同时搜索足球新闻
- 系统能支持 20 名用户同时进行足球信息的条件查询
- 系统系统应支持 100 名用户并发使用，并保证性能不受影响；
- 在网页中，系统生成的所有 Web 页面可以在不超过 4 秒的时间内可以全部下载下来。

9.3 数据管理要求

系统既要与其他系统有接口，又必须保证本系统的独立性与完整性。即应防止未经授权的各类人员对本系统进行设置和修改或访问系统内部数据。

系统服务器软件必须提供可靠的数据备份和恢复手段，在服务器软件或硬件出现严重故障时，能够根据备份的数据和账户信息迅速恢复正常运行环境。同时，软件开发者不得在系统中预留任何特殊账户和密码，保证其安全性。

除此之外，系统应具备加密登录、数据加密传输等安全方面的保障，保证数据在不用系统间传输过程中的保密性与安全性。

以下为具体细则：

- 当系统崩溃后，系统应能在 24 小时内恢复运行；
- 数据库可支持表的最大行数达到 600 行；
- 本系统用于日志等记录的数据增长约为 20MB/月，具体增长速度由用户的使用频率及所发生业务的数据量决定；
- 当出现重大事故造成数据丢失后，系统应能在 48 小时内恢复数据；
- 系统管理员每两个月应至少维护备份一次数据；
- 系统服务器应具备至少 20GB 的存储空间；

9.4 权限与安全需求

在我们的系统中，对于安全与权限进行了如下设计：

- 所有涉及功能信息或个人信息的网络事务，都应进行加密操作

- 未登录的用户，即游客仅能查看课程部分信息与留言讨论版
- 只有系统管理员有权查看及修改底层数据库数据，且行为应被系统日志记录，用户无法非法修改数据库
- 教师拥有对助教拥有权限的设置权限
- 当流量过大时，优先限制游客流量防止恶意访问
- 对可能造成等待时间较长的操作应该提供取消功能
- 对错误操作支持可逆性处理，如取消系列操作

9.5 输入要求

系统对用户输入的结果要进行合理的解析和处理，同时过滤掉一些错误的、非法的输入，同时要防止 SQL 注入攻击。

此外，系统应通过程序控制出错几率，减少系统因用户人为的错误引起的破坏，开发者应当尽量周全地考虑到各种可能发生的问题，使出错的可能降至最小。

9.6 可视化需求

- 要将搜索结果进行良好的可视化呈现并进行一定的排序
- 要将球员的个人信息和比赛数据使用美观简洁的表格、统计图等形式呈现
- 要将球队和比赛信息用良好的交互界面进行可视化呈现
- 要将足球新闻用良好的可视化界面呈现

9.7 防护性需求

- 数据库误删除时，可以使用撤销删除修复；
- 系统应该保护未开放下载权限的资料不被下载；
- 重复操作导致卡死时，系统提出警告；
- 系统应该提供验证码防止恶意使用(如爬虫等)；
- 系统应该及时信息备份防止病毒攻击；
- 系统应该能检测到恶意操作，提出警告并在一段时间内不允许操作。
- 访问无权限时，系统发出提示并禁止用户访问。

9.8 软件质量属性

- 兼容性：系统只需要保证在主流浏览器（Chrome、IE）上正常浏览和使用即可
- 易用性：系统界面应该简洁明了、操作简单，功能按钮的位置符合用户的日常习惯。此外，系统应该要有导航和清晰简短的用户使用手册；
- 可用性：系统保证早上 6 点到晚上 24 点之间可用，但在发生紧急情况时允许维护一段时间；
- 可维护性：系统运行时要保存运行日志，用来维护分析。每两周的维护日凌晨 1 点到 5 点为维护时间，在此期间用户不能使用系统。
- 可扩充性：系统在设计上考虑到了网站可能的后续发展，在后端设计和前端设计上尽可能地在满足所有需求的同时，增强了网站的可扩充性。一旦有扩展需要，客户可以联系系统维护人员，维护人员需要在 1-4 个工作日内完成客户的内容扩充需求，主要包括增加新的功能、增加新的模块、界面优化、系统性能提升等。

9.9 其他需求

- 当用户进行误操作或不合法操作，软件能给出提示信息。
- 软件对用户的使用应该有较为清晰的引导。

10 数据字典

10.1 数据元素定义表

编号	数据元素名	类型	值域	说明
E1	查询字符串	字符	/	用户在搜索引擎输入框中输入的查询字符串
E2	球员名称	字符	/	/

E3	球员国籍	字符	/	/
E4	球员俱乐部	字符	/	/
E5	球员位置	字符	/	/
E6	球员号码	字符	/	/
E7	球员年龄	整型	0-100	/
E8	球员身高	整型	0-2.5	/
E9	球员体重	整型	0-150	/
E10	球员能力值	浮点型	0.0-10.0	/
E11	球员比赛数据	字符	/	/
E12	球员荣誉记录	字符	/	/
E13	球员转会记录	字符	/	/
E14	球员伤病记录	字符	/	/
E15	球队名称	字符	/	/
E16	球队成立时间	时间	1900 -2021	/
E17	球队国家	字符	/	/
E18	球队荣誉记录	字符	/	/
E19	球队赛程	字符	/	/

表 10-1 数据元素定义表

10.2 数据精确表

编号	数据元素名	类型	精 度 要 求	说明	示例
E1	查询字符串	字符	128 个字 符以内	用户在搜索引擎输入框中输入的查询字符串	梅西

E2	球员名称	字符	64 个字符以内	/	里卡多-洛佩斯
E3	球员国籍	字符	32 个字符以内	/	巴西
E4	球员俱乐部	字符	32 个字符以内	/	上海海港
E5	球员位置	字符	8 个字符以内	/	前锋
E6	球员号码	字符	8 个字符以内	/	10 号
E7	球员年龄	整型	0-100	/	30 岁
E8	球员身高	整型	0-1000	/	186 cm
E9	球员体重	整型	0-1000	/	77 KG
E10	球员能力值	浮点型	0.0-10.0	/	8.2
E11	球员比赛数据	字符	128 个字符以内	/	赛季：2021 俱乐部：上海海港上场：5 首发：5 进球：4 助攻：1 黄牌：0 红牌：0 替补：0
E12	球员荣誉记录	字符	/	/	亚洲足球冠军联赛 X 1
E13	球员转会记录	字符	/	/	2020-02-06 全北现代 - 上海上港 550 万欧元
E14	球员伤病记录	字符	/	/	膝盖前十字韧带受伤
E15	球队名称	字符	/	/	山东泰山

E16	球队成立时间	时间	1900 - 2021	/	1993
E17	球队国家	字符	/	/	中国
E18	球队荣誉记录	字符	/	/	中国足球冠军联赛 X 1
E19	球队赛程	字符	/	/	06-27 18:00 中超第 7 轮 沧州雄狮 山东泰山

表 10-2 数据精确表

11 业务规则与业务算法

11.1 业务规则

标识符	规则定义	规则类型	静态或动态	来源
BR-01	用户仅能搜索相关球员、球队的信息	约束	静态	数据限制
BR-02	用户仅能通过页面显示的搜索范围类型进行高级搜索	约束	静态	管理策略
BR-03	用户搜索关键字支持与或非逻辑运算	动作处罚规则	动态	用户
BR-04	用户搜索时语言仅限于中文(包括拼音)以及英文	约束	静态	管理策略
BR-05	用户在搜索得出的结果列表中需要点击方可进入详情页面	动作触发规则	动态	用户
BR-06	搜索结果列表默认按照搜索的相关度排序	约束	静态	管理策略
BR-07	搜索框默认显示用户最近两次搜索的词条	约束	静态	管理策略

BR-08	知识图谱问答时输入必须和足球知识有关	约束	静态	用户
-------	--------------------	----	----	----

表 11-1 业务规则

11.2 业务算法

针对用户利用关键字的与或非逻辑运算来搜索的业务规则，“与”运算的算法应满足只有相“与”的关键字全部出现时，所检索到的结果才算符合条件;“或”运算的算法应满足只要相“或”的关键字中有任何一个出现，所检索到的结果均算符合条件;“非”运算的算法应满足搜索结果中不应含有“非”后面的任何关键字。

12 界面原型



图 12-1 搜索主页

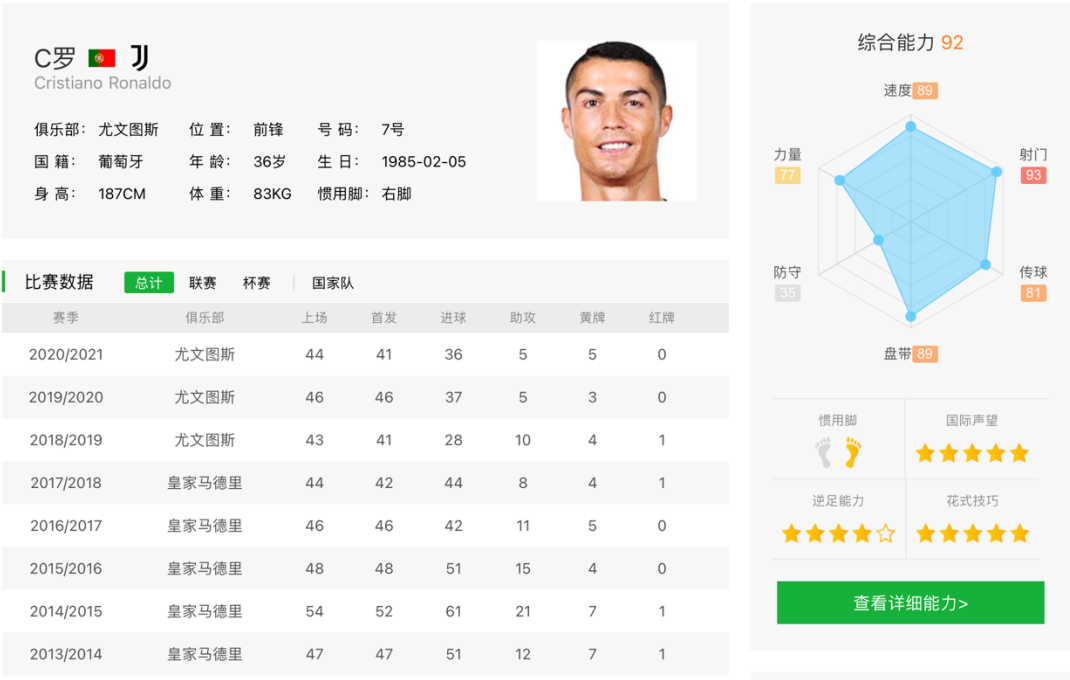


图 12-2 球员信息展示



图 12-3 球队信息展示

13 上下文图

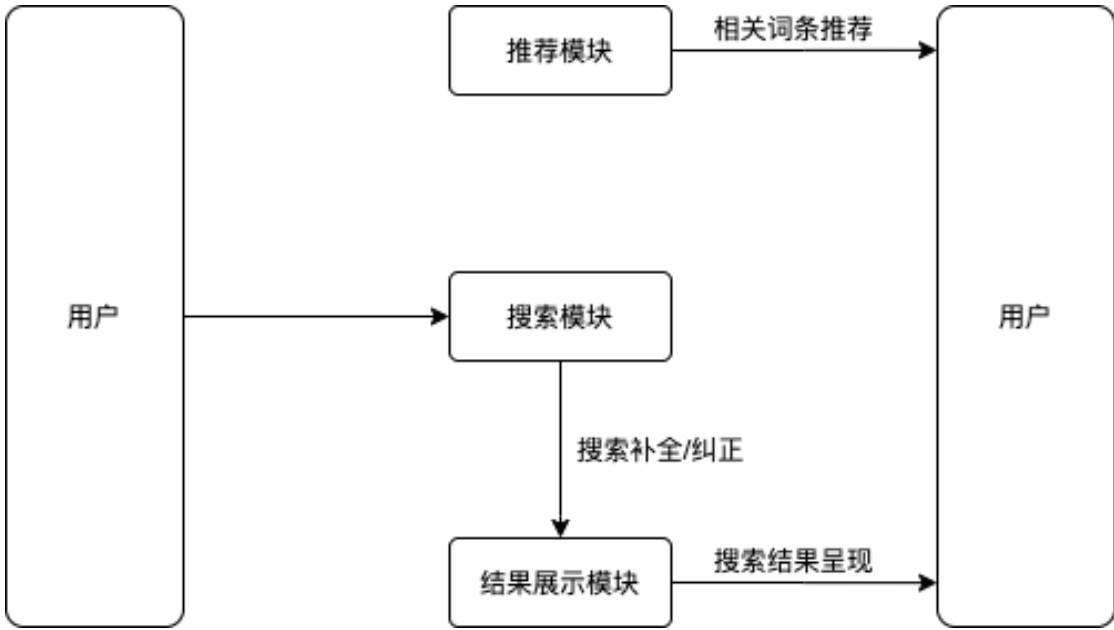


图 13-1 上下文图

14 附录

14.1 术语表

名词	解释
软件	软件是一系列按照特定顺序组织的计算机数据和指令的集合。
软件工程	软件工程是将系统化的、严格约束的、可量化的方法应用于软件的开发、运行和维护，即将工程化应用于软件；同时对上述方法所进行的研究。
软件生存周期	软件生存周期是软件的产生直到报废的生命周期，周期内有问题定义、可行性分析、总体描述、系统设计、编码、调试和测试、验收与运行、维护升级到废弃等阶段。
软件质量	软件与明确的和隐含的定义的需求相一致的程度。
质量认证	质量认证也叫合格评定，是国际上通行的管理产品质量的有效方法。
软件过程	软件过程为一个为建造高质量软件所需完成的任务的框架，即形成软件产品的一系列步骤，包括中间产品、资源、角色及过程中采取的方法、工具等范畴。

软件需求	<p>软件需求是：</p> <ol style="list-style-type: none"> 1 用户解决问题或达到目标所需条件或权能(Capability); 2 系统或系统部件要满足合同、标准、规范或其它正式规定文档所需具有 的条件或权能; 3 一种反映上面 1 或 2 所述条件或权能的文档说明。它包括功能性需求及 非功能性需求,非功能性需求对设计和实现提出了限制, 比如性能要求、 质量标准、或者设计限制。
业务需求	<p>业务需求(Business Requirements)反映了组织机构或客户对系统或产品 高层次的目标要求, 它们在项目视图与范围文档中予以说明。</p>
用户需求	<p>用户需求(User Requirements)描述了用户使用产品必须要完成的任务, 可以在用例模型或方案脚本中予以说明。</p>
功能需求	<p>功能需求(Functional Requirements)定义了开发人员必须实现的软件功能, 使得用户能完成他们的任务, 从而满足了业务需求。</p>
非功能需求	<p>非功能需求(Non-functional Requirements)是从各个角度对系统的约束和限制, 反映了应用对软件系统质量和软件需求规格说明书, 用于反映高校教学平台的额外要求。</p>
需求工程	<p>需求工程是指应用已证实有效的技术、方法进行需求分析, 确定客户需求, 帮助分析人员理解问题并定义目标系统的所有外部特征的一门学科。 它通过合适的工具和记号系统地描述待开发系统及其行为特征和相关约束, 形成需求文档, 并对用户不断变化的需求演进给予支持。</p>
时序图	<p>时序图(Sequence Diagram)又名序列图、循序图, 是一种 UML 交图。它通过描述对象之间发送消息的时间顺序显示多个对象之间的动态协作。 它可以表示用例的行为顺序, 当执行一个用例行为时, 其中的每条消息对应一个类操作或状态机中引起转换的触发事件。</p>
项目管理	<p>项目管理是通过合理地组织和利用一切可以利用的资源, 按照计划的成本和计划的进度, 完成一个计划的目标, 它包含</p>

	团队管理、风险管理、采购管理、流程管理、时间管理、成本管理和质量管理等。
UML	UML(Unified Modeling Language)又称统一建模语言或标准建模语言，是始于 1997 年一个 OMG 标准，它是一个支持模型化和软件系统开发的图形化语言，为软件开发的所有阶段提供模型化和可视化支持。

14.2 待确定问题清单

暂无