

律

镜

2023-12-3

需求说明书

AI + 法律领域垂直搜索引擎

《软件工程管理》G04 小组

组长：徐毕颖

组员：章越、沈书豪、黄文杰、王梓轩



目录

1. 引言	3
1.1 编写目的	3
1.2 软件项目背景	3
1.3 目标读者和阅读建议	3
1.3.1 组成部分	3
1.3.2 读者对象	4
1.4 参考文献与资料	5
2. 总体描述	5
2.1 产品基本信息	5
2.2 产品前景	6
2.3 用户类及其特征	7
2.4 产品功能概述	7
2.4.1 搜索词条模块	7
2.4.2 搜索结果模块	8
2.4.3 历史搜索模块	8
2.4.4 详情展示模块	8
2.4.5 AI 辅助模块	9
2.4.6 知识图谱模块	9
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	10
2.7 开发、测试和运行环境	11
2.7.1 软件运行环境	11
2.7.2 硬件运行环境	12
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	13
2.9 产品开发项目设计人员	13
3. 系统需求分析概述	14
3.1 系统总用例图	14
3.2 需求优先级分析	15
3.3 系统模块架构图	16
4. 功能需求	16
4.1 关键词搜索	16
4.2 搜索类型选择	18
4.3 高级搜索	20
4.4 搜索结果排序	22
4.5 结果摘要	24
4.6 用户反馈	25
4.7 搜索历史显示	27
4.8 个性化推荐	29
4.9 法律文档展示	30
4.10 相关文档	32

4.11 关键词提取	34
4.12 文档总结	35
4.13 关键词图	37
4.14 图谱转换与跳转搜索	39
5. 数据流图 & 数据字典	41
5.1 数据流图	41
5.2 数据字典	41
5.2.1 数据元素定义表	41
5.2.2 数据精确表	42
6. 上下文图	44
7. DM 图	44
8. ER 图	45
9. 外部接口需求	45
9.1 服务器接口	45
9.2 客户端接口	46
9.3 通信接口	47
10. 非功能性需求	47
10.1 性能需求	47
10.2 数据传输与并发要求	48
10.3 数据管理要求	49
10.4 输入要求	50
10.5 可视化需求	51
10.6 防护性需求	52
10.7 软件质量属性	53
10.8 其他需求	55
11. 业务规则与业务算法	56
11.1 业务规则	56
11.2 业务算法	57
12. 界面原型	57
12.1 搜索引擎首页	57
12.2 搜索结果列表页	58
12.3 详情页	58
13. 附录	58
13.1 术语表	59
13.2 待确定问题清单	59

软件需求规格书

1. 引言

1.1 编写目的

本文档的编写目的是明确记录法律垂直搜索引擎项目的功能和性能需求，以提供项目团队、开发人员和相关利益方全面的项目理解。通过规范的需求描述，旨在确保在软件开发生命周期中的清晰沟通，为开发团队提供准确的指导，从而最大程度地满足用户需求，保证项目的成功实施。本文档将成为项目团队沟通、决策和验证的基础，以确保软件系统的一致性、可追踪性和可管理性。

1.2 软件项目背景

法律垂直搜索引擎项目背景源于对法律信息获取存在局限性的认识，以及对提高法律专业人士和公众获取法律信息效率的强烈需求。当前，法律领域的信息分散在不同的数据库、文献和在线资源中，使得检索法律信息变得繁琐且不高效率。本项目旨在通过开发一款专注于法律领域的搜索引擎，集成并优化法律信息的检索，为用户提供便捷、精准的法律信息查询服务。通过这一项目，我们追求为法律专业人士、研究人员和公众提供更智能、高效的法律信息获取体验，推动法律领域信息化水平的提升。

1.3 目标读者和阅读建议

1.3.1 组成部分

本软件需求规格说明书主要包括以下几个部分，每个部分旨在提供对法律垂直搜索引擎项目的详细描述和规范：

i. 总体描述部分：

- 描述法律垂直搜索引擎项目的基本信息，包括目标、范围和背景。
- 突出项目的前景，解释项目的重要性和预期影响。
- 识别用户类特征，包括法律专业人士、研究人员和公众，以确保项目满足不同用户群体的需求。
- 介绍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。

ii. 系统需求分析概述部分：

- 进行对法律垂直搜索引擎系统主要用例的需求优先级分析。

- 提供系统的模块架构图，清晰展示各个模块之间的关系。
- 在功能需求部分，根据不同用户类型进行详细的需求分析，并通过用例时序图展示功能交互。
- 强调本软件需求规格说明书的核心部分。

iii. 数据库 ER 图部分：

- 提供法律垂直搜索引擎项目最基本的数据库设计，包括各个数据表的关系和结构。

iv. 数据流程图部分：

- 图形化表达每个子系统的逻辑流动方向，以便更好地理解系统的数据流动。

v. 外部接口需求部分：

- 描述法律垂直搜索引擎与外部硬件、软件和通信接口的交互关系，确保系统的顺利集成。

vi. 非功能性需求部分：

- 详细描述法律垂直搜索引擎项目的性能、安全性、可靠性等方面的非功能性需求，以确保系统的整体质量。

vii. 数据字典部分：

- 提供系统的数据流定义、数据元素定义和数据精度等详细信息，以确保数据的准确性和一致性。

viii. 业务规则与业务算法部分：

- 列举法律垂直搜索引擎项目的操作规则和算法，以确保系统的操作一致性和效率。

每个部分都将为项目的不同方面提供详细的规范和指导，以促进团队之间的沟通和开发的顺利进行。

1.3.2 读者对象

读者对象及阅读建议：

- i. 项目经理：项目经理可通过本文档深入了解法律垂直搜索引擎项目的功能需求，为系统设计和项目管理提供指导。关注总体描述和系统需求分析概述，以全面理解项目的核心特征。
- ii. 设计人员：设计人员可依据本文档对法律垂直搜索引擎的需求进行分析，进而进行系统设计，包括系统架构和数据库表的设计。重点关注系统需求分析概述和数据库 ER 图，以获取详细的功能和数据要求。
- iii. 程序开发人员：开发人员可根据需求规格说明书和系统总体设计计划了解系统功能，推动系统的开发和用户手册的编写。详细研读功能需求和系统架构设计，以确保代码开发符合预期。
- iv. 销售人员：销售人员可根据本文档了解法律垂直搜索引擎的各项功能和性能，从而有效推广产品。主要参考总体描述和系统需求分析概述，以获取对产品核心卖点的了解。

- v. 测试人员：测试人员可基于本文档编写详尽的测试用例，并进行功能性和非功能性测试。重点关注功能需求、非功能性需求和数据流图，以确保系统质量和性能。
- vi. 用户：最终用户可以根据本文档了解法律垂直搜索引擎的功能和性质，并与分析人员进行需求的讨论和协商。主要关注总体描述和系统需求分析概述，以获取对产品的整体认知。
- vii. 其他相关人员：部门领导、公司领导、政府机关领导等领导可以通过本文档了解产品的功能和性能。主要参考总体描述和系统需求分析概述，以获取对项目的高层次了解。

阅读建议：

在阅读本软件需求规格说明书时，建议首先了解产品的整体功能概貌，随后根据具体需求逐一了解每个功能的详细描述。对于不同读者对象，建议重点关注与其职责和关注领域相关的部分，以获取最有益的信息。

1.4 参考文献与资料

《软件工程开发国家标准》
《软件工程项目开发文档范例》
《软件需求（第三版）》
《软件工程——实践者的研究方法》

2. 总体描述

2.1 产品基本信息

- 产品名称：“律镜”法律领域垂直搜索引擎
- 任务提出者：本项目由浙江大学项目实训课程组提出，任课教师为邵健与邹强。
- 开发者：“律镜”法律领域垂直搜索引擎项目由浙江大学项目实训 G04 小组负责开发，项目组成员将共同努力完成项目的设计、开发和测试。
- 用户：该系统的用户主要是法律从业人员、学生、研究者和一般公众，他们有检索法律信息的需求。
- 功能：“律镜”法律领域垂直搜索引擎旨在满足用户对法律信息的检索需求。用户可以通过关键字检索法律文书、法规、案例，以及与之相关的法律新闻资讯。搜索功能将支持高级搜索、模糊搜索、拼音搜索等多项功能，提供全面而灵活的检索体验。
- 性能：系统设计目标是能够支持至少1000名用户的使用请求，以保证系统在高负载时的稳定性和性能。
- 完成期限：初步开发计划在12月21日前完成，最终系统计划在12月25日前完成。这个时间表充分考虑了项目的复杂性和各个阶段的任务，以确保项目按时推进。

2.2 产品前景

法律领域垂直搜索引擎在法律信息检索领域具有广泛的前景和应用潜力。随着社会的法治化程度提高，法律领域的信息量不断增加，法律从业人员和普通公众对快速获取准确法律信息的需求日益迫切。本项目将为用户提供高效、智能的法律信息搜索服务，有望在以下几个方面展现显著的前景：

- a. 高效检索：法律领域垂直搜索引擎将通过多种搜索方式，包括高级搜索、模糊搜索和拼音搜索，提高用户对法律信息的检索效率。用户能够迅速准确地定位所需信息，节省大量检索时间。
- b. 全面覆盖：项目旨在涵盖法律文书、法规、案例等多方面的法律信息，满足用户对不同类型法律资料的全面检索需求。全面的覆盖将使该搜索引擎成为法律从业人员和法学研究者的重要工具。
- c. 智能推荐：未来系统有望通过智能算法，根据用户的检索历史和偏好，向用户推荐相关性更高的法律信息，提升用户体验，使检索更加智能化。
- d. 用户群体广泛：面向法律从业人员、学生、研究者以及普通公众，法律领域垂直搜索引擎将服务于广泛的用户群体，推动法律信息化的普及与应用。
- e. 未来拓展：本项目具备在未来不断拓展的潜力，可以引入更多智能化功能，例如自然语言处理、知识图谱等，提升系统的智能程度和全面性。

在法律信息爆炸性增长的趋势下，法律领域垂直搜索引擎将成为法律领域信息化的重要推动力，为用户提供高效、全面、智能的法律信息检索服务，进一步促进法律领域的数字化转型。

2.3 用户类及其特征

用户群体	特征	需求
法律专业人员	法官、律师、法务人员等专业从业人员，对于法律信息的准确性和全面性有高要求。	主要关注法律文书、案例法规的检索，需要高效获取最新的法律信息，以支持法律案件的研究和决策。
法学研究者和学生	法学院教授、研究生、本科生等法学学术界人员。	对法学理论和法律研究资料的需求较大，期望系统提供全面而深入的法学知识，以支持学术研究和教学。
公众用户	普通公众中对法律事务感兴趣或有特定法律需求的个体。	主要涉及法律常识、法律新闻、法律咨询等，期望系统提供易懂的法律信息，帮助他们更好地理解法律事件和法规。
研究人员和智库机构	法律研究机构、智库研究人员等专业机构和人员。	对于法律研究的深度和广度要求较高，期望系统提供详实的法律文献、统计数据等支持研究工作。

2.4 产品功能概述

2.4.1 搜索词条模块

搜索词条模块是法律领域垂直搜索引擎的核心功能之一。该模块允许用户输入关键词或短语，从而对系统后端数据库中的文档进行搜索。以下是该模块的功能描述：

i. 关键词搜索：

用户可以在搜索框中输入一个或多个关键词或短语，或者一段长文本，用于搜索指定的内容。

ii. 搜索类型选择：

用户可以选择搜索的类型，例如全文搜索、标题搜索、文档类型搜索、作者搜索等。

iii. 高级搜索：

用户可以对搜索的方式进行一定的限定，如限制时间范围、进行模糊匹配等，来获取更精确、符合需求的搜索结果。除此之外，用户还可以根据需要的搜索结果类型（如法律法规、裁判文书、案件描述）指定需要的搜索结果。

2.4.2 搜索结果模块

在用户输入关键词并点击搜索按钮后，系统会基于用户的关键词，给出相应的搜索结果。以下是该模块的功能描述：

i. 搜索结果排序：

搜索引擎可以根据相关性对搜索结果进行排序，基于用户的关键词、搜索方式等因素为结果设置不同的权重，将最相关的结果显示在前面。

ii. 结果摘要：

对于每个搜索结果，系统可以生成一个摘要，包含相关文档的片段，以使用户快速了解文档内容。

iii. 用户反馈：

用户可以对搜索结果进行反馈（点赞点踩等），规定只有登录的用户才能点赞/点踩，防止恶意刷赞/踩。

2.4.3 历史搜索模块

i. 搜索历史显示

用户通过搜索引擎进行的法律领域的各种查询会被记录下来，形成用户的搜索历史，以便重新访问之前查找的法律信息，省去了重新输入搜索词的步骤。

ii. 个性化推荐

个性化推荐功能通过分析用户的搜索历史、点击行为、喜好等信息，为用户提供定制化的推荐内容。根据用户的兴趣和需求，搜索引擎会推荐相关的法律文档、案例、法规等内容，以满足用户更深层次的信息需求。

2.4.4 详情展示模块

内容展示模块是法律领域垂直搜索引擎的关键部分。该模块用于展示搜索结果中的具体法律文档内容，以使用户查看详细信以下是该模块的功能描述：

i. 法律文档展示：

根据用户选中的搜索结果，系统可以展示具体的法律文档内容，并以合适的字体、排版来显示文档的标题、作者、导语、正文等部分的内容，使文档更易阅读和理解。

ii. 文档导航*：

（增强功能）对于长篇法律文档，系统可以根据段落和小标题提供导航功能，方便用户快速浏览和定位到感兴趣的部分。可能有一定实现难度，在后期考虑实现。

iii. 相关文档推荐：

在当前文档的最后，搜索引擎会基于该文档的内容与类型，提供相关性较高的其他文档的链接，供用户快速查阅到相关的内容。

2.4.5 AI 辅助模块

本系统引入了人工智能技术赋能搜索，该模块负责AI模型接口的调用以及系统和模型的对接。以下是该模块的功能描述：

i. 关键字提取：

通过调用相关 API，对用户输入的查询进行分析，提取其中的关键字，用于优化搜索算法，提高搜索结果的相关性，从而更好地满足用户的信息需求。

ii. 文档总结：

在阅读文档详情时，可以利用 AI 对搜索结果文档进行智能总结，以提供用户更为精炼和易理解的信息，生成简明扼要的摘要，有助于用户快速了解文档内容，减少阅读时间，提高信息获取效率。

2.4.6 知识图谱模块

i. 关键词图

对于任何一次搜索，可以生成它知识图谱形式的关键词图：提取得到它的1-2个关键词，作为图的中心词，然后返回其深度1-2的其他关键词节点，组成一张美观的词图。

ii. 图谱转换与跳转搜索

点击词图中的各节点，可以将其转换为中心词（提供动态转换效果），并且用户可以选择是否以该节点上的词为输入值，再次搜索。

2.5 设计和实现上的约束

系统的设计、编码以及维护将严格遵照文档的描述来进行。

此外，在具体的设计与实现上，团队也将按照以下的各项约束进行。

2.5.1 数据存储

本系统将使用MySQL数据库系统作为数据存储引擎，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，本系统也将采用Elastic Search作为搜索引擎框架，会对爬取的数据进行一定的格式化处理后，存储进Elastic Search引擎中，供搜索引擎进行检索时使用，因此也会起到数据存储的作用。

2.5.2 网络服务吞吐

本系统需要保证在高负载时的稳定性和性能，至少应当能够支持1000名用户的并发访问，同时系统正常的响应时间应当不超过2s，在系统繁忙时的响应时间应当不超过5s。

2.5.3 数据安全

本系统需要从完整性、保密性、可用性三个角度，保障用户数据的安全。

- 数据完整性是指存储在数据库中的数据正确无误并且相关数据具有一致性，保证数据在未解授权和校验的情况下无法被修改，且所有的数据在系统运行时不会被篡改、盗用和丢失。对此，系统将在后端与数据库中对数据采取一定的校验手段，限制数据的输入和修改，防止数据的一致性被破坏。
- 数据保密性是指系统中的数据只有授权者才能访问和使用，防止数据泄露或被非法获取。对此，系统将对前后端传输的数据以及在数据库中存储的敏感数据，若用户密码等，采取一定的加密技术，避免数据的明文传输与存储。同时，系统的前后端也将添加特定的身份验证与访问控制机制，确保只有授权者才能获取到特定的数据。
- 数据可用性是指确保数据能够在需要时可靠地访问和使用，防止数据丢失或不可用。针对这一点，系统所使用的DBMS本身就具有一套完整成熟的解决方案，来保证存储在数据库中的数据是可靠与可恢复的。

2.6 假设和依赖

假设与依赖编号	假设与依赖描述
AS-1	系统可以从国家法律法规数据库、中国裁判文书网等网站爬取到充足的数据
AS-2	系统可以得到足够的服务器支持
AS-3	系统数据的更新信息能够得到实时刷新

AS-4	系统的用户掌握了基本的电脑与浏览器操作
AS-5	系统多轮的迭代开发过程有相应的人员负责维护和更新
AS-6	系统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事件
AS-8	团队具有充足的时间用于系统的开发与测试

2.7 开发、测试和运行环境

2.7.1 软件运行环境

项目	名称	版本
操作系统	Windows 10 及以上， Linux, MacOS	/
网站服务器	Nginx	1.22.1以上
数据库服务器	Linux socket	/
数据库服务器类型	MySQL / MongoDB / Elasticsearch	8.0
浏览器	Chrome, Safari, Edge, Firefox等主流浏览器	/

2.7.2 硬件运行环境

项目	名称
操作系统	CPU: CORE i5 及以上
内存：4G 及以上	/
硬盘：500G 及以上	/
应用服务器	内存：2GB 及以上
数据库服务器邮件服务器文件服务器	硬盘：50G 及以上
通讯设备	网线：具有良好的数据传输能力

2.8 项目文档

在“律镜”法律领域垂直搜索引擎项目中，为了支持项目的开发、管理和维护，涉及到多种文档，分为以下三类：描述类文档、过程类文档和参考类文档。

2.8.1 描述类文档

i. 需求规格说明书（SRS）：

- 描述：详细记录“律镜”法律领域垂直搜索引擎项目的功能需求、性能需求、非功能性需求等各方面的规格。
- 目的：作为开发人员、测试人员和项目经理的参考，确保项目的开发和实施与用户需求一致。

ii. 软件设计文档（SDD）：

- 描述：描述系统的整体架构、模块设计、数据库设计等方面的详细设计信息。
- 目的：为开发人员提供系统设计的指导，确保开发过程中的一致性和可维护性。

2.8.2 过程类文档

i. 项目计划书：

- 描述：包括项目的整体计划、开发阶段、测试阶段、上线部署等各个阶段的计划安排。
- 目的：为项目经理和团队成员提供清晰的项目时间表，帮助项目的有序推进。

ii. 测试计划书：

- 描述：描述测试的整体策略、测试目标、测试资源、测试计划和风险评估。
- 目的：为测试团队提供详细的测试计划，确保全面、系统地测试项目的功能和性能。

2.8.3 参考类文档

i. 用户手册：

- 描述：提供给最终用户的文档，包含系统的使用说明、操作指南和常见问题解答。
- 目的：帮助用户正确、高效地使用律镜法律领域垂直搜索引擎，解决使用过程中的问题。

ii. 技术文档：

- 描述：包括系统的架构图、数据库设计文档、代码注释等技术性的文档。
- 目的：为开发人员、维护人员提供详细的技术资料，方便系统的维护和进一步开发。

以上文档将在不同阶段发挥关键作用，确保律镜法律领域垂直搜索引擎项目的有序推进、满足用户需求和确保系统质量。这些文档将是项目管理、开发和维护的重要参考。

2.9 产品开发项目设计人员

序号	角色	职责	人员
1	项目经理	负责项目的总体规划、组织、协调和控制，按实际情况分配和安排各人员工作，确保项目按时、按质量要求完成。	 徐毕颖
2	业务分析师	负责市场调研和收集用户需求，将业务需求转化为可执行的软件需求，确定软件功能的具体范围。	 王梓轩
3	设计总监	建立系统框架，数据库设计，概要设计，技术评审	 章越
4	测试经理	组织编写测试计划和测试方案，组织系统测试；参加技术评审	 沈书豪
5	质量经理	带领质量监督组成员制定质量保证计划，对监督组反应的质量问题进行汇总，和产品经理以及项目经理进行交流沟通	 黄文杰
6	开发人员	负责进行编码工作和单元测试，进行系统集成，及时解决测试的时候出现的问题	全体人员
7	测试人员	编写测试方案和测试用例，进行系统测试，向开发组反馈 bug	全体人员
8	软件质量监督	实时对质量经理以及项目经理提供项目进度与项目实际开发时候的差异提出报告，指出差异原因和改进方法	全体人员

3. 系统需求分析概述

3.1 系统总用例图

系统用例图实际上是对系统功能需求的进一步明确和整理。在下面的总用例图中，我们整理并展示了本系统涉及到的所有功能需求（浅色椭圆），将其按模块归类（深色椭圆）并用连线上的注释内容描述各用例间的关系：



要对各个需求进行优先级分析，我们可以从功能**使用频率**和**关键程度**两个维度展开四象限气泡图，实现更直观的需求优先级分析：

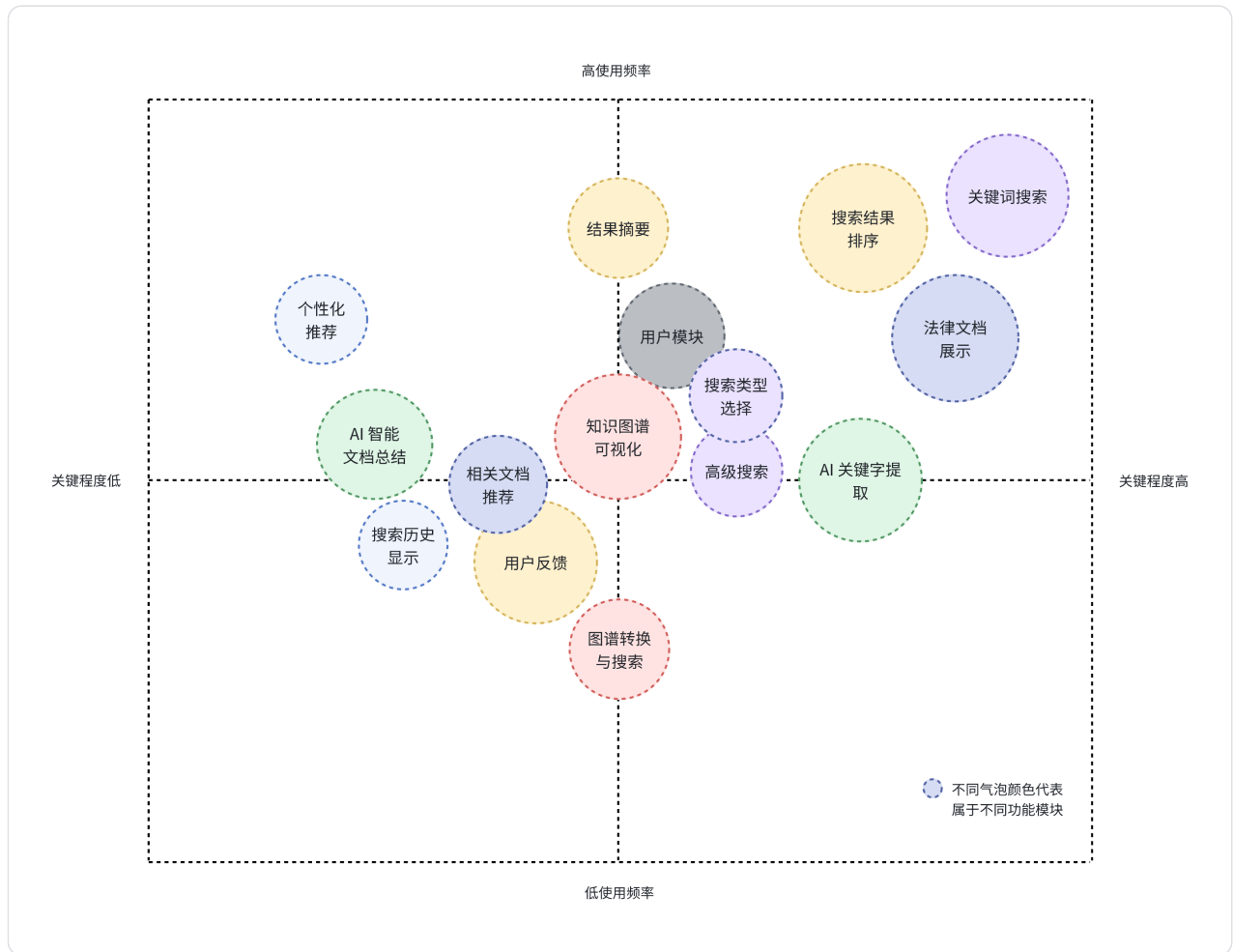


图3-2 需求优先级分析的四象限气泡图

说明：在这张需求优先级分析象限图中，高使用频率且关键程度高（右上角）的功能需求应该被作为核心功能优先实现和满足，然后是关键程度次高并且使用频率适中（中部略偏右上）的那些需求，最后再实现关键程度和使用频率都较低（偏左下一带）的功能需求，需求优先级便由此完成确认了。

3.3 系统模块架构图

在完成了上述总体分析之后，本系统的功能模块构成已经划分得比较清晰了，在此用架构图的形式将其表示出来，本图也可为4.功能需求一节提供参考。

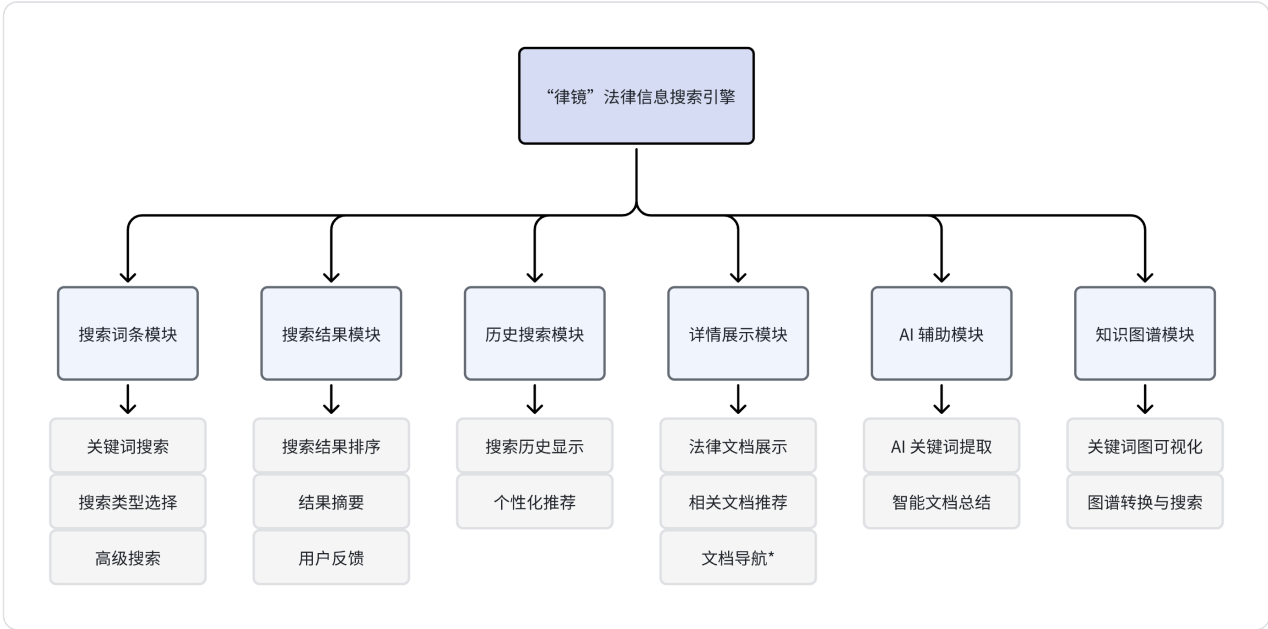


图3-3 系统模块架构图

4. 功能需求

4.1 关键词搜索

用例编号	SE-01	用例名称	关键词搜索
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	用户在搜索框中输入一个或多个关键词或短语，或者一段长文本，点击搜索按钮搜索指定的内容。		
前置条件	1. 系统可以正常运行 2. 用户用有一定的电脑操作技术，会使用输入法		
触发器	用户进入搜索引擎主页		
后置条件	若用户已登录，记录用户搜索历史		
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页面
分支过程	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“查询异常，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：系统显示“系统异常，请联系系统维护人员”，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词和短语
输出	搜索结果显示页面
包括用例	无
优先级	高
使用频率	高
备注	无

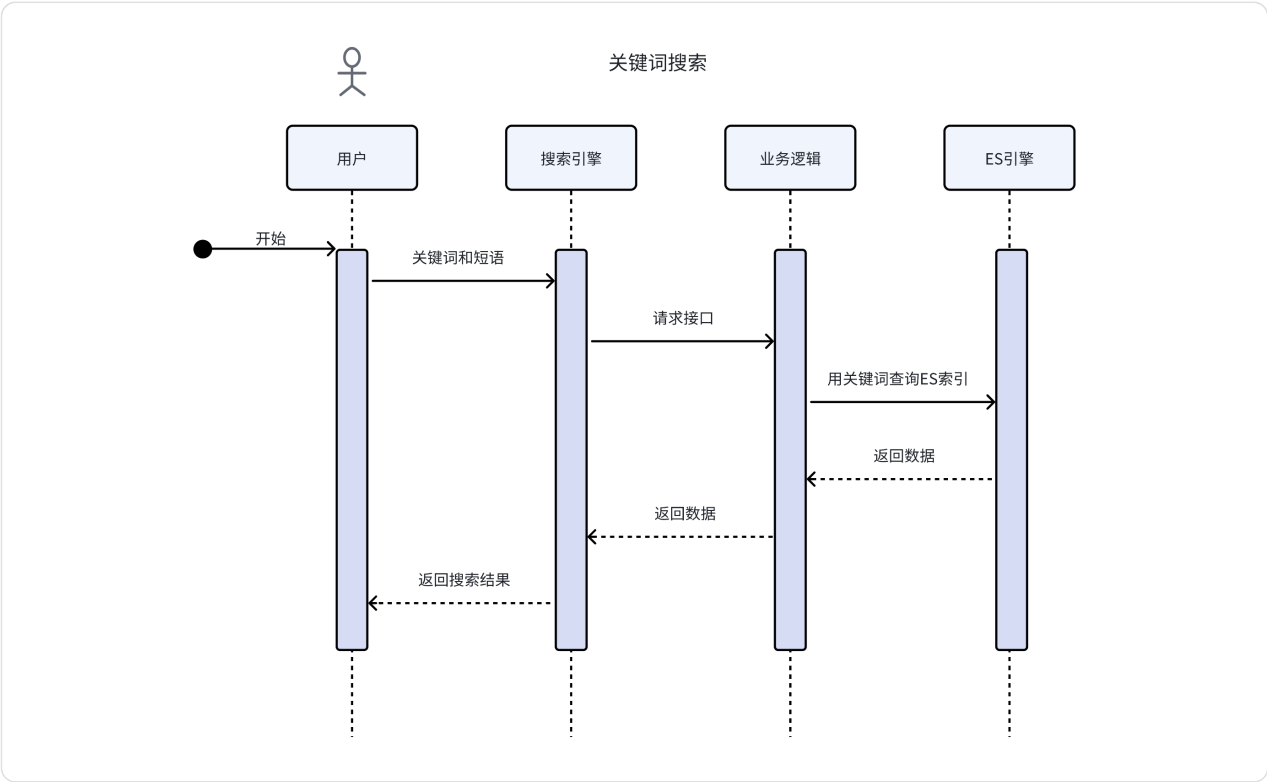


图 4-1 关键词搜索时序图

4.2 搜索类型选择

用例编号	SE-02	用例名称	搜索类型选择
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	用户选择搜索的类型，如全文搜索、标题搜索、文档类型搜索、作者搜索等，点击搜索按钮搜索指定的内容。		
前置条件	<ol style="list-style-type: none"> 1. 系统可以正常运行 2. 用户用有一定的电脑操作技术，会使用输入法 		
触发器	用户进入搜索引擎主页		
后置条件	若用户已登录，记录用户搜索历史		
输入信息	选中指定的搜索类型并输入搜索关键词		
主干过程	<ol style="list-style-type: none"> 1. 用户进入搜索引擎主页 2. 用户选中指定的搜索类型 3. 用户输入关键词 4. 用户点击搜索按钮 5. 搜索成功后跳转至结果显示页面 		
分支过程	无		
异常	<ol style="list-style-type: none"> 1. 用户输入危险信息（发生于主干过程3） 处理：系统提示“查询异常，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5） 处理：系统显示“系统异常，请联系系统维护人员”，系统结束用例 		
假设	没有任何异常发生		
输入	用户选择的搜索类型，想要搜索的关键词和短语		
输出	搜索结果显示页面		
包括用例	关键词搜索		

优先级	中
使用频率	中
备注	无

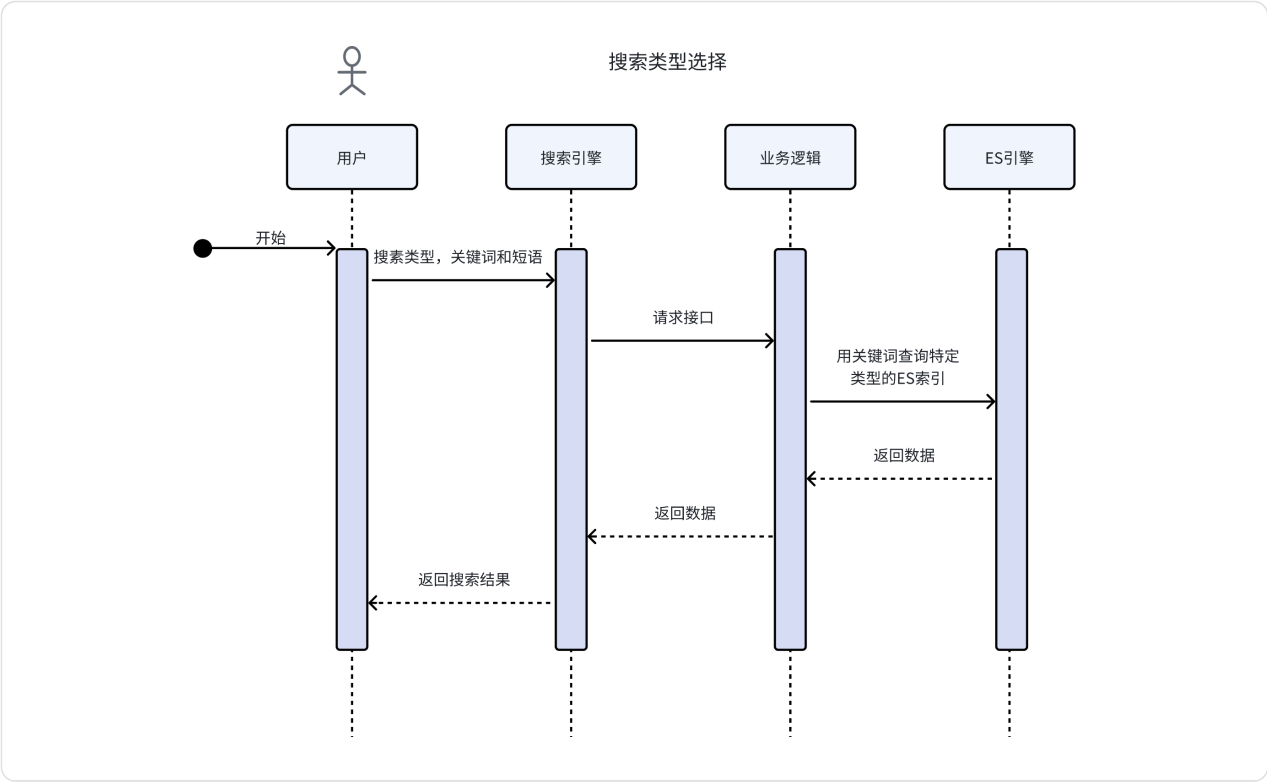


图 4-2 搜索类型选择时序图

4.3 高级搜索

用例编号	SE-03	用例名称	高级搜索
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	用户对搜索的方式进行一定的限定，如限制时间范围、进行模糊匹配等，或根据需要的搜索结果类型（如法律法规、裁判文书、案件描述）指定需要的搜索结果。		
前置条件	1. 系统可以正常运行 2. 用户用有一定的电脑操作技术，会使用输入法 3. 关键词搜索功能已经实现		

触发器	用户进入搜索引擎主页
后置条件	若用户已登录，记录用户搜索历史
输入信息	选中指定的搜索类型并输入搜索关键词
主干过程	<ol style="list-style-type: none">1. 用户进入搜索引擎主页2. 用户输入关键词3. 用户设置搜索条件限制4. 用户点击搜索按钮5. 搜索成功后跳转至结果显示页面
分支过程	无
异常	<ol style="list-style-type: none">1. 用户输入危险信息（发生于主干过程2） 处理：系统提示“查询异常，请重新输入”，系统结束用例2. 用户设置矛盾的条件限制（发生于主干过程3） 处理：系统提示“条件限制错误，请重试”，系统结束用例3. 搜索后页面出现异常（发生于主干过程5） 处理：系统显示“系统异常，请联系系统维护人员”，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词和短语，查询条件限制
输出	搜索结果显示页面
包括用例	关键词搜索
优先级	中
使用频率	中
备注	无

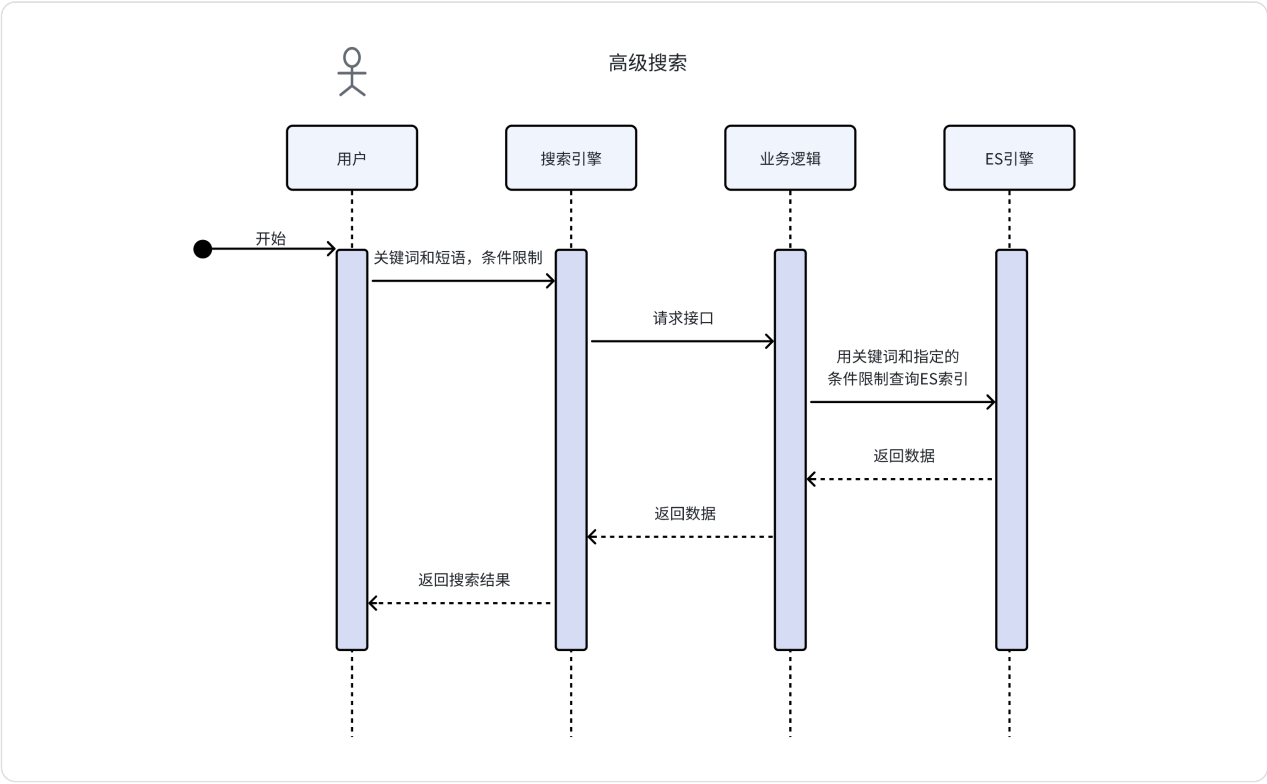


图 4-3 高级搜索时序图

4.4 搜索结果排序

用例编号	SE-04	用例名称	搜索结果排序
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	搜索引擎根据相关性对搜索结果进行排序，基于用户的关键词、搜索方式等因素为结果设置不同的权重，将最相关的结果显示在前面。		
前置条件	1. 系统可以正常运行 2. 用户用有一定的电脑操作技术，会使用输入法 3. 用户可以正常打开相关界面		
触发器	用户进入搜索引擎主页		
后置条件	无		
输入信息	搜索关键词、短语		

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. 结果界面按照相关度显示排序后的搜索结果
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：系统显示“系统异常，请联系系统维护人员”，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词和短语
输出	搜索结果显示页面
包括用例	无
优先级	高
使用频率	高
备注	无

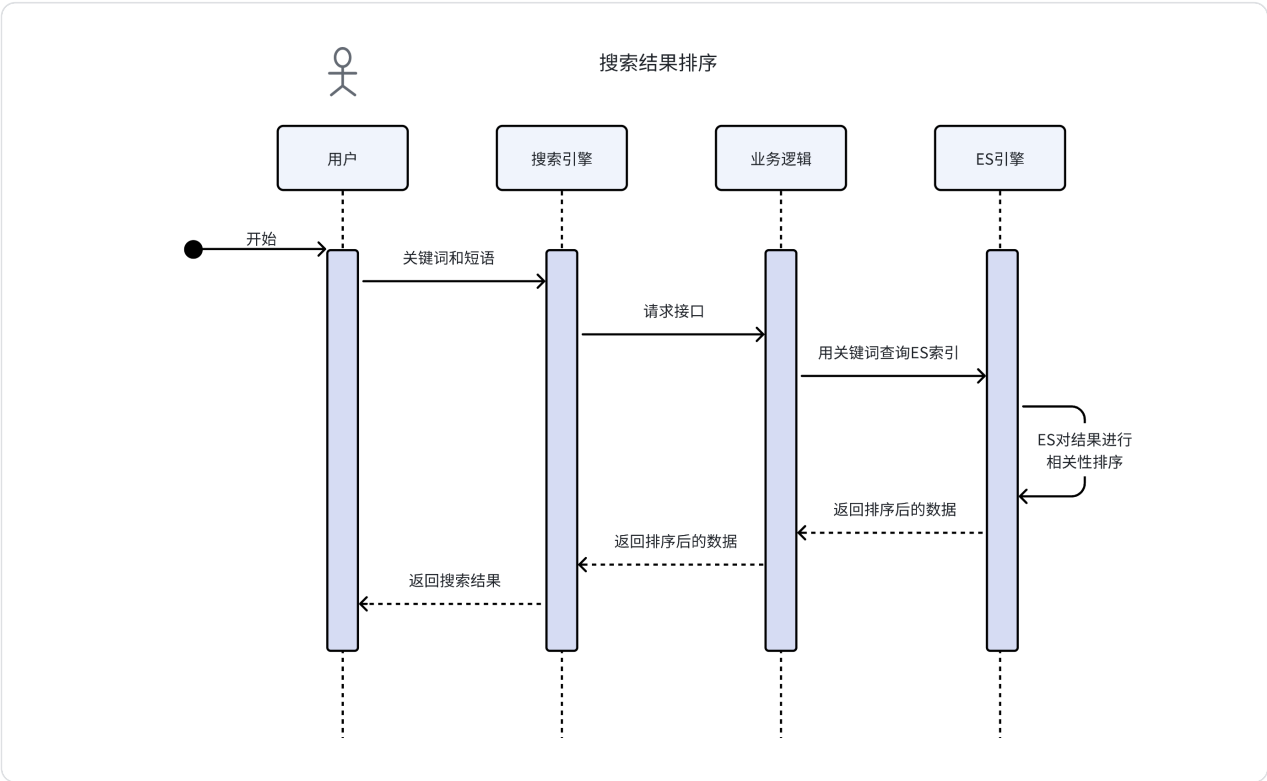


图 4-4 搜索结果排序时序图

4.5 结果摘要

用例编号	SE-05	用例名称	结果摘要
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	对于每个搜索结果，系统将生成一个摘要，包含相关文档的片段，以结果卡片的形式显示给用户。		
前置条件	<ol style="list-style-type: none">1. 系统可以正常运行2. 用户拥有一定的电脑操作技术，会使用输入法3. 用户可以正常打开相关界面		
触发器	用户进入搜索引擎主页		
后置条件	无		
输入信息	搜索关键词、短语		
主干过程	<ol style="list-style-type: none">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. 结果界面显示每一条搜索结果的结果卡片列表		
分支过程	无		
异常	<ol style="list-style-type: none">1. 搜索后页面出现异常（发生于主干过程4） <p>处理：系统显示“系统异常，请联系系统维护人员”，系统结束用例</p>		
假设	没有任何异常发生		
输入	用户想要搜索的关键词和短语		
输出	搜索结果显示页面		
包括用例	关键词搜索		
优先级	中		

使用频率	高
备注	无

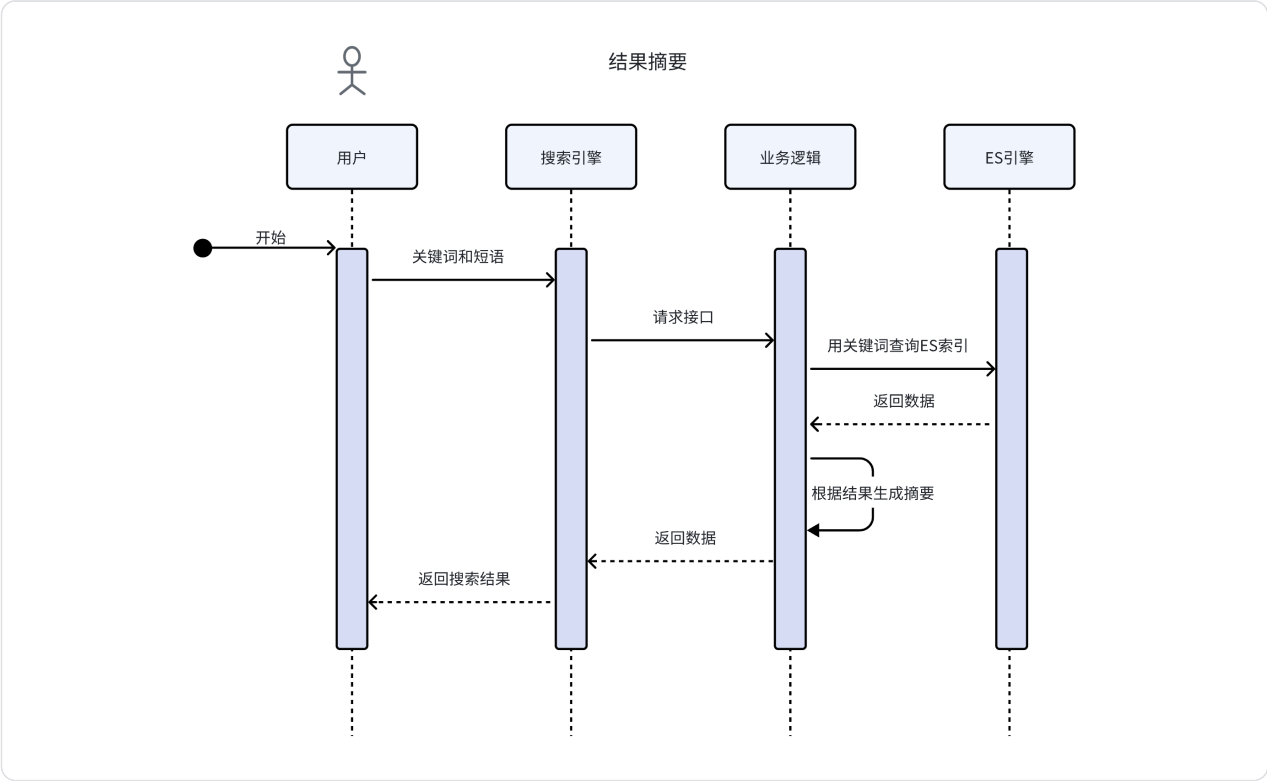


图 4-5 结果摘要时序图

4.6 用户反馈

用例编号	SE-06	用例名称	用户反馈
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	用户可以对搜索结果进行反馈（点赞点踩等），规定只有登录的用户才能点赞/点踩。		
前置条件	<div>1. 系统可以正常运行</div> <div>2. 用户用有一定的电脑操作技术，会使用输入法</div> <div>3. 用户可以正常打开相关界面</div> <div>4. 用户已经登录</div>		

触发器	用户进入搜索引擎主页
后置条件	将点赞/点踩数据保存至数据库
输入信息	点赞/点踩及对应的文档
主干过程	<ol style="list-style-type: none"> 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. 结果界面显示每一条搜索结果的结果卡片列表 5. 用户对指定的结果卡片点赞或点踩
分支过程	无
异常	<ol style="list-style-type: none"> 1. 搜索后页面出现异常（发生于主干过程4） 处理：系统显示“系统异常，请联系系统维护人员”，系统结束用例 2. 用户在未登录的情况下试图执行点赞/点踩操作（发生于主干过程5） 处理：系统提示“请先登录”，系统结束用例
假设	没有任何异常发生
输入	用户点赞/点踩及对应的文档
输出	点赞/点踩成功提示
包括用例	无
优先级	低
使用频率	低
备注	无

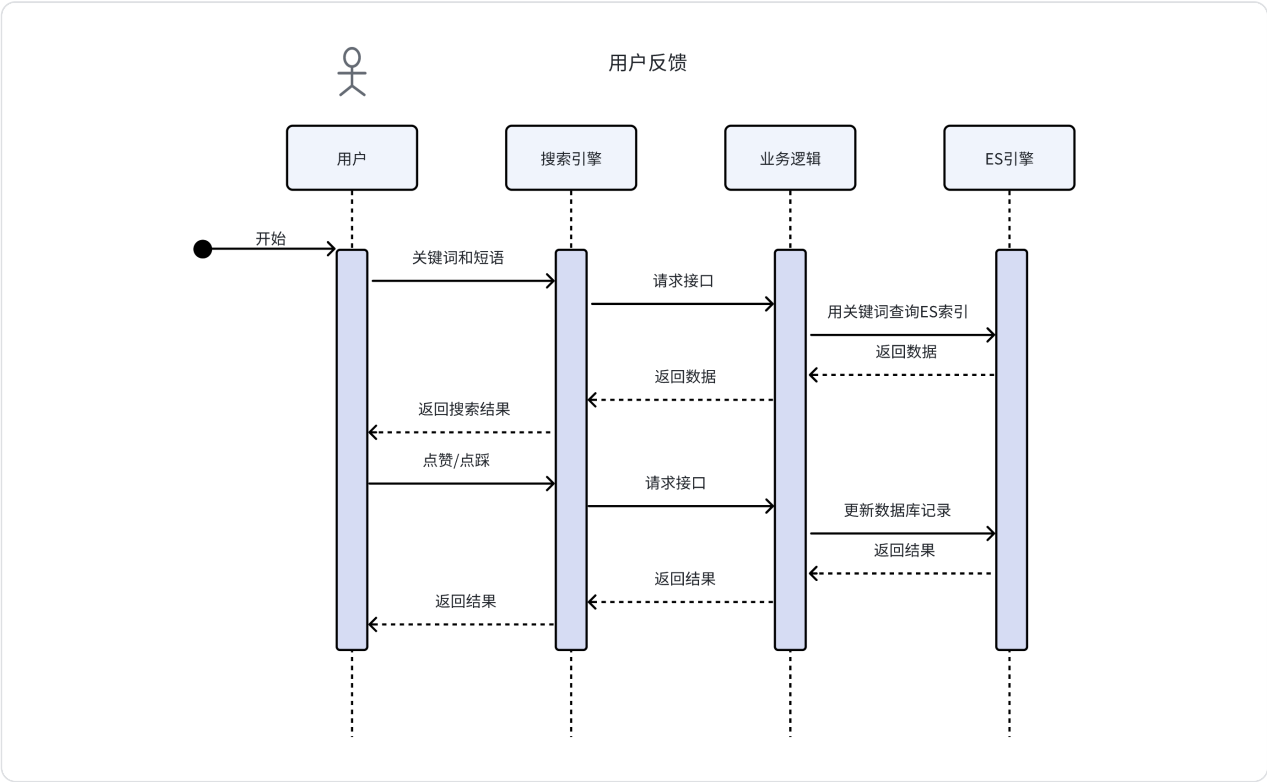


图 4-6 用户反馈时序图

4.7 搜索历史显示

用例编号	SE-07	用例名称	搜索历史显示
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	记录已登录用户的查询历史，在查询界面提供搜索历史提示，用户可以点击历史查询记录进行快捷查询。		
前置条件	1. 系统可以正常运行 2. 用户可以正常打开相关界面 3. 用户已经登录		
触发器	用户进入搜索引擎主页		
后置条件	无		
输入信息	选中历史查询条目		

主干过程	1. 用户进入搜索引擎主页 2. 用户点击历史查询条目 3. 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3） 处理：系统显示“系统异常，请联系系统维护人员”，系统结束用例
假设	没有任何异常发生
输入	用户点击历史查询条目
输出	查询结果页面
包括用例	关键词搜索
优先级	低
使用频率	低
备注	无

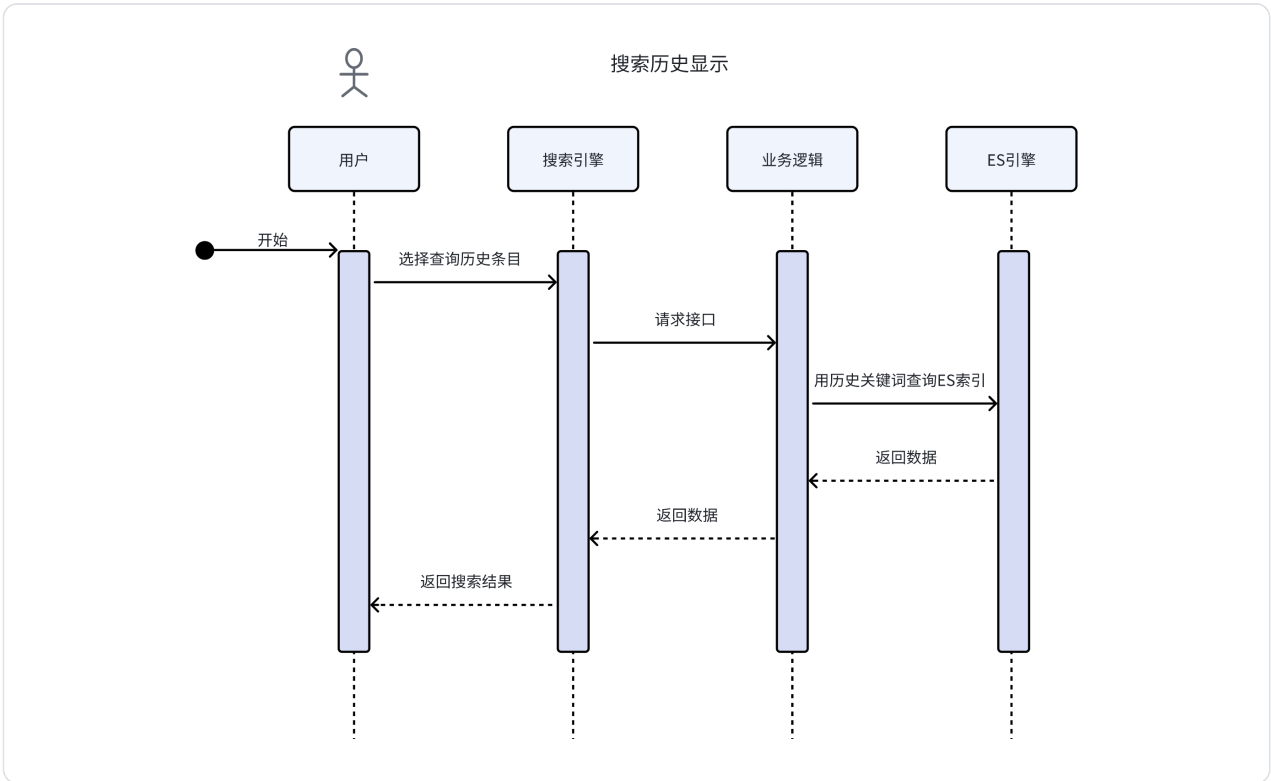


图 4-7 搜索历史显示时序图

4.8 个性化推荐

用例编号	SE-08	用例名称	个性化推荐
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	通过分析用户的搜索历史、点击行为、喜好等信息，为用户提供定制化的推荐内容。根据用户的兴趣和需求，搜索引擎推荐相关的法律文档、案例、法规等内容		
前置条件	<ol style="list-style-type: none">1. 系统可以正常运行2. 用户可以正常打开相关界面3. 用户已经登录		
触发器	用户进入搜索引擎主页		
后置条件	无		
输入信息	无		
主干过程	<ol style="list-style-type: none">1. 用户进入搜索引擎主页2. 搜索引擎显示个性化推荐内容		
分支过程	无		
异常	无		
假设	没有任何异常发生		
输入	无		
输出	个性化推荐内容		
包括用例	无		
优先级	低		
使用频率	低		
备注	无		

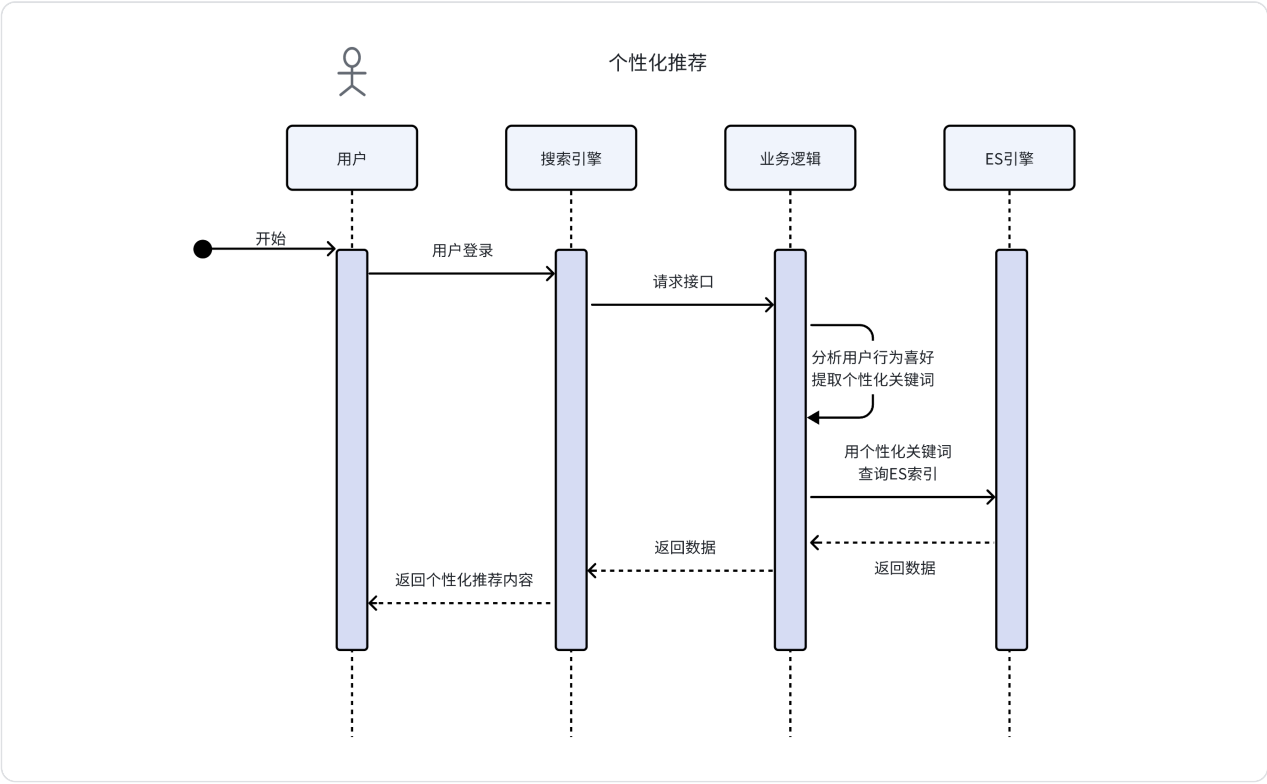


图 4-8 个性化推荐时序图

4.9 法律文档展示

用例编号	SE-09	用例名称	法律文档展示
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	根据用户选中的搜索结果，系统展示具体的法律文档内容，并以合适的字体、排版来显示文档的标题、作者、导语、正文等部分的内容。		
前置条件	1. 系统可以正常运行 2. 用户可以正常打开相关界面		
触发器	用户进入结果显示界面		
后置条件	无		
输入信息	待显示的文档		
主干过程	1. 用户进入搜索引擎主页		

	<ol style="list-style-type: none"> 2. 用户输入关键词 3. 用户点击搜索按钮 4. 跳转至结果显示界面 5. 用户点击指定的结果条目 6. 跳转至文档展示界面
分支过程	无
异常	<ol style="list-style-type: none"> 1. 搜索后页面出现异常（发生于主干过程4） 处理：系统显示“系统异常，请联系系统维护人员”，系统结束用例 2. 获取指定文档的内容失败（发生于主干过程6） 处理：系统显示“文档获取失败，请联系系统维护人员”，系统结束用例
假设	没有任何异常发生
输入	用户选中要查看的文档
输出	法律文档详细内容
包括用例	无
优先级	高
使用频率	高
备注	无

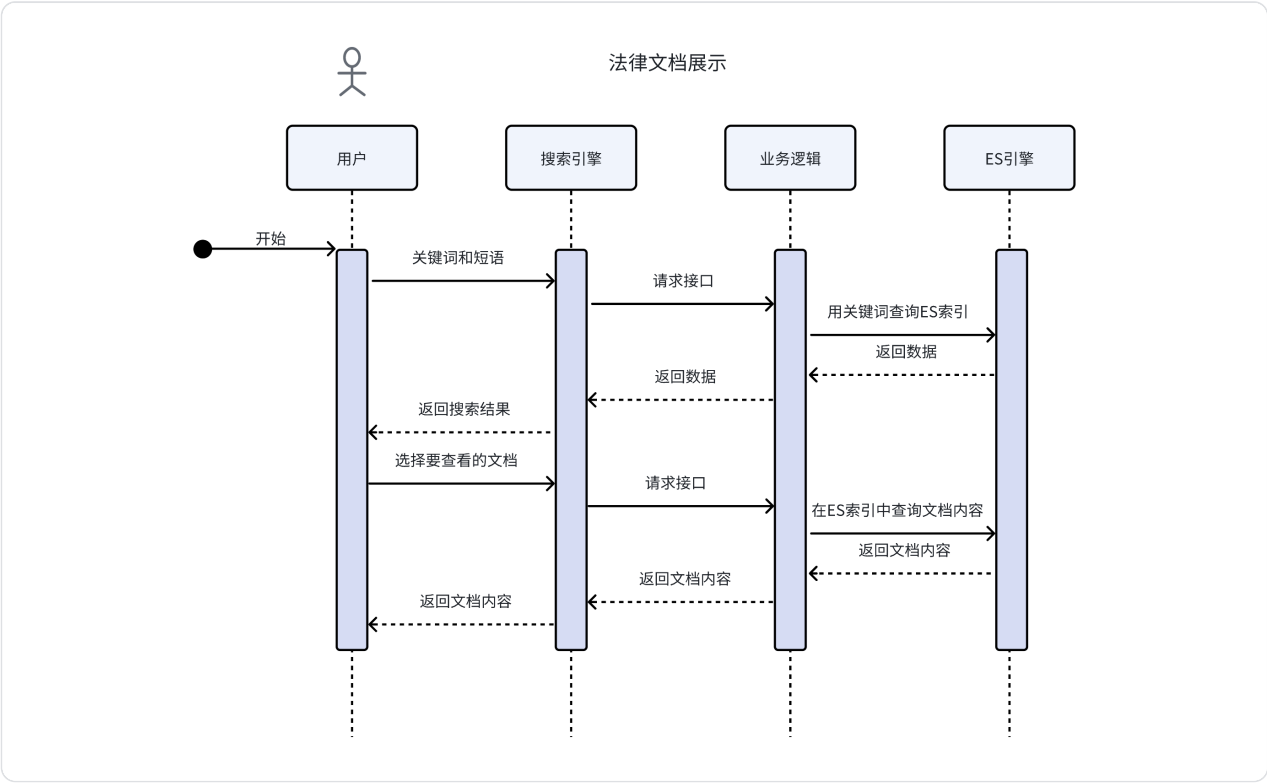


图 4-9 用户反馈时序图

4.10 相关文档

用例编号	SE-10	用例名称	相关文档
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	在当前文档的最后，搜索引擎基于该文档的内容与类型，提供相关性较高的其他文档的链接。		
前置条件	1. 系统可以正常运行 2. 用户可以正常打开相关界面		
触发器	用户进入文档展示界面		
后置条件	无		
输入信息	当前显示的文档		
主干过程	1. 用户进入文档展示界面		

	2. 用户点击相关文档链接 3. 跳转至相关文档展示界面
分支过程	无
异常	1. 获取指定文档的内容失败（发生于主干过程3） 处理：系统显示“文档获取失败，请联系系统维护人员”，系统结束用例
假设	没有任何异常发生
输入	用户当前查看的文档
输出	相关文档链接
包括用例	法律文档展示
优先级	低
使用频率	中
备注	无

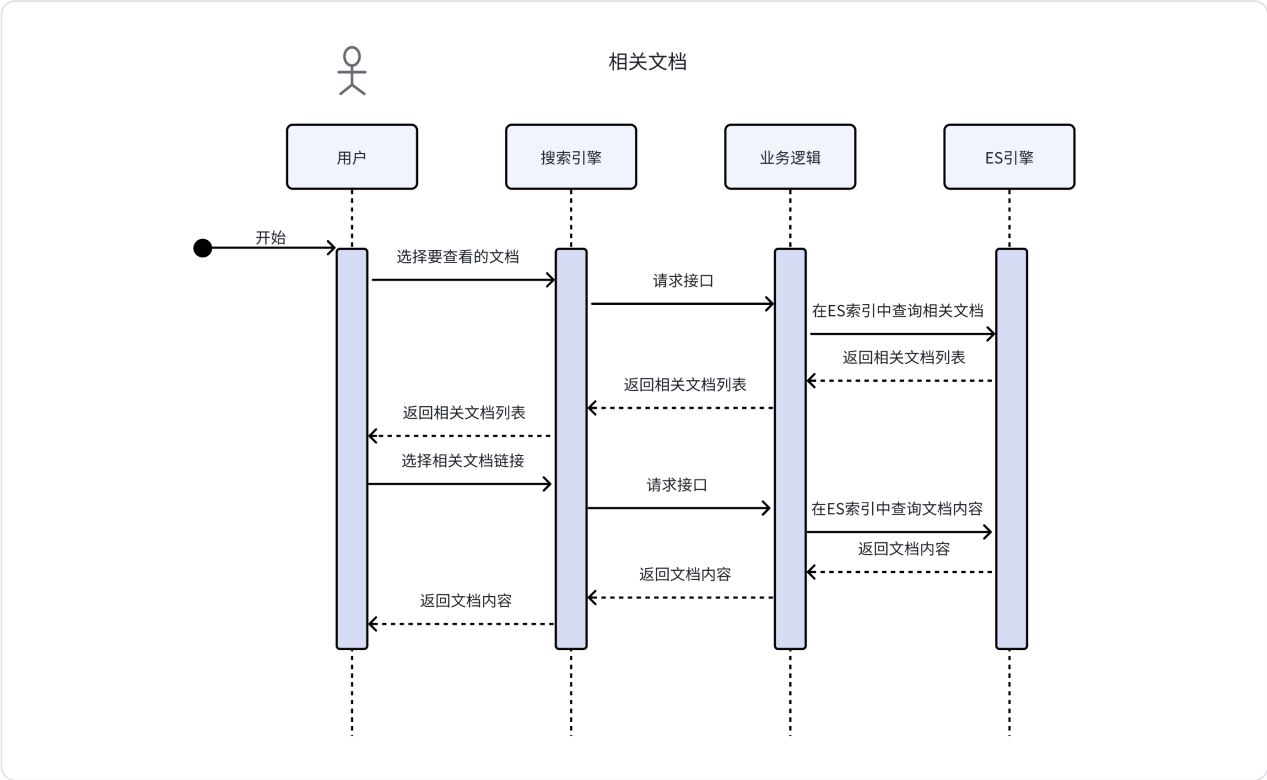


图 4-10 相关文档时序图

4.11 关键词提取

--	--	--	--

用例编号	SE-11	用例名称	关键词提取
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	调用相关 API，对用户输入的查询进行分析，提取其中的关键字，用于优化搜索算法，提高搜索结果的相关性。		
前置条件	<ol style="list-style-type: none"> 1. 系统可以正常运行 2. 用户可以正常打开相关界面 3. 系统能够正常调用AI模型api 		
触发器	用户进入搜索引擎主页		
后置条件	无		
输入信息	搜索内容描述		
主干过程	<ol style="list-style-type: none"> 1. 用户进入搜索引擎主页 2. 用户在AI对话框内输入搜索内容描述 3. 用户点击关键词提取按钮 4. 界面显示AI从搜索内容描述中提取出的关键词 		
分支过程	无		
异常	<ol style="list-style-type: none"> 1. AI模型接口调用失败（发生于主干过程4） <p>处理：系统显示“AI接口调用异常，请稍后重试”，系统结束用例</p>		
假设	没有任何异常发生		
输入	用户待搜索的搜索内容描述		
输出	提取的关键词		
包括用例	无		
优先级	低		
使用频率	低		
备注	无		

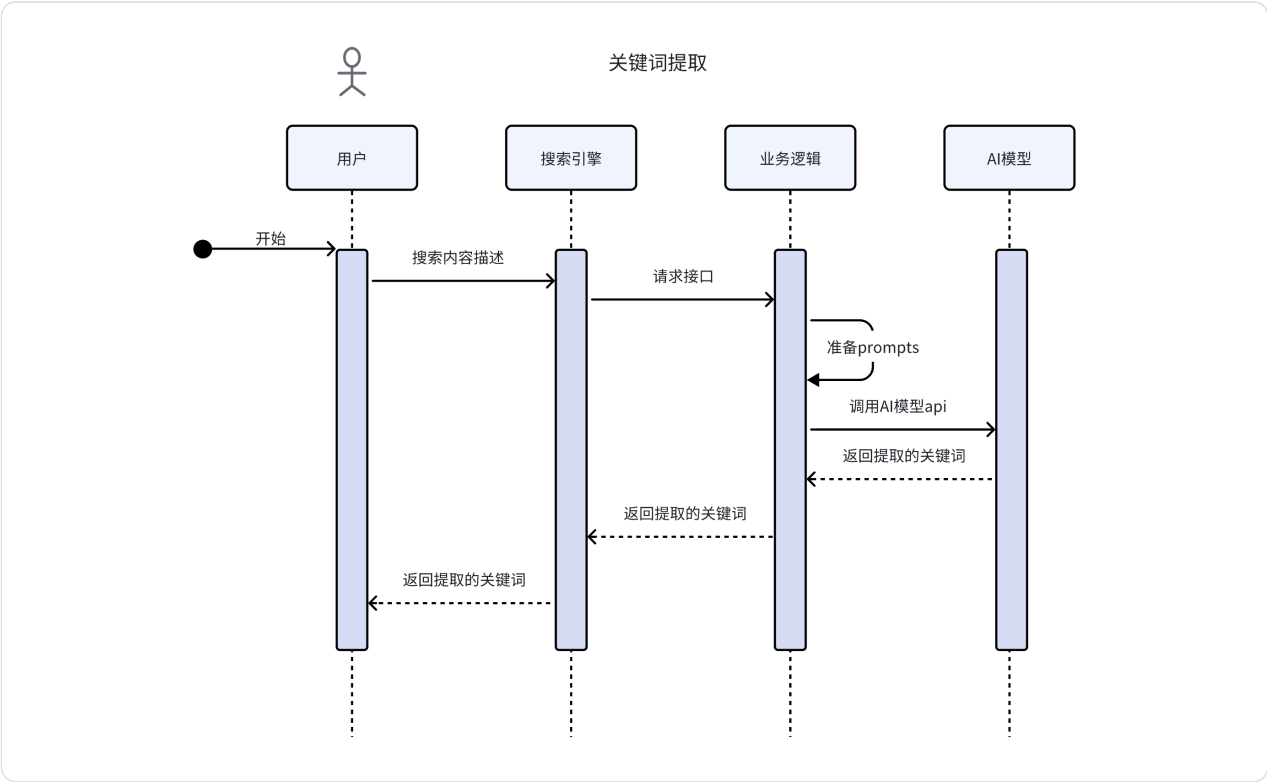


图 4-11 关键词提取时序图

4.12 文档总结

用例编号	SE-12	用例名称	文档总结
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	利用 AI 对搜索结果文档进行智能总结，提供用户更为精炼和易理解的信息，生成简明扼要的摘要，帮助用户快速了解文档内容，提高信息获取效率。		
前置条件	1. 系统可以正常运行 2. 用户可以正常打开相关界面 3. 系统能够正常调用AI模型api		
触发器	用户进入文档展示界面		
后置条件	无		
输入信息	文档内容		

主干过程	1. 用户进入文档展示界面 2. 用户点击文档总结按钮 3. 界面显示AI生成的文档总结
分支过程	无
异常	1. AI模型接口调用失败（发生于主干过程3） 处理：系统显示“AI接口调用异常，请稍后重试”，系统结束用例
假设	没有任何异常发生
输入	用户正在查看的文档内容
输出	AI生成的文档总结
包括用例	无
优先级	低
使用频率	低
备注	无

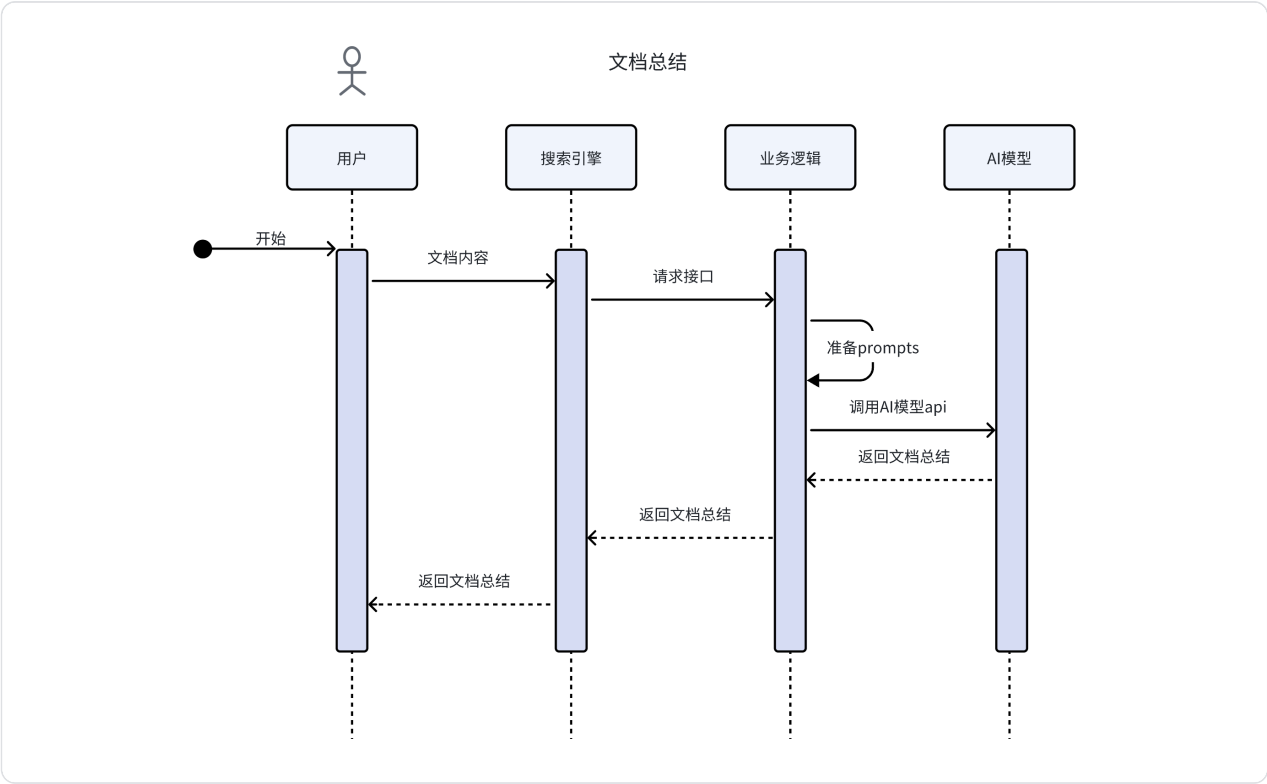


图 4-12 文档总结时序图

4.13 关键词图

用例编号	SE-13	用例名称	关键词图
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	对于任何一次搜索，系统将生成它知识图谱形式的关键词图：提取得到它的1-2个关键词，作为图的中心词，然后返回其深度1-2的其他关键词节点，组成一张美观的词图。		
前置条件	<ol style="list-style-type: none"> 1. 系统可以正常运行 2. 用户用有一定的电脑操作技术，会使用输入法 3. 用户可以正常打开相关界面 		
触发器	用户进入搜索引擎主页		
后置条件	无		
输入信息	搜索关键词、短语		
主干过程	<ol style="list-style-type: none"> 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. 结果界面显示关键词图 		
分支过程	无		
异常	<ol style="list-style-type: none"> 1. 搜索后页面出现异常（发生于主干过程4） 处理：系统显示“系统异常，请联系系统维护人员”，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词和短语		
输出	关键词图		
包括用例	无		
优先级	中		
使用频率	中		
备注	无		

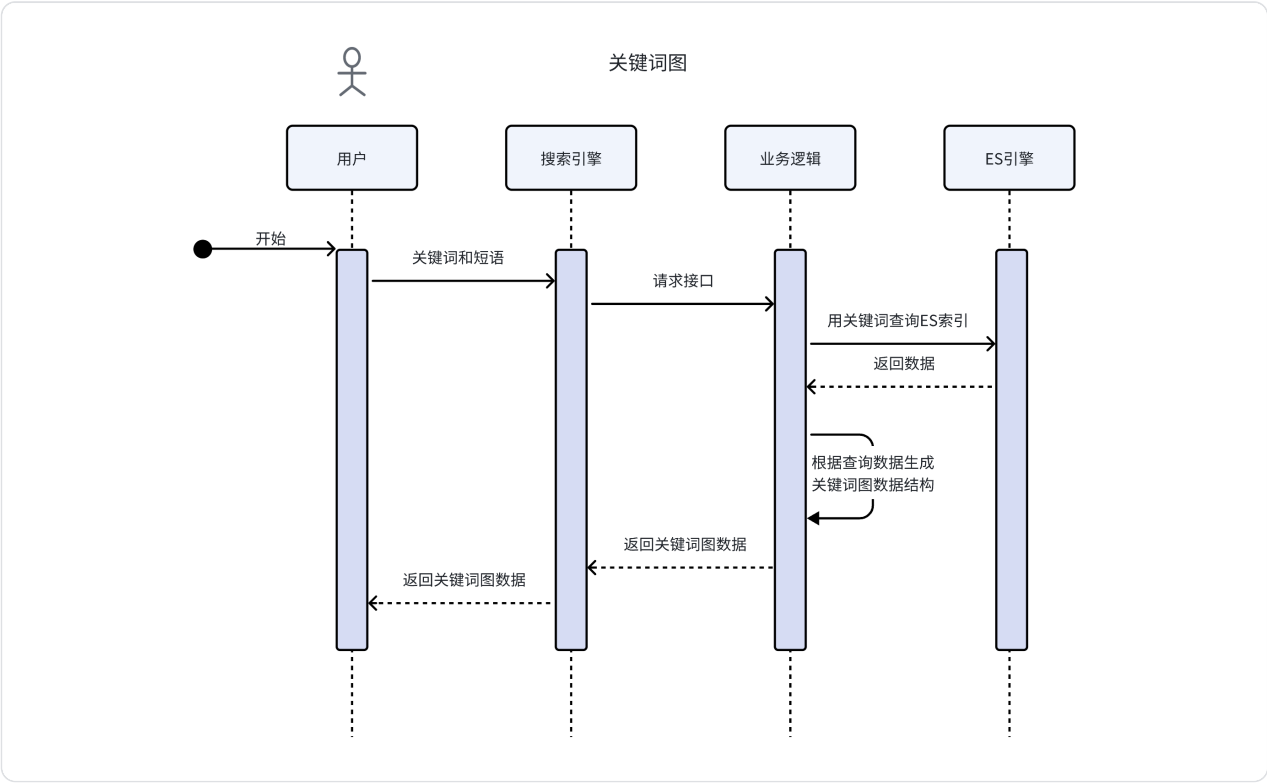


图 4-13 关键词图时序图

4.14 图谱转换与跳转搜索

用例编号	SE-14	用例名称	图谱转换与跳转搜索关键词图
创建人	沈书豪	最后修改人	沈书豪
创建日期	2023.11.22	最后修改日期	2023.11.22
角色	普通用户	需求来源	需求分析和小组会议
主要参与者	普通用户		
描述	点击词图中的各节点，可以将其转换为中心词（提供动态转换效果），并且用户可以选择是否以该节点上的词为输入值，再次搜索。		
前置条件	1. 系统可以正常运行 2. 用户用有一定的电脑操作技术，会使用输入法 3. 用户可以正常打开相关界面		
触发器	用户进入搜索引擎主页		
后置条件	无		
输入信息	点击关键词图节点		

主干过程	<ol style="list-style-type: none">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. 结果界面显示关键词图5. 用户点击关键词图节点6. 跳转至新的结果显示界面
分支过程	无
异常	<ol style="list-style-type: none">1. 搜索后页面出现异常（发生于主干过程4） 处理：系统显示“系统异常，请联系系统维护人员”，系统结束用例2. 关键词图节点跳转异常（发生于主干过程6） 处理：系统显示“”，系统结束用例
假设	没有任何异常发生
输入	用户点击想跳转的关键词图节点
输出	新的搜索结果页面
包括用例	无
优先级	中
使用频率	中
备注	无

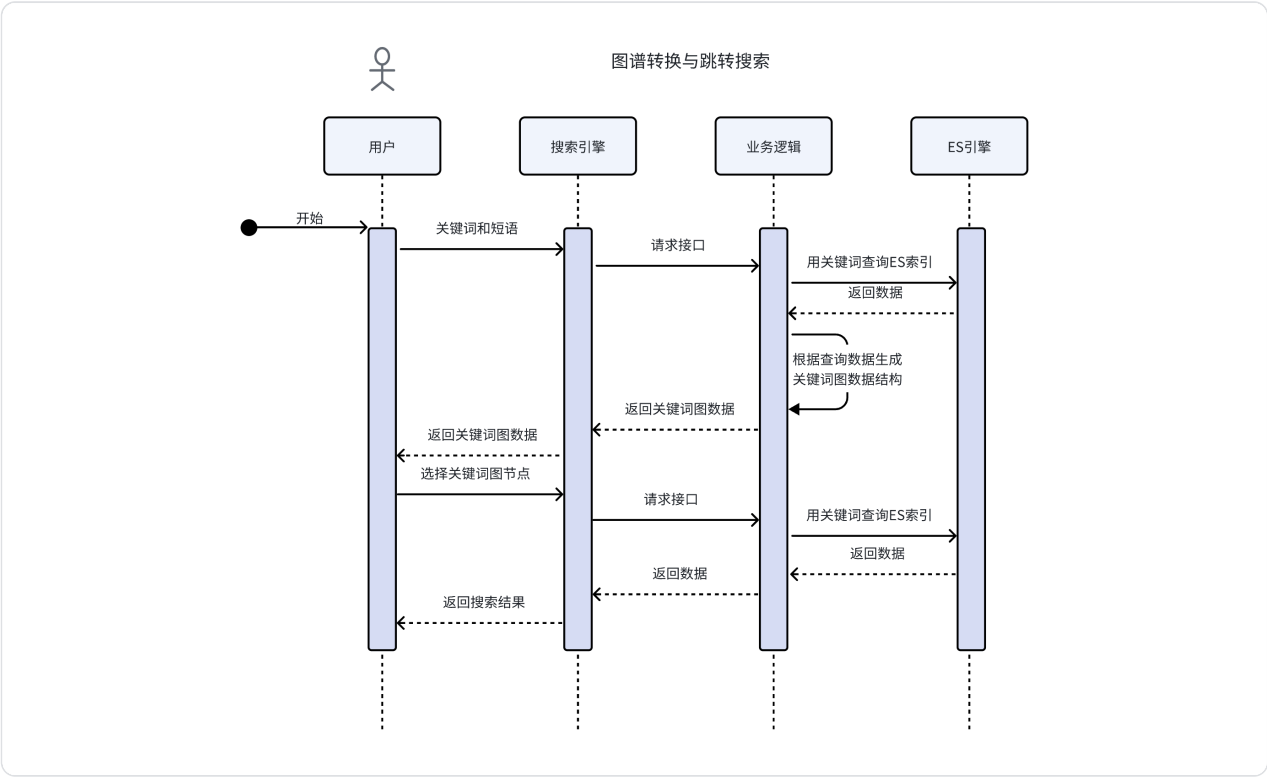


图 4-14 图谱转换与跳转搜索时序图

5. 数据流图 & 数据字典

5.1 数据流图

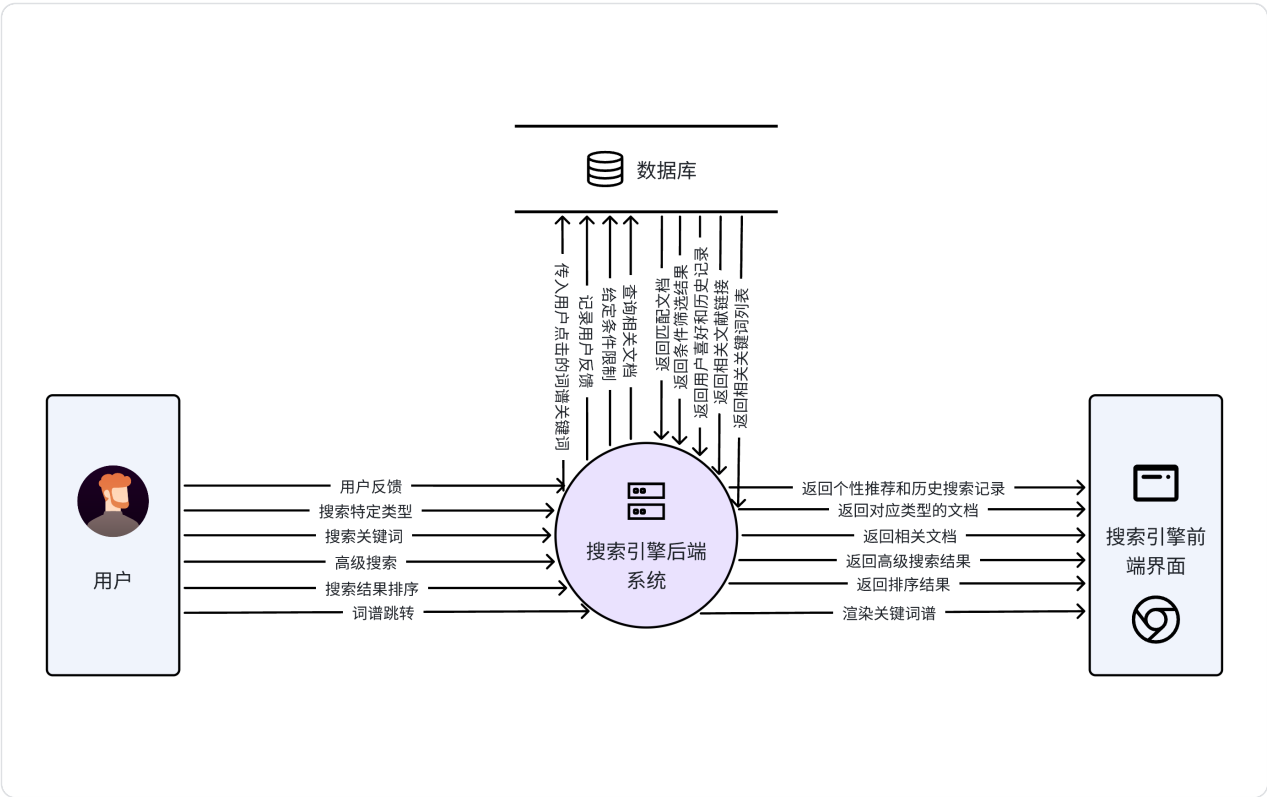


图5-1 数据流图

5.2 数据字典

5.2.1 数据元素定义表

编号	数据元素名	类型	值域	说明
E1	搜索关键词	字符串	/	用户输入的搜索关键词
E2	搜索类型	枚举型	1-N(N为搜索类型总数)	用户选择的搜索类型
E3	搜索结果	字符串	/	返回的搜索结果
E4	用户反馈的内容	枚举型	0或1（赞或踩）	用户对搜索结果的反馈
E5	搜索历史记录	字符串列表	/	用户的搜索历史记录
E6	推荐内容	字符串	/	系统基于用户行为推荐的内容
E7	文档标题	字符串	/	法律文档的标题
E8	文档作者	字符串（或字符串列表）	/	法律文档的作者
E9	文档类型	枚举型	1-N（N为文档类型的总数）	法律文档的类型
E10	文档内容	字符串	/	法律文档的具体内容
E11	关键词提取结果	字符串列表	/	从查询中提取的关键词
E12	文档摘要	字符串	/	文档的摘要内容
E13	关键词图谱	字符串列表	/	每次搜索生成的关键词图谱

5.2.2 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	用户在搜索框中输入的查询字符串	"合同法"
E2	搜索类型	枚举型（整形）	1-N(N为搜索类型总数)	用户选择的搜索类型	1（全文搜索）

E3	搜索结果	字符	256个字符以内	返回的搜索结果	"《合同法》相关条款..."
E4	用户反馈	枚举型 (整形)	0或1 (赞或踩)	用户对搜索结果的反馈	0 (赞)、1 (踩)
E5	搜索历史记录	字符	单次记录256个字符以内	用户的搜索历史记录 (列表)	"合同法案例" "民法案例"
E6	推荐内容	字符	256个字符以内	系统推荐的内容 (列表)	"最新法规更新"
E7	文档标题	字符	128个字符以内	法律文档的标题	"新民法典"
E8	文档作者	字符	20个字符以内 (单个名字)	法律文档的作者 (列表)	"张律师"
E9	文档类型	枚举型 (整形)	1-N (N为文档类型的总数)	法律文档的类型	"司法解释"
E10	文档内容	字符	-	法律文档的具体内容	"《合同法》第一条..."
E11	关键词提取结果	字符	32个字符以内 (单个)	从查询中提取的关键词 (列表)	"合同"、"履行"
E12	文档摘要	字符	512个字符以内	文档的摘要内容	"本文解释了《合同法》..."
E13	关键词图谱	字符	-	每次搜索生成的关键词 图谱 (列表)	"合同"、"民法"、"契约精神"
E14	AI处理数据	字符	-	AI处理后的数据	"搜索算法优化结果..."
E15	用户身份验证信息	字符	256个字符以内	用户的身份验证信息	"user123"
E16	用户偏好信息	字符	256个字符以内	用户的搜索偏好、设置	"偏好：合同法"

6. 上下文图

上下文图是一种高层次的建模工具，用于描述系统与外部实体之间的交互关系。它将整个系统的功能完全抽象，看做一个黑箱，因此系统的功能就只能通过与外部实体的交互来体现。上下文图通常

本系统的参与者只有用户一类，上下文图表示如下：



对话图（DM 图）是用来描述系统内部各个组件之间，以及系统与外部实体之间的交互的图。与上下文图不同的地方在于。它还展示了内部各组件（模块）之间的交互。

The diagram illustrates the architecture of the "Law Mirror" platform, showing the flow from a user request to a user response.

Platform Components and Flow:

- 用户 (请求) (User (Request))**: Initiates the process.
- 搜索历史 (Search History)** and **个性推荐 (Personalized Recommendation)**: These components feed into the **关键词 (Keywords)** input.
- 新的搜索历史 (New Search History)**: Receives input from the **搜索输入 (Search Input)** component.
- 搜索输入 (Search Input)**: Receives **关键词 (Keywords)** and provides **输入内容 (Input Content)**, **筛选条件 (Filter Conditions)**, and **搜索类型 (Search Type)** to the **搜索结果列表 (Search Results List)**.
- 搜索结果列表 (Search Results List)**: Displays results and receives **赞/踩 (Like/Dislike)** feedback. It provides **选中文档/内容 id (Selected Document/Content ID)** to the **内容详情 (Content Details)** component.
- 知识图谱 (Knowledge Graph)**: Receives **关键词 (Keywords)** and provides **转换 (Conversion)** feedback. It also receives **用户输入 (User Input)**.
- AI 辅助 (AI Assistance)**: Receives **用户输入 (User Input)** and provides **文档内容 (Document Content)** and **内容总结 (Content Summary)** to the **内容详情 (Content Details)** component.
- 内容详情 (Content Details)**: Provides the final **用户 (响应) (User (Response))**.

The flow is summarized as: **用户 (请求) → ... → 用户 (响应)**.

39

8. ER 图

知识图谱相关数据表的ER图：

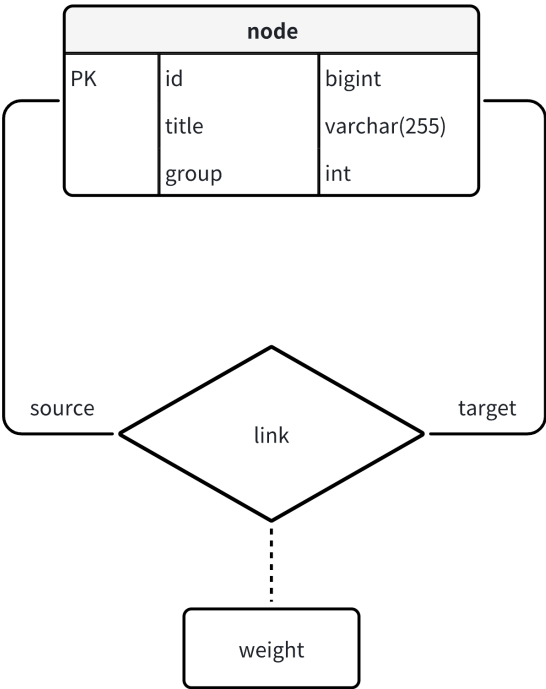


图8 ER图

Elastic Search采用非关系型数据库存储，无ER图。

9. 外部接口需求

9.1 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 i5 芯片及以上
内存	4G 及以上
存储	1T 及以上
网卡	速率 10Mbps 及以上
备份	RAID 5

- 服务器软件接口需求

项目	信息
操作系统	Linux Ubuntu 22.04
服务器	Nginx
搜索引擎框架	ElasticSearch
数据库	MySQL

9.2 客户端接口

- 客户端硬件需求

项目	信息
处理器	Intel i5 芯片及以上
内存	2G 及以上
存储	任意主流存储介质
网卡	速率 10Mbps 及以上
备份	RAID 5

- 客户端软件需求

项目	信息
操作系统	Windows 7 及以上、Linux Ubuntu 18.04 及以上
浏览器	Chrome、Firefox、Edge等主流浏览器

9.3 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 及 ZJUWLAN-Secure，有条件时可进行适当扩展。

10. 非功能性需求

10.1 性能需求

a. 系统响应时间：

- 律镜法律领域垂直搜索引擎对用户请求的响应时间应控制在2秒以内，以确保用户体验流畅。
- 高并发情况下，系统的响应时间不应超过5秒，以保证在用户访问量增加时仍能提供稳定的服务。

b. 并发用户量：

- 系统应能支持最少1000名并发用户同时使用，以满足大量用户同时访问的需求。
- 在极端情况下，系统应有扩展性，支持最少2000名并发用户，以防止因用户数量激增而导致的系统性能下降。

c. 吞吐量：

- 律镜法律领域垂直搜索引擎的吞吐量应能够处理每秒至少100个请求，以确保高效的信息检索服务。
- 在高峰期，系统吞吐量需达到每秒至少200个请求，以满足用户在特定时间段的集中查询需求。

d. 系统稳定性：

- 系统运行连续时间应能够达到每月99.9%的可用性，确保系统在绝大多数时间内可供用户访问。
- 系统应具备快速恢复的能力，故障发生后系统可在30分钟内恢复正常运行。

e. 数据安全性：

- 律镜法律领域垂直搜索引擎应采用加密技术，确保用户的检索数据传输过程中的机密性。
- 用户数据在系统中的存储应采取安全措施，防止未经授权的访问或数据泄露。

f. 系统日志：

- 系统应记录用户的检索历史和操作日志，以支持系统运行状态的监控和故障排除。
- 日志记录应包括关键操作、异常事件、以及系统性能指标等信息，以便进行系统性能分析和优化。

g. 可扩展性：

- 律镜法律领域垂直搜索引擎应具备良好的可扩展性，能够灵活地应对未来业务规模的扩大。
- 在系统升级或功能拓展时，应保证系统的稳定性，不影响已有功能的正常使用。

以上性能需求将确保律镜法律领域垂直搜索引擎在使用过程中能够高效、稳定地提供服务，满足用户的期望和需求。

10.2 数据传输与并发要求

a. 数据传输效率：

- 律镜法律领域垂直搜索引擎应通过优化数据传输协议和压缩算法，确保在不同网络环境下，数据传输效率最大化。
- 数据传输时的平均带宽利用率应在80%以上，以保证在网络状况较差时仍能提供相对较快的响应速度。

b. 并发访问控制：

- 系统应实现合理的并发访问控制机制，确保在高并发环境下不同用户之间的数据不会混淆或受到干扰。
- 对于敏感数据的并发访问，系统应采取适当的锁机制，以保证数据的完整性和一致性。

c. 并发事务处理：

- 在并发事务处理方面，系统应支持多用户同时进行事务，并能正确处理并发事务带来的问题，如死锁等。
- 系统的并发事务处理性能应在高并发场景下保持稳定，避免事务阻塞导致系统性能下降。

d. 数据一致性：

- 在并发读写操作中，系统应确保数据的一致性，即使在高并发环境下，用户获得的数据也应是最新的。
- 对于数据更新，系统应采用合适的机制，如乐观锁或悲观锁，以维护数据的一致性。

e. 连接池管理：

- 系统应实现连接池管理机制，有效利用数据库连接资源，避免频繁的连接创建和销毁。
- 连接池应具备动态调整的能力，以根据系统负载和资源情况自动调整连接池的大小。

f. 流量控制：

- 律镜法律领域垂直搜索引擎应具备流量控制机制，防止恶意攻击和异常流量对系统的影响。
- 在系统超负荷情况下，应对请求进行合理的排队和拒绝处理，以保障系统的稳定性。

g. 数据加密与解密：

- 在数据传输和存储过程中，系统应采用合适的加密算法，确保敏感信息的安全性。
- 数据的解密和加密过程应对系统性能影响较小，不影响正常的用户体验。

以上数据传输与并发要求将确保律镜法律领域垂直搜索引擎在处理大量用户请求时能够高效、稳定地运行，同时保护用户数据的安全性。

10.3 数据管理要求

a. 数据安全性与独立性：

- 系统应确保与其他系统的接口能够安全交互，同时保持本系统的独立性和完整性。
- 未经授权的人员不得修改系统设置、访问系统内部数据，确保系统的安全性和数据保密性。

b. 数据备份与恢复：

- 系统服务器软件必须提供可靠的数据备份和恢复手段，确保在软件或硬件故障时能够迅速恢复正常运行环境。
- 软件开发者不得在系统中预留任何特殊账户和密码，以确保系统的安全性。

c. 安全性细则：

- 加密登录：系统应实现用户登录信息的加密存储和传输，确保用户身份的安全性。
- 数据加密传输：在数据传输过程中应使用安全的加密算法，保障数据在传输过程中的保密性和完整性。

具体细则**a. 系统恢复时间：**

当系统崩溃后，系统应能在24小时内完成数据恢复和重新运行。

b. 数据库容量：

数据库应支持单表最大行数达到600行，以适应系统的数据存储需求。

c. 数据增长速度：

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和业务数据量决定。

d. 重大事故数据恢复：

当出现重大事故导致数据丢失时，系统应在48小时内完成数据的恢复。

e. 备份维护频率：

系统管理员每两个月应至少维护一次备份，以确保备份数据的及时性和可靠性。

f. 服务器存储空间：

系统服务器应具备至少20GB的存储空间，以满足系统运行和数据存储的需求。

以上详细规定将确保律镜法律领域垂直搜索引擎在数据管理方面能够达到高水平的安全性、稳定性和可维护性，保障系统的正常运行和用户数据的安全。

10.4 输入要求**a. 用户友好性：**

提供直观、清晰的用户界面，使用户能够轻松理解和操作系统。

b. 输入格式验证：

对用户输入的数据进行格式验证，包括但不限于日期、数字、文本等，确保输入的数据符合系统要求。

c. 输入响应时间：

保持用户输入的响应时间在2秒以内，以提升用户体验。

d. 错误处理提示：

当用户输入错误时，给予明确的错误提示，并提供解决方案或建议，以引导用户纠正错误。

e. 输入安全性：

对用户输入的数据进行严格的安全性检查，防范常见的输入安全漏洞，如SQL注入、跨站脚本攻击等。

以上非功能性需求将确保律镜法律领域垂直搜索引擎在用户输入方面能够提供友好、安全、有效的操作体验，同时保障系统的稳定性和数据的完整性。

10.5 可视化需求

a. 搜索结果可视化：

系统应以直观的图形方式呈现搜索结果，包括但不限于图表、关系图等，以方便用户快速理解检索数据的相关性。

b. 法律关系图谱：

提供法律领域的关系图谱可视化，展示法条、案例、法规等之间的关联，帮助用户深入理解法律体系的复杂关系。

c. 趋势分析图表：

支持搜索结果的趋势分析，以图表形式展示法律领域的发展趋势、热点问题变化等，帮助用户了解法律动态。

d. 用户操作反馈：

在用户进行搜索、筛选等操作时，提供直观的反馈动画或提示，增强用户对系统操作的可感知性。

e. 主题词云展示：

搜索结果应以词云的形式展示主题关键词，帮助用户迅速把握搜索结果的核心信息。

f. 支持多维度过滤：

提供可视化的多维度过滤界面，使用户能够直观、灵活地进行信息筛选，提高搜索效率。

g. 实时数据更新：

在可视化展示中，确保数据的实时更新，反映法律领域的最新动态和变化。

以上需求将使律镜法律领域垂直搜索引擎在可视化方面更具吸引力和实用性，为用户提供直观而深入的法律信息呈现。

10.6 防护性需求

a. 用户身份认证：

实施严格的用户身份认证机制，包括但不限于多因素认证，确保只有授权用户能够访问敏感信息和功能。

b. 防止恶意攻击：

部署防火墙、入侵检测系统等安全设备，有效防范网络攻击、恶意扫描等安全威胁。

c. 数据加密传输：

所有用户与系统之间的数据传输应采用安全的加密协议，保障数据在传输过程中的机密性。

d. 防范SQL注入和XSS攻击：

实施有效的输入验证和过滤机制，防范SQL注入和跨站脚本（XSS）等常见攻击手段。

e. 系统审计与监控：

建立系统审计和监控体系，记录用户操作日志、系统异常日志等，及时发现并应对潜在的安全问题。

f. 安全升级机制：

提供安全升级机制，确保系统能够及时应对已知安全漏洞，保障系统的安全性。

g. 权限控制：

实施严格的权限控制，确保用户只能访问其授权范围内的数据和功能，防止越权访问。

h. 安全培训与意识：

开展定期的安全培训，提高团队成员对安全问题的意识，减少内部安全风险。

i. 敏感数据脱敏：

对于敏感数据，采用脱敏技术，以减少数据泄露的潜在风险。

j. 安全漏洞响应：

设立安全漏洞响应团队，确保在发现安全漏洞时能够迅速做出应对措施，及时通知用户并修复问题。

细化与处理措施

a. 数据库误删除修复机制：

当发生数据库误删除情况时，系统应实施撤销删除修复机制，允许管理员在一定时间内恢复被误删除的数据。

b. 下载权限保护：

对于未开放下载权限的资料，系统应采用严格的权限控制，确保未经授权的用户无法下载相关资料。

c. 防重复操作警告：

在用户进行可能导致重复操作的情况下，系统应给予警告提示，防止用户误操作导致系统卡死。

d. 验证码防恶意使用：

实施验证码机制，特别是在用户频繁请求、操作时，要求用户输入验证码，以防止恶意使用，如爬虫等。

e. 及时信息备份：

定期进行系统数据备份，并确保备份的完整性。在发生病毒攻击等情况下，系统能够及时恢复至备份状态。

f. 检测恶意操作与警告：

实施行为分析系统，能够检测到可能的恶意操作，对于异常操作行为，系统应提出警告并在一段时间内限制用户操作。

g. 访问无权限警告与禁止：

当用户尝试访问无权限的资源时，系统应发出明确的警告，并禁止用户继续访问受限资源。

以上防护性需求将确保律镜法律领域垂直搜索引擎在安全性方面具备全面的保护机制，有效预防潜在的安全威胁和风险。

10.7 软件质量属性

a. 性能优化：

实现快速的搜索响应时间，确保用户在进行搜索和操作时能够获得即时的反馈，响应时间控制在2秒以内。

b. 可伸缩性：

支持系统的横向扩展，以应对用户数量的增加，保障系统在高负载时仍能保持稳定性和性能。

c. 可维护性：

采用模块化和清晰的代码结构，使系统具备良好的可维护性，方便团队进行后续的功能扩展和修改。

d. 可靠性与稳定性：

系统应保持高可靠性，最小化系统崩溃和故障的发生，保障用户能够稳定、持续地使用系统。

e. 灵活性与定制化：

提供丰富的搜索筛选选项，允许用户定制搜索参数，提高搜索的灵活性，以满足不同用户的个性化需求。

f. 兼容性：

确保系统在不同浏览器和平台上的兼容性，使用户能够在各种设备上流畅使用系统。

g. 安全性与隐私保护：

保障用户数据的安全性，采用加密手段保护敏感信息，确保用户隐私不受侵犯。

h. 易用性与用户体验：

设计简洁直观的用户界面，提供良好的用户体验，降低用户的学习成本，使系统易于上手。

i. 可测试性：

提供有效的单元测试、集成测试和系统测试机制，以确保系统的稳定性和功能的正确性。

j. 故障恢复与紧急处理：

实现系统的自动故障恢复机制，及时处理系统异常，减小故障对用户的影响。

细化要求

a. 兼容性：

系统只需确保在主流浏览器（如Chrome、IE）上正常浏览和使用，提供一致的用户体验。

b. 易用性：

界面设计应简洁明了，操作简单且符合用户日常习惯。系统应包含导航功能，并提供清晰简短的用户使用手册，以便用户更好地理解系统操作和功能。

c. 可用性：

系统保证每天早上6点到晚上24点之间可用。在发生紧急情况需要维护时，允许进行一定时间的维护操作，确保用户在维护期间得到通知。

d. 可维护性：

系统运行时保存运行日志，以便进行维护和分析。每两周的维护日定于凌晨1点到5点，用户在此期间不能使用系统。

e. 可扩充性：

系统在设计上考虑到了网站可能的后续发展，尽可能在满足所有需求的同时增强了网站的可扩充性。对于扩展需求，客户可以联系系统维护人员，维护人员需在1-4个工作日内完成客户的内容扩充需求，包括但不限于增加新功能、增加新模块、界面优化、系统性能提升等。

以上软件质量属性将确保律镜法律领域垂直搜索引擎在性能、可维护性、可靠性、安全性等方面具备高水平的软件质量，提升用户体验和系统可用性。

10.8 其他需求

a. 隐私保护：

系统需遵循相关法律法规，保护用户隐私，不得未经授权收集、使用用户个人信息。提供隐私政策并明确用户数据的使用范围。

b. 多语言支持：

系统应提供多语言支持，确保用户可以选择使用他们偏好的语言进行搜索和浏览。

c. 离线模式：

提供离线模式，允许用户在无网络连接时访问已缓存的搜索结果和信息。

d. 反垃圾保护：

实施反垃圾机制，防止恶意爬虫、刷点击等行为，确保搜索结果的真实性和可信度。

e. 系统通知与提醒：

支持系统通知和提醒功能，向用户发送重要信息、更新和个性化提醒，提高用户参与度。

f. 访问日志保留：

系统应保存用户访问日志，以便进行系统性能分析和问题排查。

g. 自动更新机制：

实现系统的自动更新机制，确保用户使用的始终是最新版本，包括功能改进和安全性修复。

以上其他需求将进一步丰富律镜法律领域垂直搜索引擎的功能和用户体验，保障系统的稳定性、可用性和用户满意度。

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	用户仅能搜索法律相关的信息	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	动作触发规则	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法律知识有关	约束	静态	用户
-------	--------------------	----	----	----

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

界面原型部分，我们针对每个界面的功能设计了界面的低保真灰模图（即用灰度区块表示不同的功能分区），与系统设计报告中的页面设计部分不同，在这些图中，重点不在于页面外观的美观度与交互性设计，而重在体现页面上的功能分区。

本系统中的主要页面设计如下：

12.1 搜索引擎首页

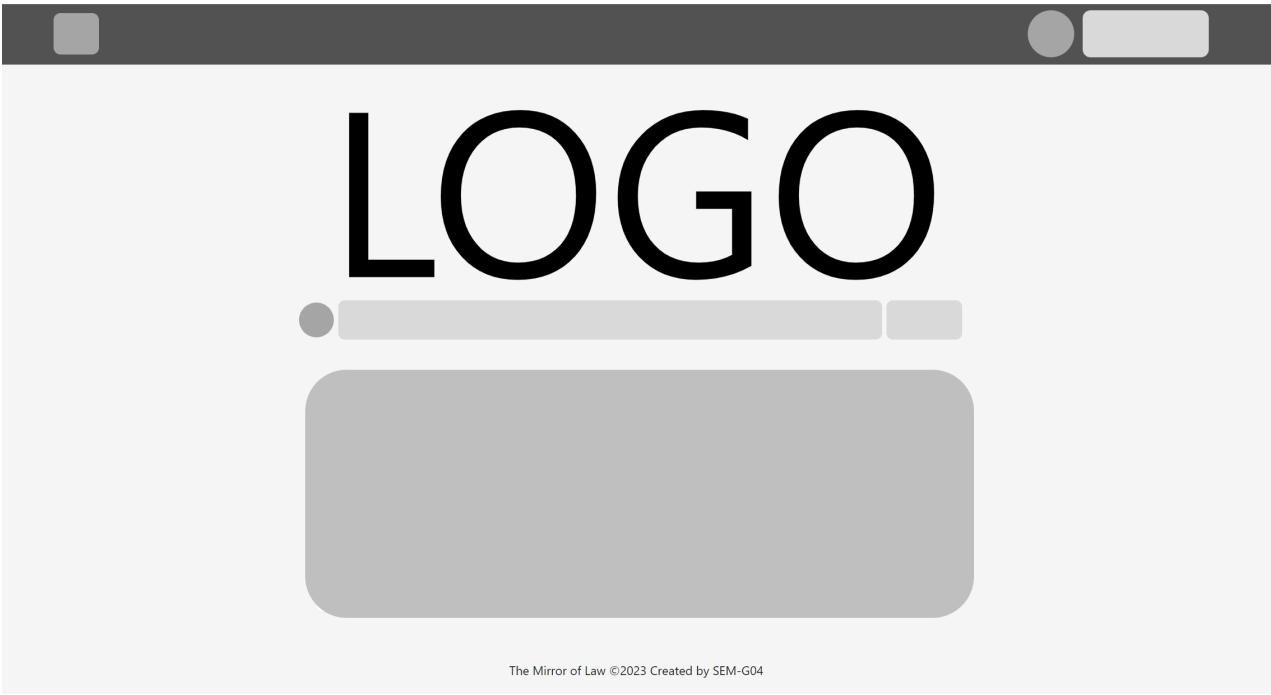


图12-1 搜索引擎首页原型设计

12.2 搜索结果列表页

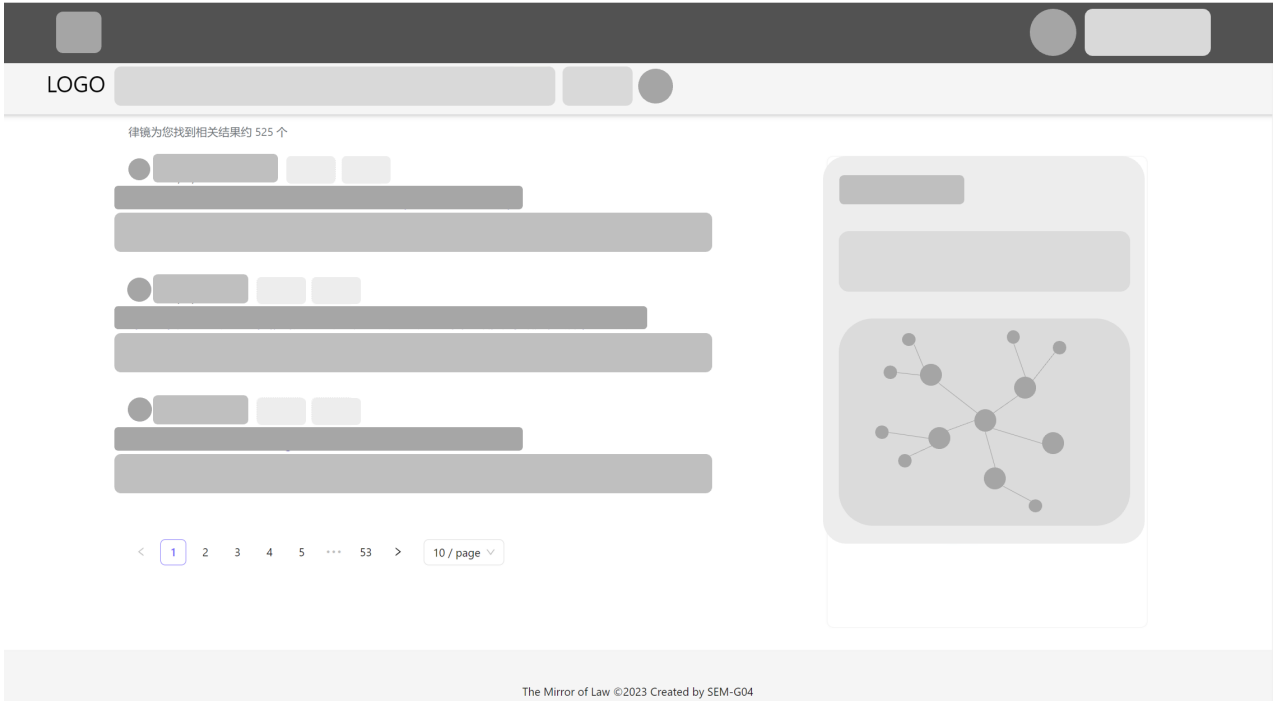


图12-2 搜索结果列表页原型设计

12.3 详情页



图12-3 详情页原型设计

13. 附录

13.1 术语表

a. Elastic Search

Elasticsearch是一个开源的分布式搜索和分析引擎，用于在大规模数据集中进行快速、实时的搜索和分析。它基于Apache Lucene搜索引擎构建，通过提供简单的RESTful API和丰富的查询语言，使用户能够轻松地在各种类型的数据中执行复杂的搜索和分析操作。Elasticsearch被广泛用于构建实时应用程序，如日志分析、全文搜索、业务分析等，因其高性能、可扩展性和容错性而备受欢迎。

b. 垂直搜索引擎

垂直搜索引擎是一种专注于特定主题或领域的搜索引擎。与通用搜索引擎（如Google或Bing）不同，垂直搜索引擎致力于提供深度而专业化的搜索结果，以满足特定领域的用户需求。这些领域可以涵盖特定行业、主题、媒体类型或其他特定的信息范畴。通过专注于特定领域，垂直搜索引擎可以提供更准确、有针对性的搜索结果，帮助用户更有效地获取他们所需领域的信息。

c. 爬虫

爬虫（Spider）是一种程序或脚本，被设计用来自动地在互联网上获取信息。爬虫的主要任务是通过遍历网络上的页面并提取有用的数据，以便后续的分析、索引或展示。爬虫通常按照一定的规则和算法遍历网页，跟随链接，然后从页面中提取感兴趣的信息，如文本内容、链接、图像等。

爬虫在各种应用中都有广泛的用途，包括搜索引擎的索引过程、数据挖掘、舆情监测、价格比较、信息聚合等。然而，爬虫的使用也面临一些道德和法律问题，因为它们可能会对网站的性能产生负面影响，甚至触发反爬虫机制。因此，在使用爬虫时需要遵循网站的规定并尊重网络礼仪。

d. 知识图谱

知识图谱是一种用于表示、组织和推理知识的图状数据结构。它是一种语义网络，通过节点和边的连接关系表示实体之间的语义关联。知识图谱旨在捕捉现实世界中的实体和它们之间的关系，以便计算机能够更深入地理解 and 处理信息。

13.2 待确定问题清单

a. 服务器证书问题

为了方便同步数据库，我们将Elasticsearch部署在服务器上，但该服务器证书为自签名证书且目前只签名为localhost，导致后端远程连接数据库的时候出现SSL错误，证书有效性和主机名验证失败，有待进一步解决。

b. 摘要中高亮部分处理问题

在搜索结果列表中，后端默认会返回一段包含搜索关键字的上下文摘要，其中搜索的关键字会被标签高亮显示。但出于安全考虑，后端返回的数据在前端会被视为字符串进行解析而非直接渲染，因此面临标签难以解析的问题，需要有更好的方法实现摘要中搜索关键字的高亮功能。