

VidBot

——基于视频内容挖掘与播放流量统计的视频智能学习工具



目录

一、项目简介	-----1
(一) 项目背景	
(二) 项目实施计划	
二、用户范围	-----4
(一) 视频观看者	
(二) 视频上传者	
三、市场调研和竞品分析	-----4
(一) 市场需求调研	
(二) 竞品分析	
四、总体设计	-----5
(一) 本地版概述	
(二) 在线版概述	
五、详细设计	-----7
(一) 功能设计	
(二) 数据结构设计	
(三) 界面设计	
六、创新点	-----15
七、操作指南	-----17
八、即将上线功能	-----17
九、历史版本	-----17

一、项目简介

VidBot，视学知识助手。它提供视频流量分析、视频转知识导图、视频转思维导图、视频自动分段、智能助教和共享笔记等功能，旨在改善您的在线视频体验和学习过程，帮助视频上传者了解观众行为和优化内容。无论您是教育工作者、内容创作者，还是渴望更深入理解视频内容的学习者，VidBot 都将成为您的不可或缺的工具。

（一）项目背景

近年来随着信息技术迅速发展，知识获取的方式发生了根本变化，在线教育通过信息技术改变和教育行业的运作方式，优化了流程和模式，从而提高教与学的效率。前几年在新冠疫情的影响下，国家推出“停课不停学”的政策，加速了在线教育行业的渗透率。随着国民可支配收入的提升，互联网普及度的发展，在线教育行业实现了新一轮的增长。2020 年在线教育用户规模达到 3.51 亿人，在线教育市场规模超过 4800 亿元。由此可见，线上学习正日益成为教育领域不可逆转的重要趋势。

然而，其弊端是不可忽视的。因为缺少面对面交流、参与感弱，走神概率大大提高、学习效果变差，教师也难以掌握学生观看视频的情况。针对这些弊端，VidBot 系列产品提出了相应解决措施，本地版和在线版都实现了视频各时段反复播放次数可视化，方便后来的学习者更好地抓取此视频的重难点并在这一部分集中精力；视频内容转化思维导图可以帮助学习者在观看视频前对内容有一个大体的把握，课后复习的时间也得以缩短。不仅如此，视频一经上传就会被自动分段，点击思维导图还会跳转相应位置，实现两项功能的有机结合。在线版还提供了智能助教和共享笔记等功能。简而言之，VidBot 系列产品的开发旨在帮助学习者提高学习效率、获得更好的学习体验，也为视频创作者提供用户观看热点和视频的分段标记。

“水杉在线”是数据科学与工程学院开发的一个学习平台，目前，“水杉在线”已经在数据科学与工程学院专业课和华东师大计算机公共课中全面使用，同时为 10 家以上的机构、1 万多注册用户提供了 100 多项在线课程或实训实验项目服务，访问人次超过 100 万。因此，系列产品的在线版选择首先在“水杉在线”平台落地测试使用。

（二）项目实施计划

2023.5-2023.7：确定在线版和本地版的功能。创建功能规范文档，定义详细功能列表和用户界面设计，产品命名为 StreamTextPro。

2023.7-2023.9：完成本地版视频转文本、文本转知识图谱、文本转思维导图、视频自动分段和思维导图自动跳转等多个模块。

2023.10-2023.11：对本地版进行优化，完成插件相应功能的开发，新增智能助教和共享笔记功能，部署到云服务器，产品更名为 VidBot。

2023.12-2024.1：计划发布系列产品，对用户反馈的需求进行针对性优化，新增部分相关功能。

2024.2-2024.3：计划对每个部分涉及到的算法与模型进行优化，选择部分领域深入探索，在部分功能上性能和准确率争取实现突破。

2024.4-2024.5：计划进行功能测试和调试，修复可能存在的错误和问题。进行性能测试和负载测试，确保产品在各种情况下都能正常工作。

二、用户范围

（一）视频观看者

观看教学视频的用户可能需要快速定位到视频中自己想看的知识点，还能获得一份基于视频内容的思维导图和知识图谱并了解别人反复观看的部分，如果对视频内容难以理解还可以向智能助教提问，如果看视频需要记一些笔记可以记录在共享笔记。

（二）视频创作者

上传教学视频到平台的用户可以借助这个插件实现长视频的自动分段，如果用户需要剪辑，可以参考这些分段。智能助教是教师的延伸，辅助提升教学效果。

三、市场调研和竞品分析

（一）市场需求调研

在接受调研的对象中，95.03%希望获取教学视频的内容思维导图，98.34%希望将视频内容进行分段，并标注在进度条上，64.64%想要获得教学视频的知识图谱，95.58%需要反复观看特定部分，79.56%认为播放次数多的片段更可能是视频的重点。在访谈中大部分受访者都提到了希望增加 AI 问答和笔记功能。

（二）竞品分析

功能对比	B 站	腾讯会议	飞书	BibiGPT	十行笔记
视频流量统计	无（仅有弹幕统计）	无	无	无	无
视频转知识图谱	无	无	无	无	无
视频转思维导图	无	无	有	有	有（但需用户上传）
视频自动分段	无（仅有手动分段）	有(针对录制的会议)	无	无	无
智能助教	无	无	无	无	无
共享笔记	无	无	无	无	无

四、总体设计

（一）本地版概述

开发本地版，允许用户上传本地视频进行分析。集成本地 whisper 模型和大语言模型，将视频内容转化为思维导图和知识图谱。实现思维导图的交互式显示，允许用户点击不同部分以跳转到相关视频内容。在视频进度条上自动显示分段小标题，以便用户了解视频的内容结构。统计视频分时段的观看次数，并将其以可视化方式展示在进度条上。



（二）在线版概述

开发浏览器扩展，允许用户在网页上选择和分析在线视频内容。集成本地 whisper 模型和大语言模型，将选定的网页视频内容转化为思维导图和知识图谱。实现思维导图的交互式显示，允许用户点击不同部分以跳转到相关视频内容。在视频进度条上自动显示分段小标题，以便用户了解视频的内容结构。统计视频分时段的观看次数，并将其以可视化方式展示在进度条上。加入与大模型的对话功能（智能助教）和共享笔记，方便用户随时提问和记录。

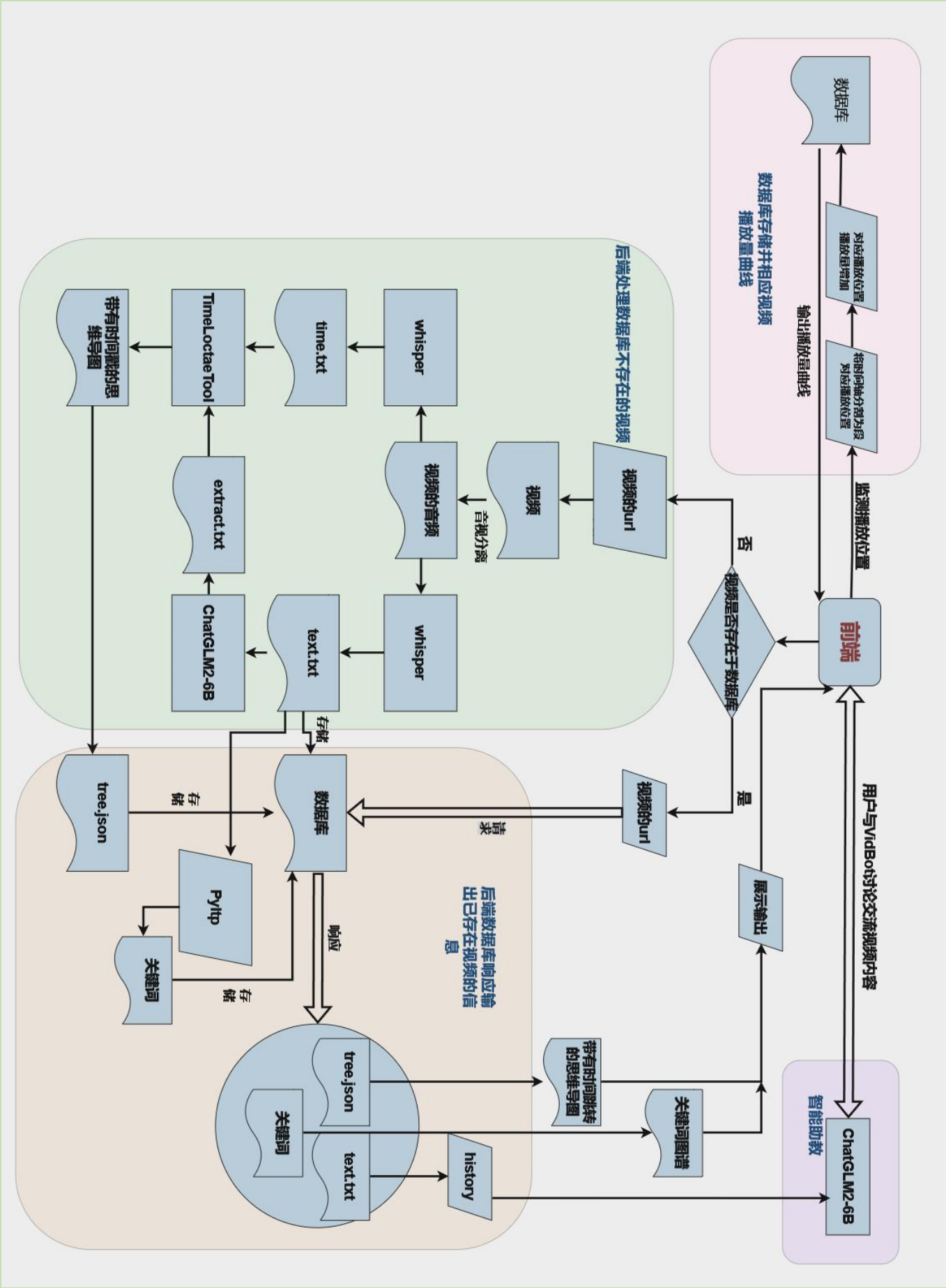
The screenshot displays the Shuishan Online Learning Platform (水杉在线) interface. The main content area shows a video player titled "大数据5V" (Big Data 5V) with a progress bar and a generated mind map. The mind map is centered on "大数据 5V 特征" (Big Data 5V Features) and branches out into five categories: 数量 (Volume), 速度 (Velocity), 种类 (Variety), 价值 (Value), and 真实性 (Veracity). Each category lists specific characteristics. For example, under "数量 (Volume)", it lists "数据量巨大" (Large data volume), "数据类型多样" (Diverse data types), and "数据增长迅速" (Rapid data growth). Under "速度 (Velocity)", it lists "数据生成速度快" (Fast data generation), "数据处理速度快" (Fast data processing), and "数据更新速度快" (Fast data update). Under "种类 (Variety)", it lists "数据类型多样" (Diverse data types), "数据来源多样" (Diverse data sources), and "数据格式多样" (Diverse data formats). Under "价值 (Value)", it lists "数据价值高" (High data value), "数据应用广泛" (Wide data application), and "数据决策支持" (Data decision support). Under "真实性 (Veracity)", it lists "数据准确性高" (High data accuracy), "数据可信度高" (High data credibility), and "数据一致性高" (High data consistency). The sidebar on the left shows the course structure, including "理论课 绪论" (Theory Course Introduction) and "1.1.1 什么是分布式系统" (What is a Distributed System?). The top navigation bar includes links to "首页" (Home), "水杉学堂" (Shuishan Academy), "水杉工坊" (Shuishan Workshop), "开源教育计划" (Open Source Education Plan), and "水杉校场" (Shuishan Campus).

五、详细设计

——功能、数据结构、页面

(一) 功能设计

项目流程图



视频观看流量可视化

该功能允许观看者和创作者了解观众行为。通过这项功能，用户可以查看有关视频的观看数据，这些数据将以图形的形式呈现，使用户能够清晰地定位观看热点。

视频总结生成思维导图

该功能使用户能够生成视频总结的思维导图。产品会分析视频内容，提取关键信息，并将其以思维导图的形式呈现。这一思维导图将有助于用户更好地理解视频的结构和主要概念，允许他们快速导航到感兴趣的部分，以及更好地理解视频内容。

视频总结生成知识图谱

该功能可以生成有关视频内容的知识图谱。插件将分析视频内容并构建一个关于视频主题的知识图谱，其中包含高频词和关键词。用户可以深入了解视频内容，获取相关知识。

视频自动分段

该功能自动将视频划分为不同的部分，并为每个部分生成小标题。这些分段小标题将在视频进度条上自动显示，使用户能够更轻松地了解视频的结构，快速跳转到感兴趣的部分，以及更好地导航和回顾视频内容。

智能助教

该功能为用户学习提供了多方面的支持和互动。包括回答用户的问题，提供定制化建议，并根据用户的需求和学习风格提供个性化的学习支持。智能助教帮助用户更深入地学习视频内容，解决问题，并更高效地实现其学习目标。

共享笔记（在线编辑器）

该功能根据视频内容自动生成笔记，包括思维导图等可编辑文本，用户可以在编辑器修改笔记，生成自己个性化的笔记，可以上传水杉仓库，形成大家共同完善的笔记，实现知识共享，老师也可以在上传视频以后获得一份自己课程内容的思维导图，这样使得课堂更加自由且可回溯。

(二) 数据结构设计

1. 数据库设计

采用SQLAlchemy 在 Python 中操作关系型数据库,数据库中主要实体字段说明:

视频转化结果信息表:

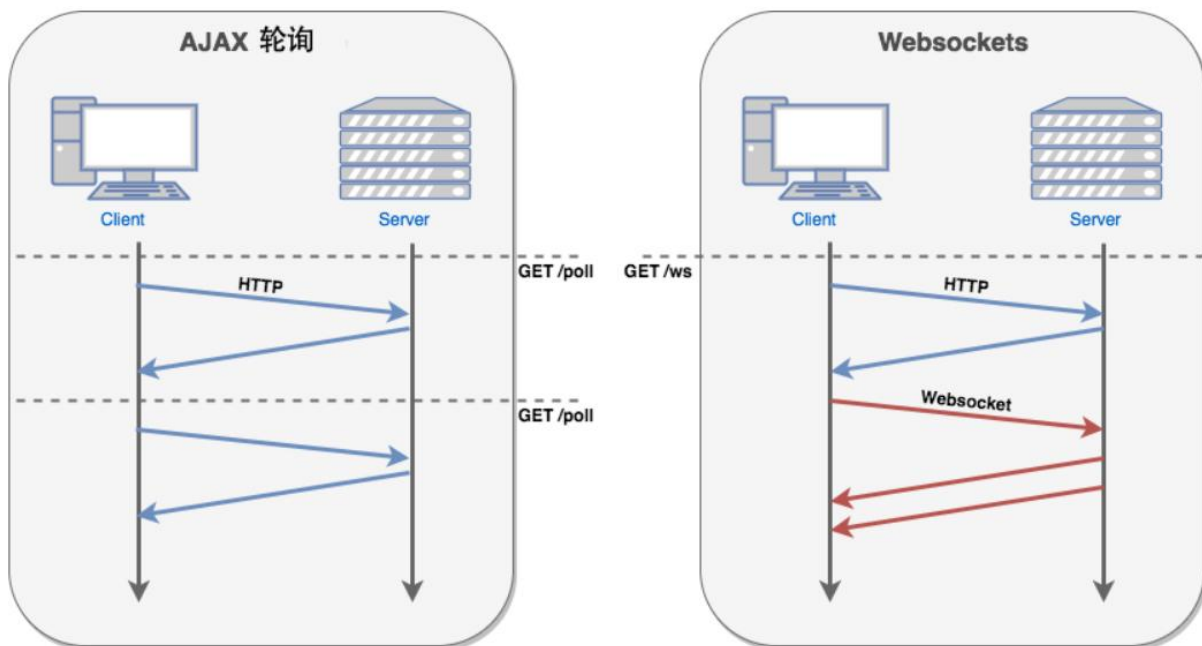
数据表名: transform_result						
字段名称	字段描述	数据类型	长度	是否允许为空	缺省值	备注
video_id	主键	bigint	256	N	NULL	通过视频 url 正则化
mindmap	思维导图路径	varchar	512	N	NULL	
graph	知识图谱路径	varchar	512	N	NULL	
summary	视频概括路径	varchar	512	N	NULL	

视频流量统计表:

数据表名: video_flowrate						
字段名称	字段描述	数据类型	长度	是否允许为空	缺省值	备注
video_id	主键	bigint	256	N	NULL	通过视频 url 正则化
time_stamp	时间点	int	11	N	NULL	将视频均分为 100 段
number_of_views	观看数量	int	11	N	NULL	统计观看次数

2.通信设计

前后端通信:



WebSocket 通信:

使用 WebSocket 来支持前端与后端的实时通信。前端将视频 URL 发送给后端，后端处理请求并从数据库中检索思维导图 JSON 数据、知识图谱 JSON 数据和观看流量数据发送回前端。智能助教则需要实时传输问答。

Flask 框架:

后端采用 Python 的 Flask 框架，接受前端的请求并向前端发送数据。

前端内部页面通信:

Content.js 与嵌入页面脚本之间的相互通信:

Content.js 通过 `chrome.runtime.sendMessage` 方法向嵌入页面脚本发送消息。嵌入页面脚本通过 `chrome.tabs.sendMessage` 方法向特定标签页的内容脚本发送消息。Content.js 与嵌入页面脚本之间通过 `chrome.runtime.onMessage` 方法监听其它脚本发送的消息。采用结构化和可维护的 Promise 对象来处理不同链接之间通信的异步性，保证 Content.js 与嵌入页面脚本对页面的正常渲染。

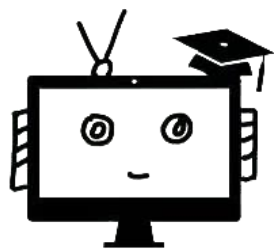
(二) 界面设计

1.界面风格设计

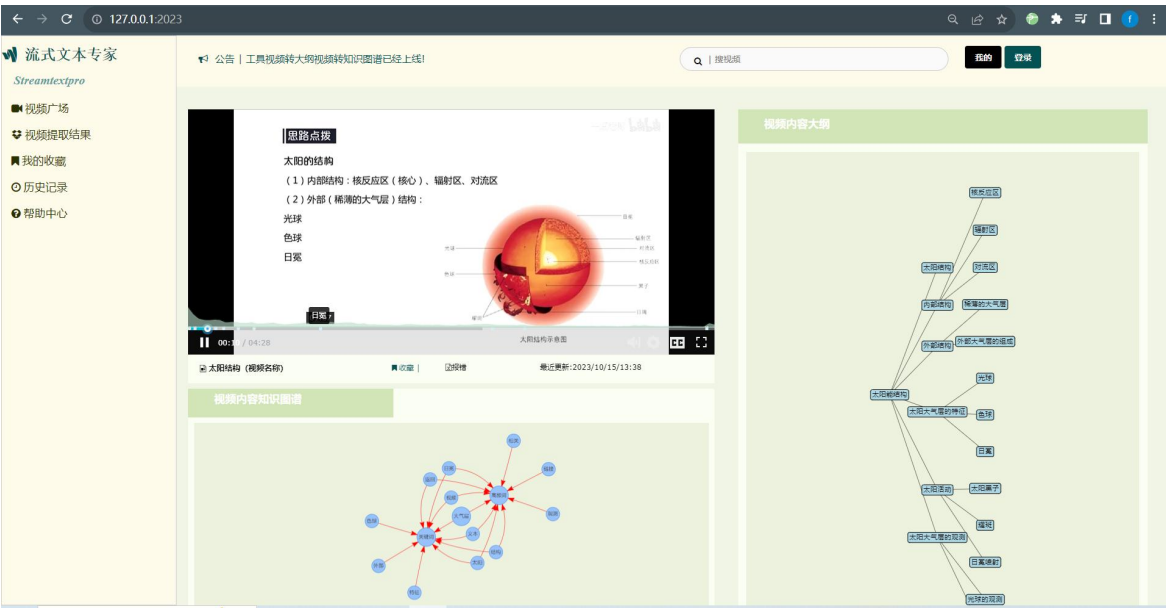
本项目的图形用户界面主要为网页，分为三个主要页面：本地界面、插件启动页、插件内容页。VidBot 系列产品以绿色为主题，旨在为用户提供轻松愉快的体验，同时增强产品的专业感和吸引力。这一主题传达了和平、轻松和积极的情感，且有助于提高产品的识别度。此外，水杉在线也以绿色为主题，整体配色协调一致。

2.主要功能界面

项目图标



本地版主页



该页面展示出本工具的四大功能：视频进度条上展示历史观看次数、自动打点并添加段落标签、中下方展示知识图谱、右侧展示思维导图。点击右侧思维导图可以跳转到视频相应位置。

插件界面

功能总览



所有生成的内容展示在内容界面，可以通过切换选项卡展示不同功能。

popup 页面



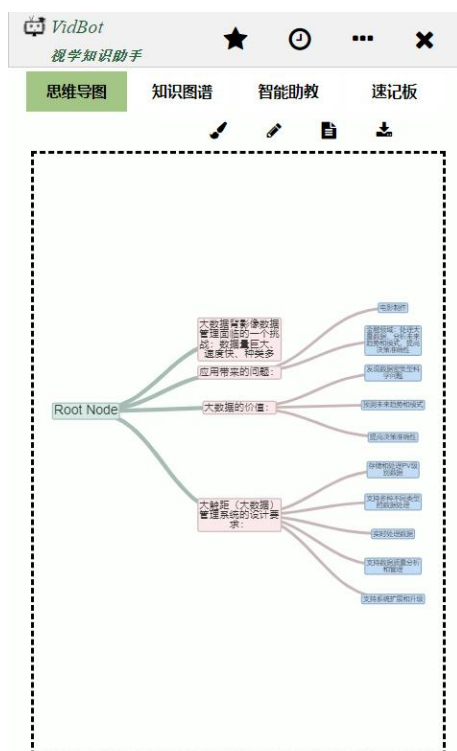
点击一键总结可唤醒内容界面，当用户对结果不满意时可以重新生成。

视频播放页

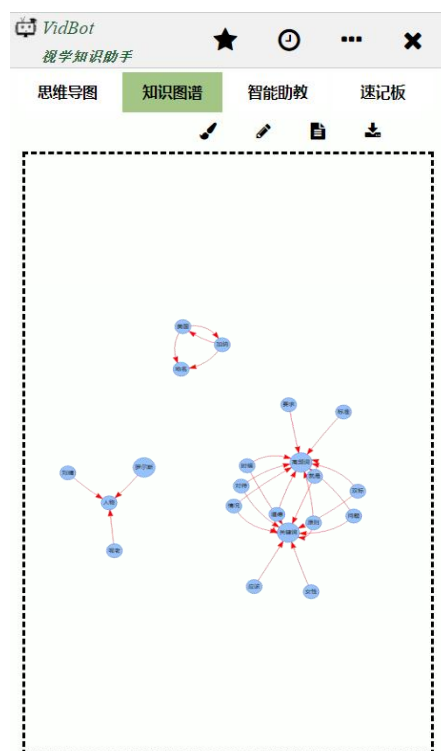


插件开启时即可对观看流量进行统计和展示，播放视频时可以看到视频的分段标签。

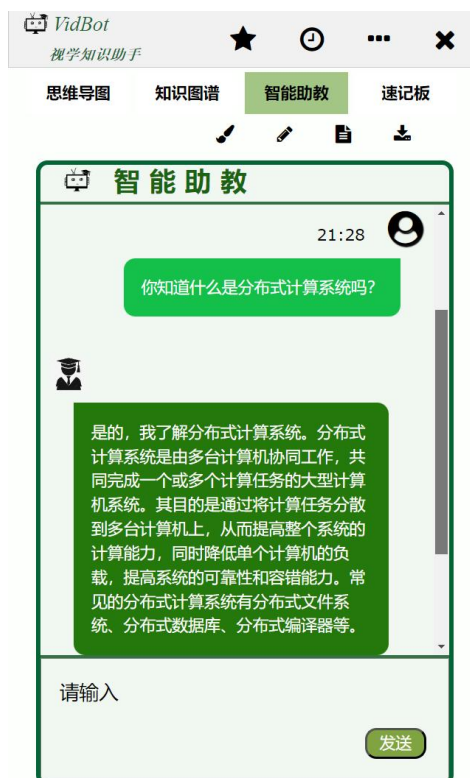
思维导图



知识图谱



智能助教



共享笔记（在线编辑器）



思维导图页：点击思维导图对应标签跳转到视频相应内容的位置。

知识图谱页：交互式界面，选中节点可拖拽并显示关联。

智能助教页：基于视频内容向智能助教提问，模型可生成实时问答。

共享笔记页：内置 markdown 编辑器，用于记录学习笔记，可以一键导出。

六、创新点

本地模型部署

将视频分析和文本处理模型本地部署在系统中。这种本地部署减少了对第三方 API 的依赖，提高了系统的可控性和安全性。用户不必担心外部服务的可用性，同时也能更好地管理和保护敏感数据。

英特尔 bigdl.llm 实现量化

使用了 bigdl.llm 使 ChatGLM2-6B 根据需求进行量化，并且使其在 cpu 上进行推理，降低了本地化运行的门槛。

清华大学 ChatGLM2-6B 模型

使用清华大学 ChatGLM2-6B 模型实现文本概括是项目的重要创新点。这个大型语言模型能够生成流畅、连贯的文本概括，使用户更容易理解视频的关键信息。同时，本地部署这个模型增强了系统的性能和速度。

交互式思维导图和知识图谱

项目引入了思维导图和知识图谱，其中知识图谱采用哈工大 LTP 模型工具生成，以视频高频词和关键词为图谱核心，展示视频内容要点；思维导图允许用户通过点击不同部分来跳转到视频的相关位置。这种交互式导航提供了更直观的方式来探索视频内容，使用户能够深入研究他们感兴趣的部分。

分段小标题和观看统计

在视频进度条上显示分段小标题和统计视频分时段的观看次数是另一个创新。这不仅能帮助用户快速了解视频内容的结构，还为视频创作者提供了观众热点行为信息，有助于其更好地了解他们的观众。

智能助教和共享笔记

智能助教具备回答用户提出的问题和解决问题的能力。当用户在观看视频时遇到困难或疑问时，他们可以向智能助教提出问题，获得即时的答案和解释。这种互动性将加强用户对学习内容的理解。在提问的过程中，观看者可将灵感记录在速记板上，共享笔记作为一个 markdown 编辑器方便用户排版，最后可以一键导出和共享。

性能优化和数据库支持

引入性能优化措施并结合数据库支持是项目的另一个创新。这些措施提高了系统的速度和稳定性，确保用户能够流畅使用产品，而不受性能问题的干扰。

九、操作指南

现有部署方式：将代码拉取到本地，打开浏览器扩展(chrome 或 Edge)，将 VidBot 文件夹加载到浏览器上，进入水杉视频播放界面，开启扩展即可使用。

未来：上线浏览器扩展商店，用户可直接安装。本地版将在 Github 开源。

八、即将上线功能

视频知识点联动：同一个知识点可能在不同视频中被提及，计划在生成思维导图的时候标注还有哪些视频提到了这个知识点，尝试实现跳转其他视频。

知识点查询：输入想要学习的知识点，给用户返回含有这个知识点的视频链接。

智能出题、向视频创造者提问等功能

七、历史版本

本地版：

v0.1 StreamTextPro（调用 api 实现视频转文本+抽取式生成摘要+流量统计）

v0.2 StreamTextPro（调用 api 实现视频转文本+调用科大讯飞 api 生成摘要+流量统计）

v0.3 StreamTextPro

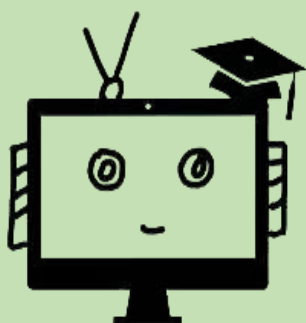
（部署本地 whisper 实现视频转文本+清华大学 ChatGLM2-6B 模型生成摘要+知识图谱+流量统计）

v0.4 StreamTextPro（新增英特尔 bigdl.llm 量化模型）

在线版：

v0.1 StreamTextPro（英特尔 bigdl.llm 量化模型+部署本地 whisper 实现视频转文本+清华大学 ChatGLM2-6B 模型生成摘要+知识图谱+流量统计）

v0.2 部署到服务器，更名为 VidBot（新增智能助教和共享笔记）



VidBot

项目贡献者（姓氏顺序）：

刘蔚璁 刘子玉 石季凡 谢沁桦 袁凡

Github: <https://github.com/VidBots>