**小组项目需求分析**

**网页端基础CAT工具**

**组员：段孝辰、秦浩洋、李善灡、潘光林**

# 一. 概述

## 1. 背景与目标

在技术与翻译逐步融合的大环境下，一批CAT工具先后出现，旨在帮助译员以更高的效率，完成客户提出的翻译任务。同时，这些工具还能帮助客户降低成本，为翻译公司扩大利润空间，从而达成一举三得。然而诸如RWS Trados、MemoQ等主流CAT工具，虽然功能强大，但使用复杂、售价高昂，这导致学习成本和使用成本都颇为高昂，使很多惯用基础文档编辑器，甚至是纸笔来进行翻译的译员望而却步。因此，我们开发本工具的目的就是为了能给入门用户提供一个学习CAT工具的基础平台，并让CAT工具惠及更多翻译从业者。

## 2. 工具概述

本工具为网页端应用，面向使用者为CAT工具的初学者。本工具旨在实现主流CAT中的主要功能，包括文件处理、翻译、多人项目管理等。要求所需功能实现完善、易用，使用逻辑符合基础用户的习惯。

## 3. 小组成员

段孝辰、秦浩洋、李善灡、潘光林

# 二. 功能需求

|  |  |  |
| --- | --- | --- |
| **文件模块** | 导入/导出 | 多文件格式导入：常见的翻译文件格式包括docx，html，xlsx，pdf等。该工具将支持读写其中主流的一些格式（pdf，xlsx，docx），并提供导入导出的功能。导入时将识别文件中需翻译的文字部分，并在导入后的文档中予以保留。导出时将提供一个xlsx格式的双语对照表格。 |
| 翻译内容识别：在本工具中导入待翻译文件时，需要剔除文件中无需翻译的部分、识别文件中需翻译部分的文字内容，并在导入后的文档中予以保留。 |
| 翻译内容分句：在本工具中导入待翻译文件时，需要将文件中的需翻译内容按句分割（segmentation），在翻译界面中予以呈现。 |
| TM重合率统计：在翻译界面中TM将在右侧栏中显示，并附有TM与当前句子的匹配度 |
| **翻译模块** | 翻译 | 复制原文和标签至译文：在翻译页面上方的功能区有“复制到译文框”按钮，点击可以将选中句段的原文全部复制到译文区中。 |
| 条目翻译确认：翻译完成后，按enter可将当前条目保存在翻译记忆库中。 |
| 机器翻译：点击某一词条，在翻译结果栏会自动显示来自百度翻译的机器翻译结果，以供参考。 |
| 在线查询：翻译过程中还可以使用在线查询功能。用户在任意句段内选中文字，点击在线查询按钮，即可弹出百度、谷歌、结果。 |
| 导入/导出 | 可以导出双语对照的xlsx文件存储在本地。 |
| **TM模块** | 导入/导出 | 可以导入/导出xlsx格式的文件，形成新的翻译记忆表。同时在平台上创建的TM也可以导出为xlsx格式在本地进行保存。 |
| 编辑 | 条目确定自动添加：翻译过程中，句段确认后自动添加到翻译记忆库中。 |
| TM独立编辑界面：翻译记忆库有独立编辑页面，可以手动添加、修改、删除某一条翻译记忆。一条TM对的属性包括源文本和目标文本。 |
| 匹配 | 点击某一句段，在翻译结果处会自动出现翻译记忆匹配结果，且会显示匹配率。 |
| **术语库模块** | 导入/导出 | 可以导入/导出xlsx格式的文件，形成新的术语表。同时在平台上创建的术语库也可以导出为xlsx格式在本地进行保存。 |
| 编辑 | 添加术语：添加术语。在翻译过程中，选中句段中的某一个词，可以点击“添加至术语库”按钮，在跳出的表单中输入术语翻译，即可添加至术语库。 |
| TB独立编辑界面：术语库有独立编辑页面，可以手动添加、修改、删除某一条术语。一条术语对的属性包括：源文本、目标文本、定义。 |
| 匹配 | 点击某一句段，在翻译结果处会自动出现翻译记忆匹配结果，且会显示匹配率。 |
|  |
| **账号模块** | 普通  译员 | 用户可以注册账号，以使用该工具功能对自己上传至数据库的数据进行相应的操作。 |
| **回收站模块** | | 项目、文件、翻译记忆表、术语表在删除后，会保存在回收站中。回收站中的内容在各部分前端页面不会显示。回收站中可恢复被删除内容，使其重新显示在原页面。也可彻底删除，彻底删除时，表对应的具体术语、翻译记忆、待翻译词条也会随之删除。 |  |

# 三. 数据需求

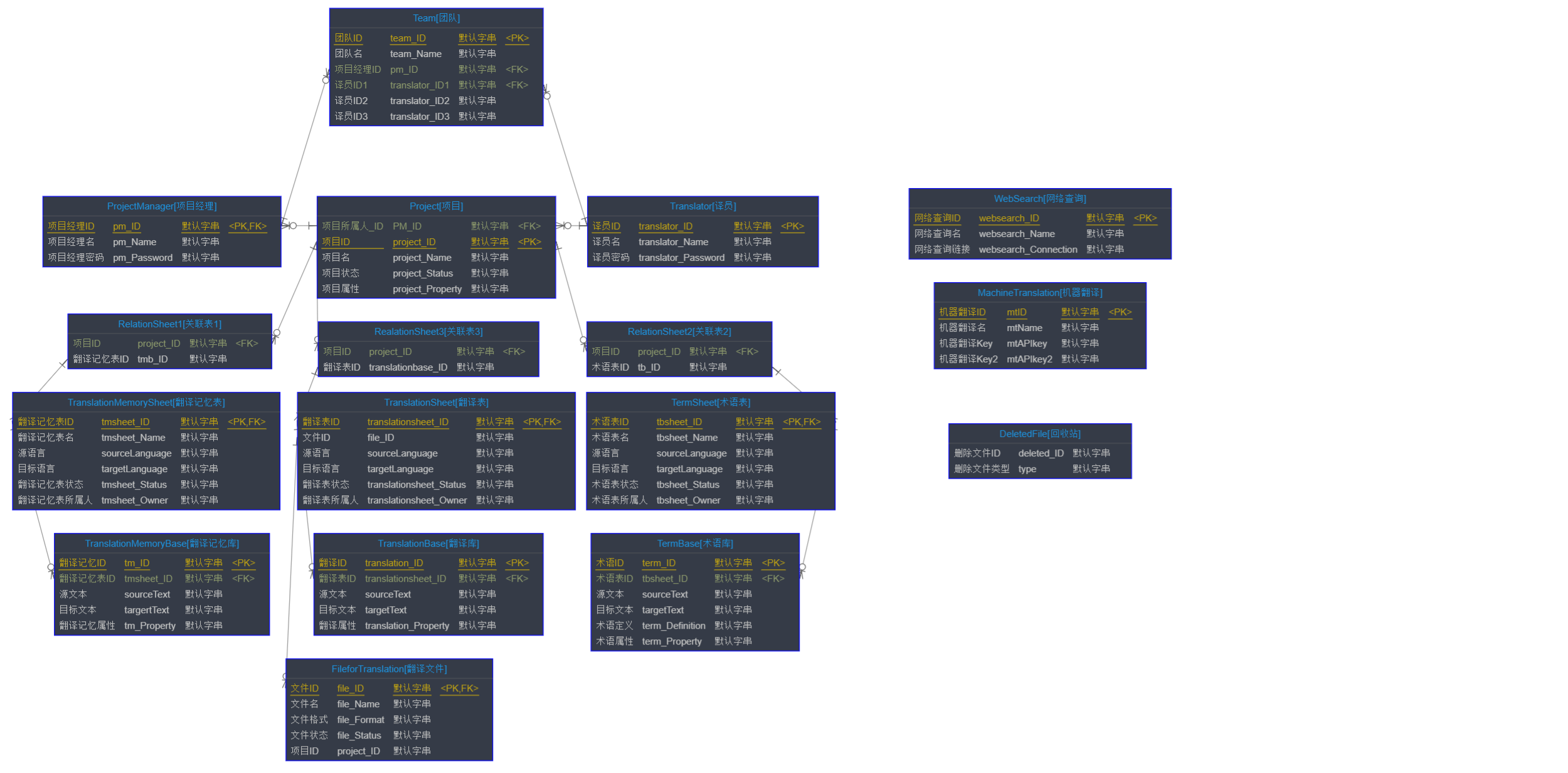


图 1 数据库-逻辑关系图

# 四. 用户界面

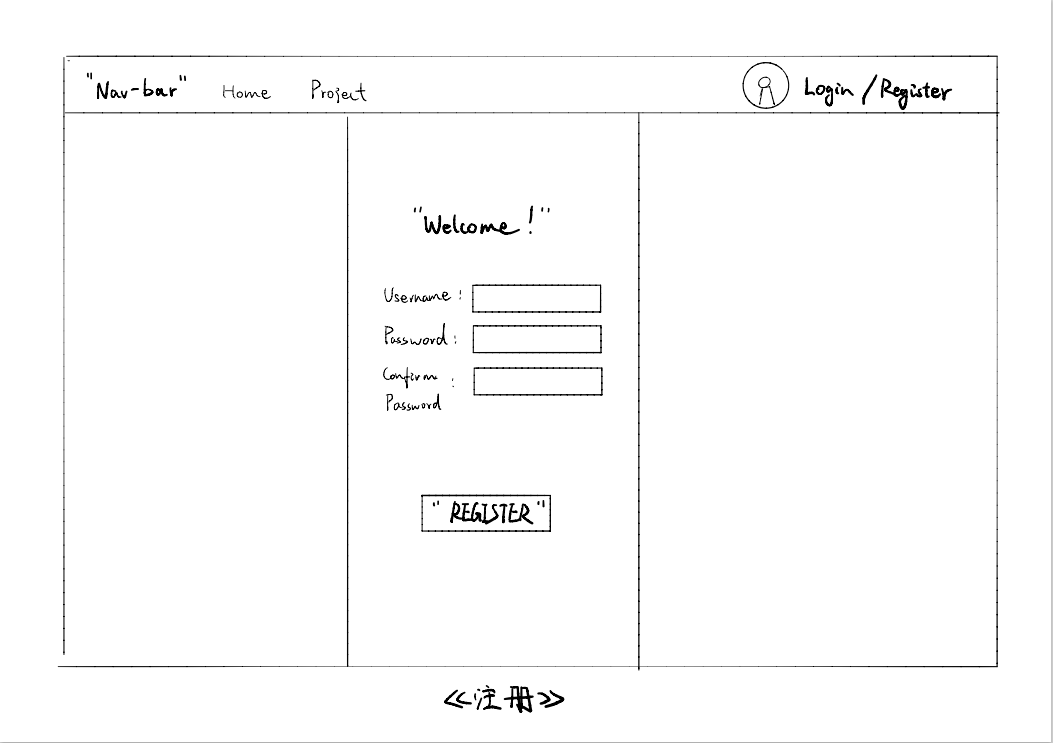
****

图 2 前端-注册

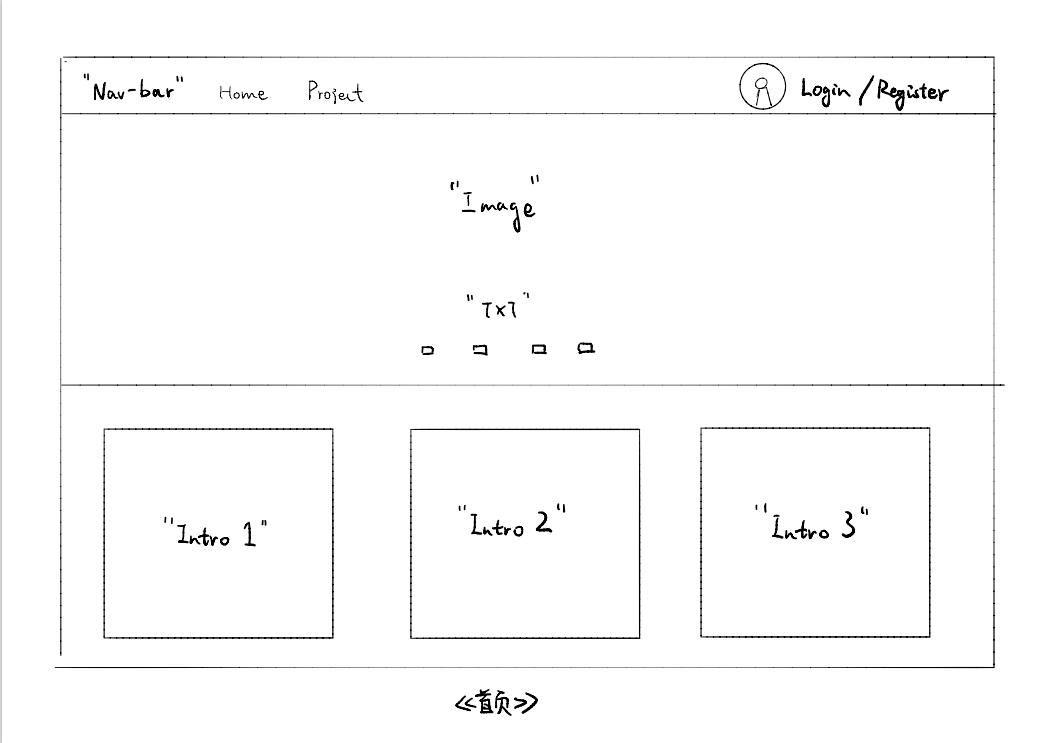
****

图 3 前端-首页

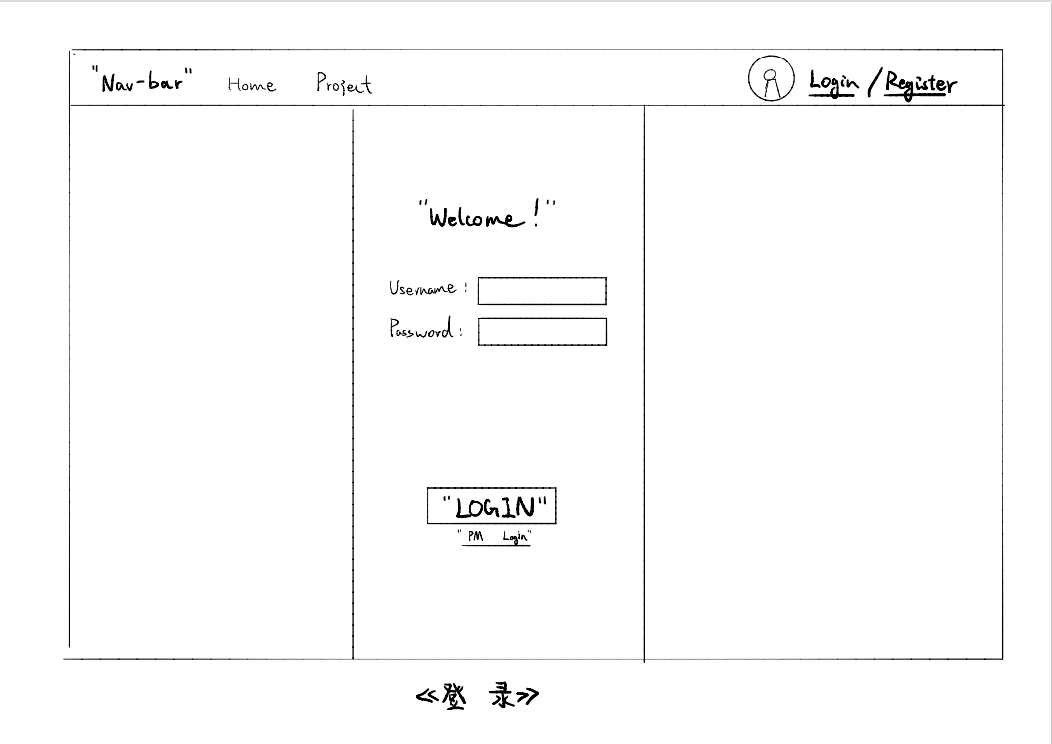
****

图 4 前端-用户登录

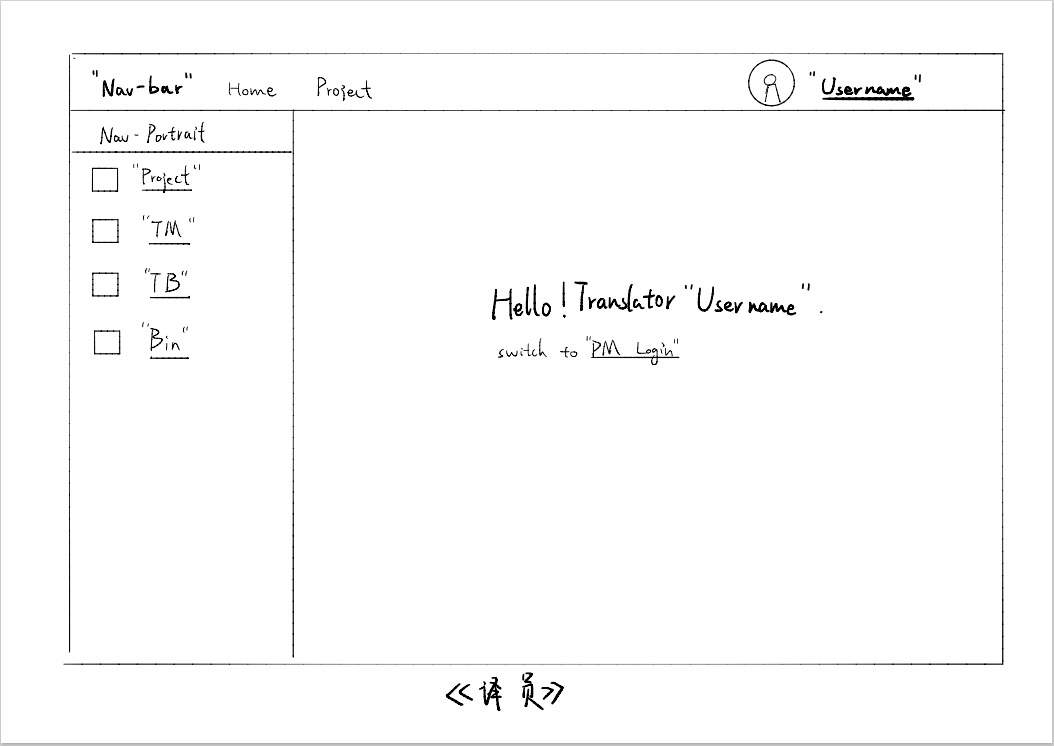
****

图 5 前端-译员界面

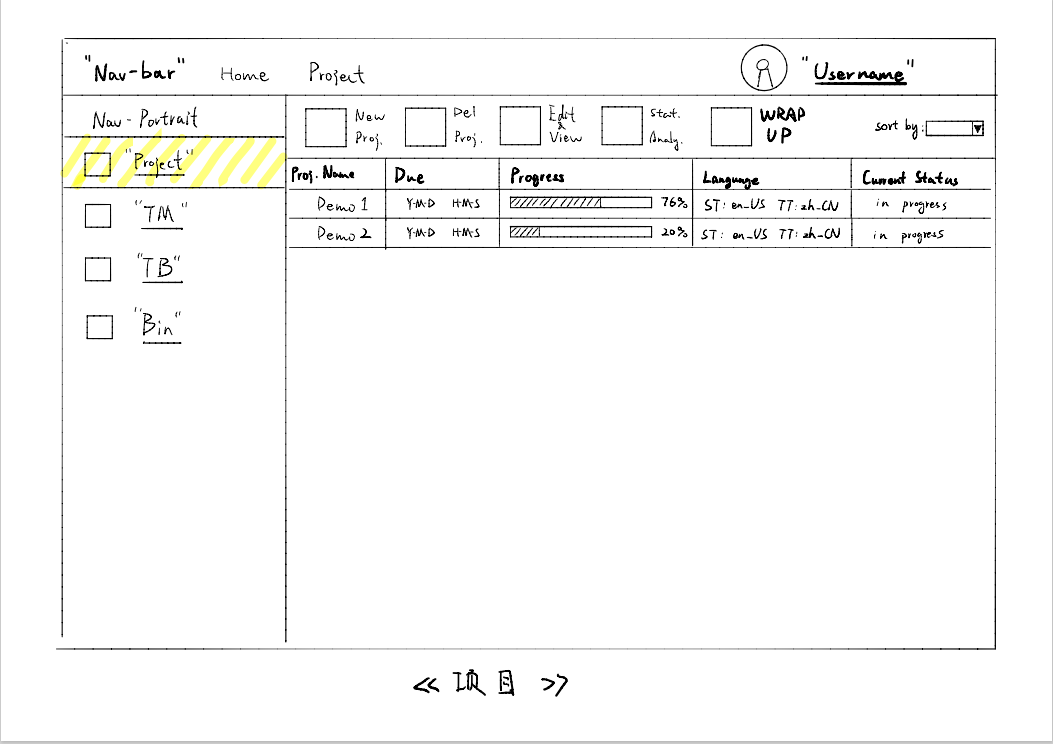
****

图 6 前端-项目界面

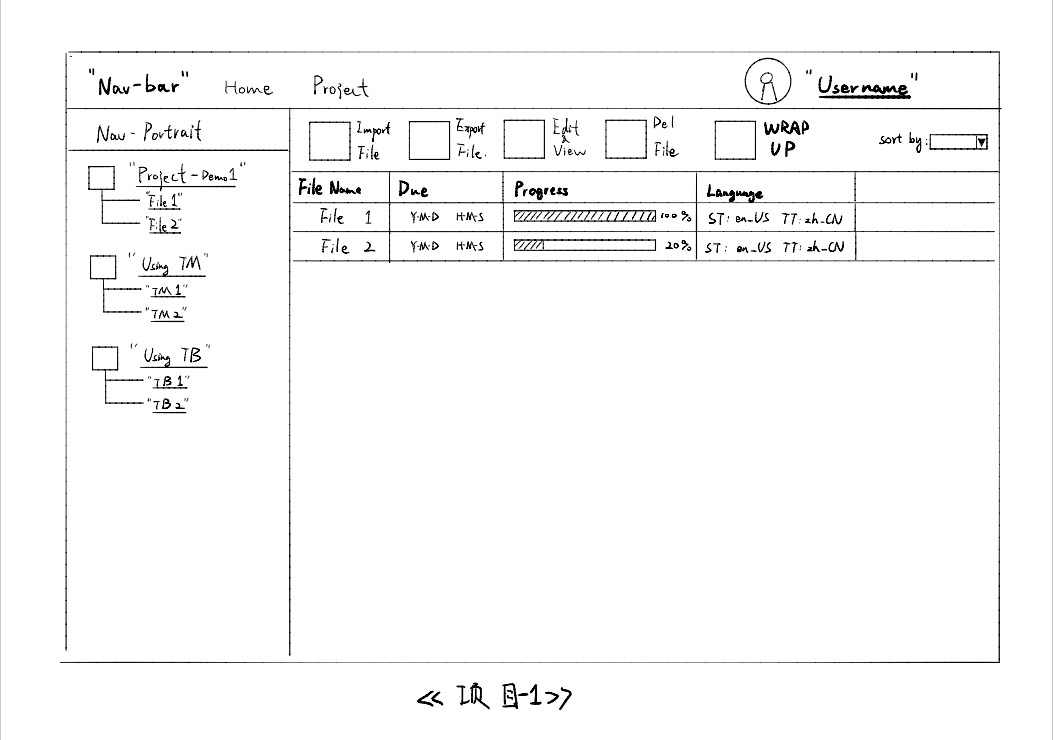
****

图 7 前端-项目内界面

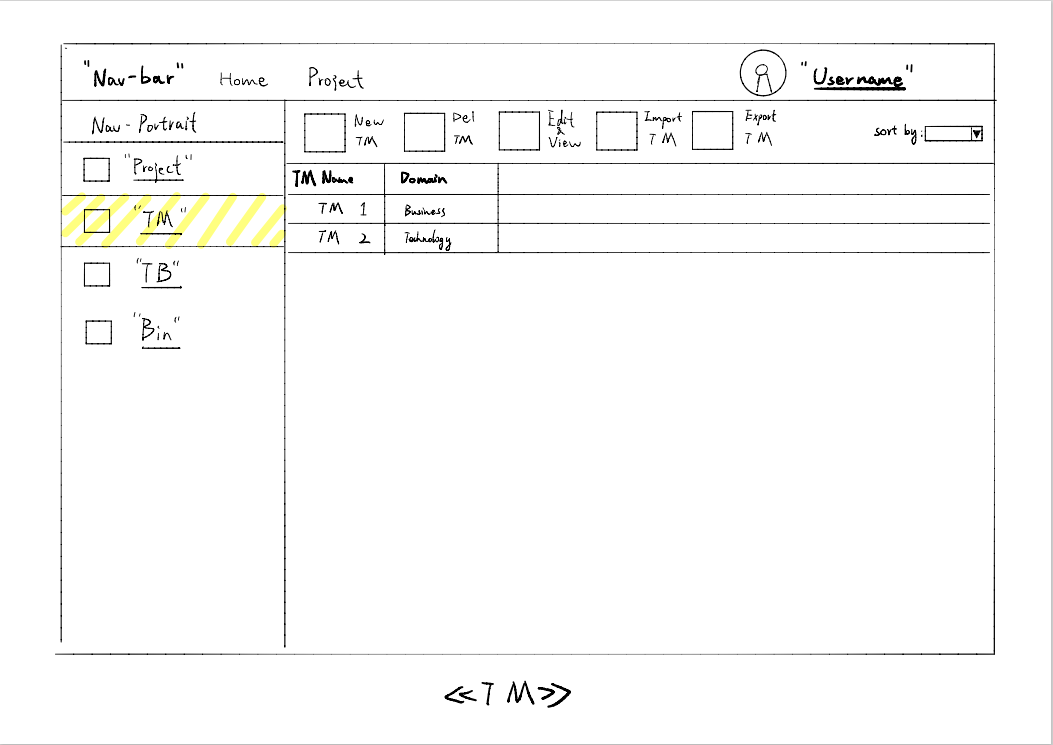
****

图 8 前端-翻译记忆库界面

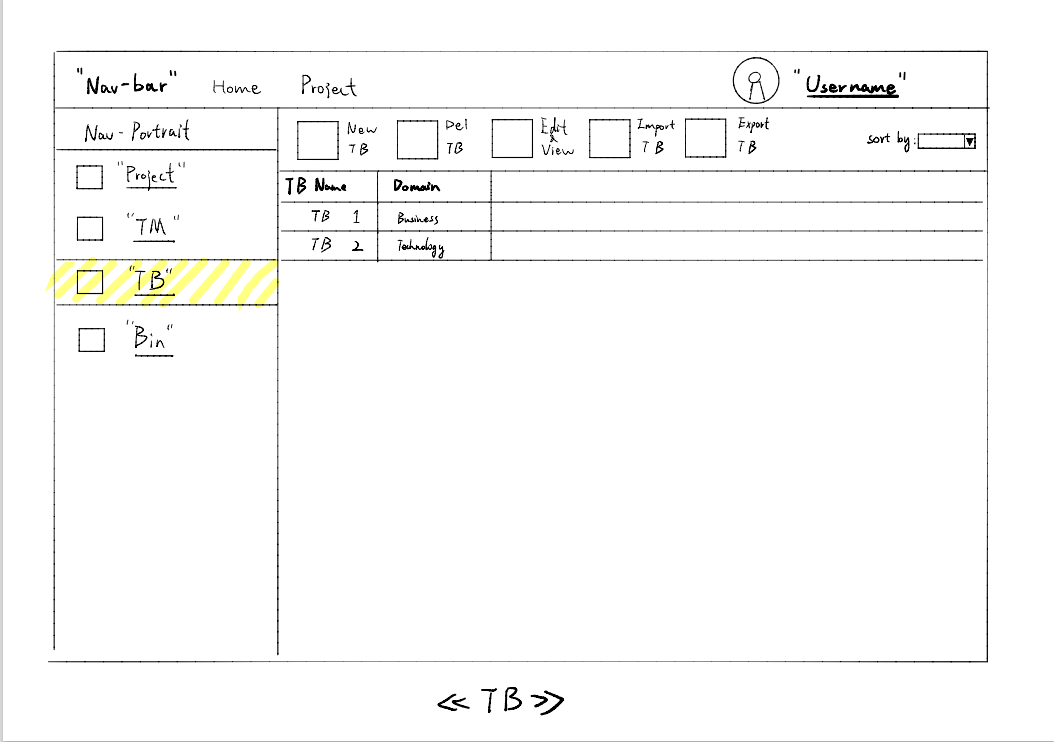
****

图 9 前端-术语库界面

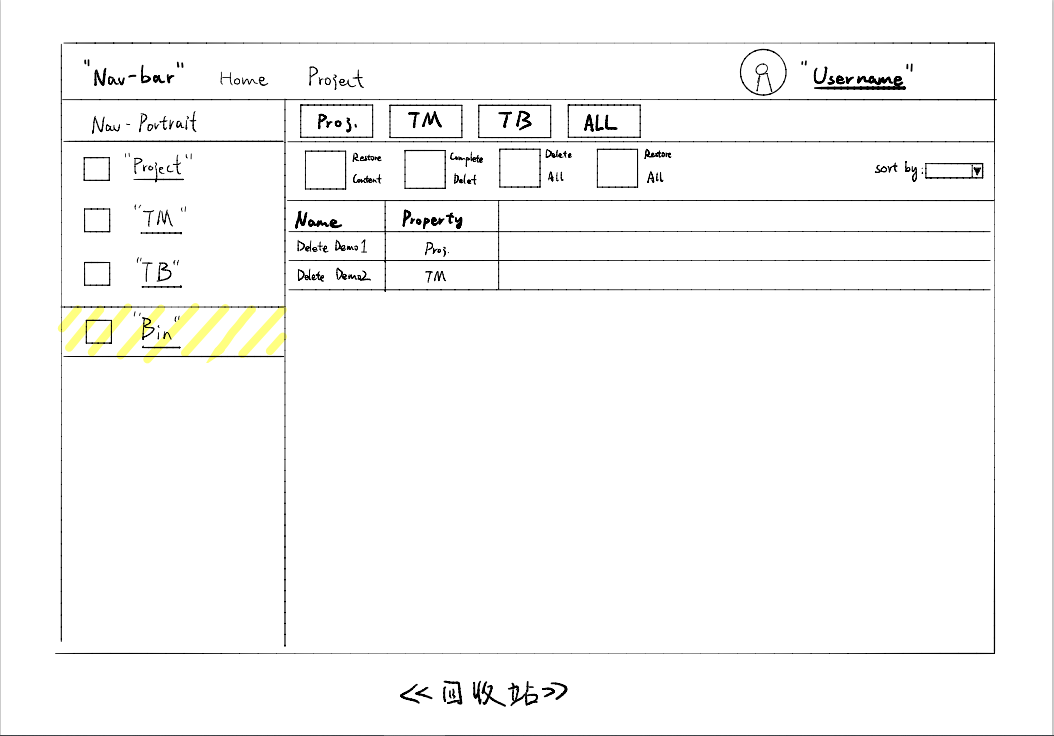
****

图 10 前端-回收站界面

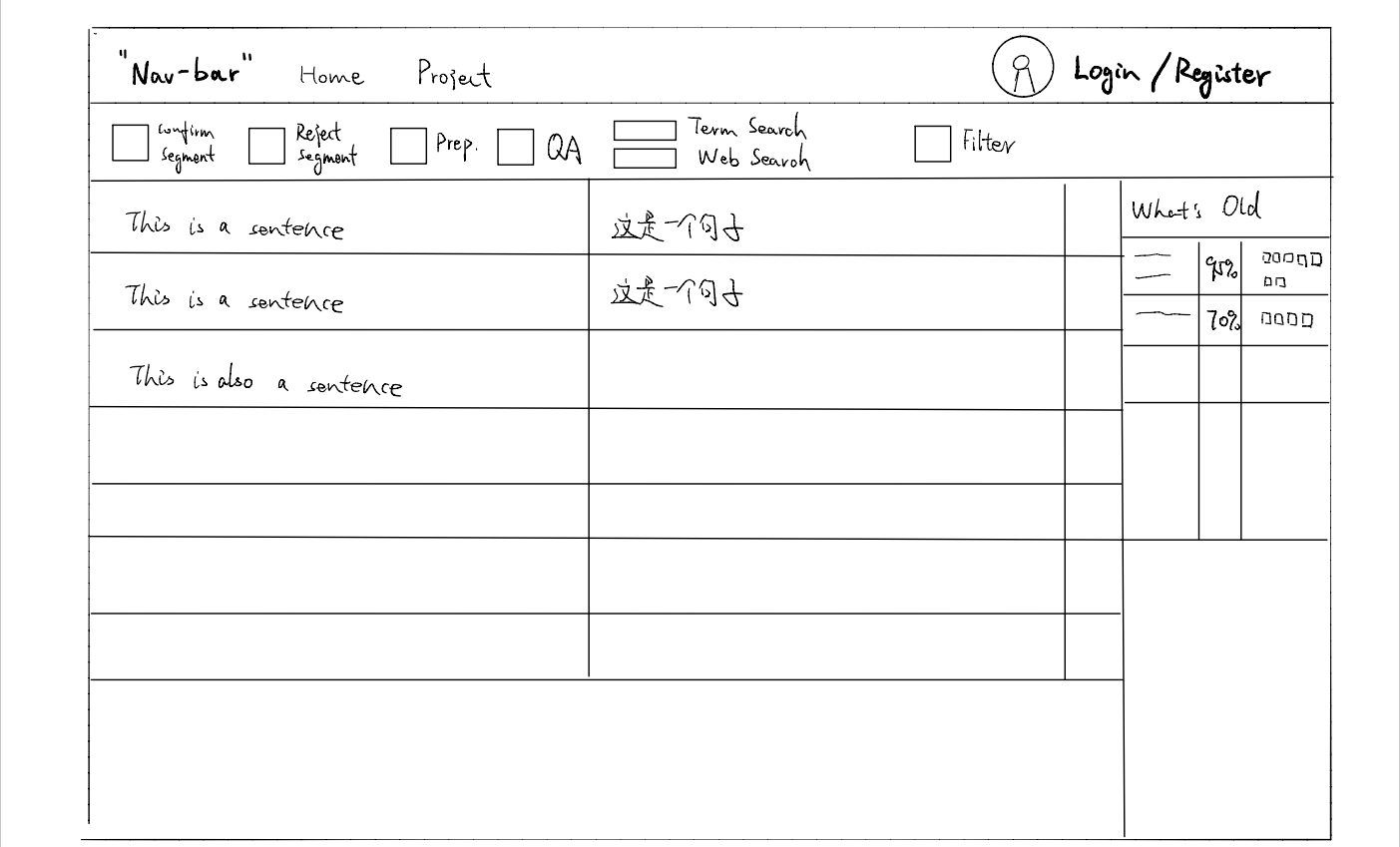
****

图 11 前端-文档翻译界面

# 五. 环境

**开发环境**

**数据库**

XAMPP，Mysql

**编程语言**

PHP，Python3，

**前端**

HTML5+css+javascript

Bootstrap前端框架

**运行环境**

Windows10系统

Google Chrome浏览器

Microsoft Edge 浏览器

# 六．功能框架图

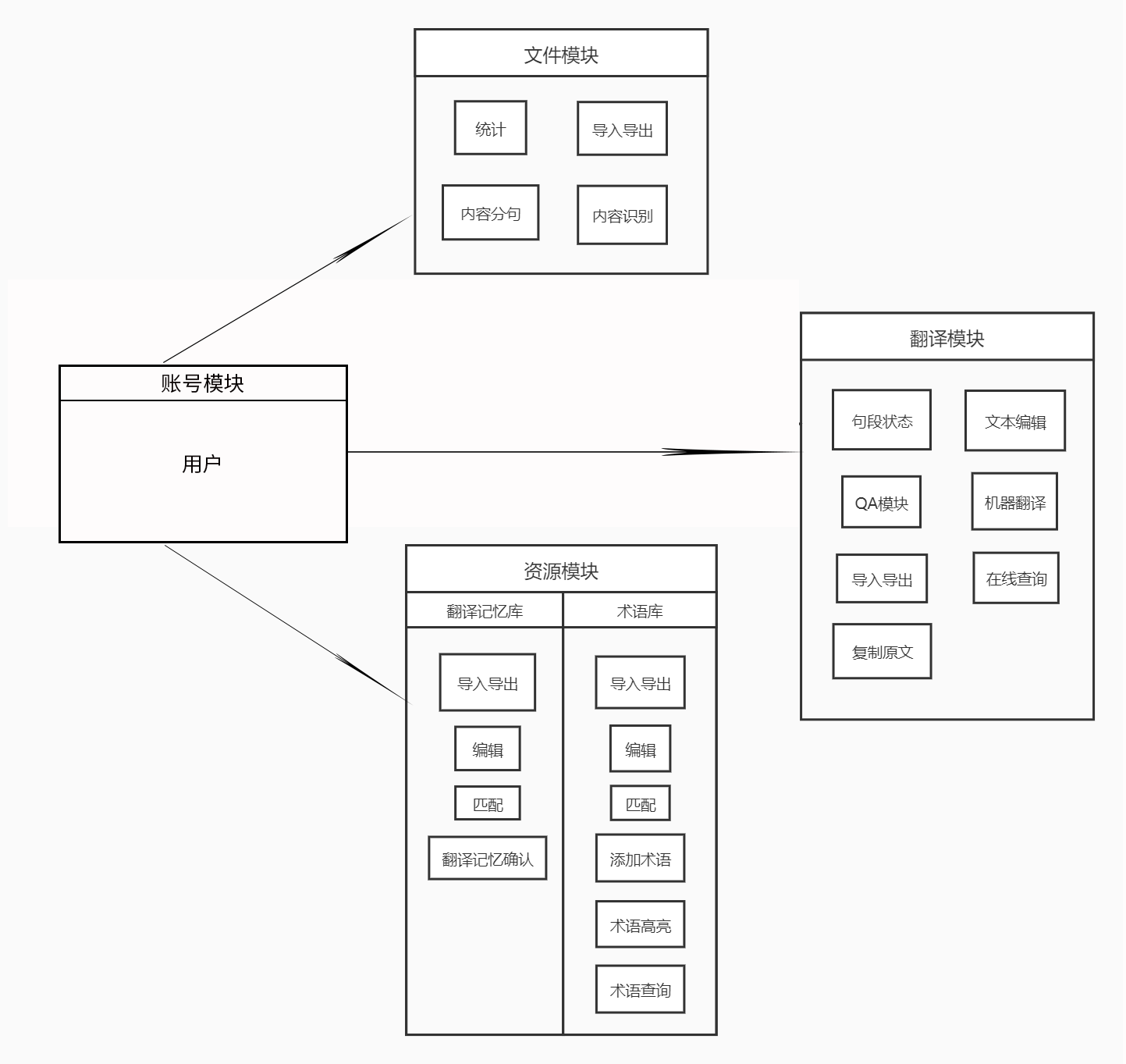
****

图 13 功能框架图