聆答 使用手册

实时字幕 · 面试 AI 助手

BigoLab 出品

安装说明 + 使用说明

2026-10-17

第一部分 安装与上手

> BigoLab 研发的面试实时字幕与 AI 助手

一个本地运行的面试助手：把电脑里听到的声音实时转成字幕，并让 AI 替你想好怎么回答。

它做什么：面试官通过腾讯会议、Zoom 这类软件说话 → 电脑把这段声音转成文字 → AI 结合你事先准备的资料，给出一段可以直接照着说的回答，显示在屏幕上的悬浮窗里。

它不做什么：不录音、不上传录音、不碰摄像头。只有文字会发给 AI 服务商。

---

1. 这台电脑能跑吗

项目	要求
系统	Windows 10 1809 以上 / Windows 11，64 位
硬盘	空出 10 GB
显卡	有 NVIDIA 显卡最好（识别快很多）；没有也能用，就是慢一些
网络	装的时候要下载约 8 GB；用的时候只有文字往返，很省

另外需要一个 **DeepSeek 账号**（AI 回答由它生成，按量付费，面试一场大约几毛钱）。 到 platform.deepseek.com 注册，充值后在「API Keys」里新建一个 Key，复制备用。

---

2. 安装

1. 把拿到的压缩包解压到一个**路径不含中文和空格**的地方，例如 `D:\` 。
2. 进入这个文件夹，**右键** `setup.ps1` → **使用 PowerShell 运行。**

如果右键没有这一项，就按 `Win+X` 打开「终端」或「PowerShell」，输入：

```
powershell -ExecutionPolicy Bypass -File D:\  \setup.ps1
```

3. 等它跑完（第一次 10–20 分钟）。**装之前先确认有 12 GB 以上空余空间**（PyTorch 约 2.5 GB、语音模型约 3 GB，解压时还要一份临时副本）。

脚本会先检查空间：默认装在 `%LOCALAPPDATA%\LingDa` ，如果那个盘不够，会自动换到空间最多的盘；想自己指定就加参数，例如 `powershell -ExecutionPolicy Bypass -File .\setup.ps1 -RuntimeDir E:\LingDa` 。它会自动装好 Python、按你的显卡装对应版本、编译悬浮窗，最后在桌面放一个「聆答面试助手」的快捷方式。

中途如果提示没有 Python，它会自己装；装完如果说「请重新运行一次」，关掉窗口再运行一次即可。

> 装完就不用再操心升级了：有新版本时聆答自己会弹框问你，点「是」十几秒就好， 只下载几百 KB 的代码，不会再花 20 分钟重装一遍，语音模型也不会重下。

---

3. 第一次启动

双击桌面的「聆答 面试助手」。

第一次会下载约 3 GB 的语音识别模型，要等几分钟，期间屏幕上没有反应是正常的。好了之后，右下角托盘里会出现图标，屏幕上出现一个悬浮窗。

接着做三件事：

(1) 填入 API Key

右键托盘图标 → **设置** → 找到「DeepSeek AI 助手」那一栏 → 把 Key 粘进去 → **保存并应用**。

Key 会加密保存在你自己的电脑上，不会上传到别处。

(2) 告诉 AI 你是干什么的

还是在设置页，「附加提示词」那一栏，用一两句话写清你的方向和岗位。这一步很关键——不写的话，AI 只能泛泛地答。例如：

> 我是做 Java 后端开发的，面的是电商平台的后端岗位。技术问题请结合 Spring、MySQL、Redis、消息队列这些常用技术栈来讲，举例优先取高并发、订单、库存这类场景。没有把握的参数和版本号不要编，直接说需要再确认。

换成你自己的方向照着写就行。

(3) 放入你的资料

设置页 →「背景资料」→ 点「打开资料文件夹」。里面有两个 `.example` 模板文件：

- `.md.example` — 复制一份、改名成 `A.md`，按里面的小标题填写你做过的项目。每个项目一份文件。
- `.txt.example` — 复制一份、改名成 `.txt`。这里写语音识别容易听错的专业词，格式是 `:` 。

> 带 `.example` 的文件不会被读取，所以模板本身可以留着不动。

填好保存，几秒内自动生效，不用重启。

资料越全，回答越像你自己说的。 建议至少写：一份简历要点、每个项目一份文档、一份术语表。

---

4. 两种模式

启动时会问「这次要做什么」：

- **面试模式**：听对话、出字幕、替你组织回答。
- **笔试模式**：不采集声音，读屏幕上的试题，逐道给出答案和分析。**电脑上不显示任何东西**，答案只推送到手机；结束后可把记录导出成一个文件。

选好立即开始；之后随时可以在托盘菜单里切换。勾上「记住选择」就不再每次询问。

5. 面试时怎么用

1. 提前 10 分钟启动，确认悬浮窗里能看到字幕。
2. 面试开始后，面试官说完一句，大约 1–2 秒后回答就会出现。
3. 悬浮窗可以拖动、缩放；按 `Ctrl+Alt+H` 一键隐藏 / 显示。

回答区怎么看：

- 每轮问答之间有分隔线，问题前面有序号。
- 最新一轮最亮，左边有蓝色竖线；旧的会变暗。
- 新回答出现时会自动顶到最上面，从第一行开始读。
- **第一句是加粗的结论**，来不及看完就只看这一句。

---

6. 常见问题

现象	处理
没有字幕	确认电脑正在出声，且会议软件用的是系统默认播放设备
手机上面试，电脑听不到	设置页「声音来源」改成「麦克风」，手机开外放放在电脑旁；详见使用说明第 5 节
有字幕但 AI 不回答	检查 API Key 填了没、账户还有没有余额
回答太泛、不像自己说的	第 3 步的「附加提示词」和资料没写够，补上
专业名词老是听错	写进 `.txt` ，一行一个
回答太长 / 太短	设置页的「附加提示词」里加一句，例如「回答控制在 150 字以内」
悬浮窗挡住东西	拖到副屏，或按 `Ctrl+Alt+H` 临时隐藏
双击快捷方式一闪而过	多半是字幕窗被 `Ctrl+Alt+H` 藏起来了，再双击一次就会出来；真失败会弹框说明原因
启动失败	看项目目录下 `.runtime\service.err.log` 的最后几行

---

7. 关于隐私

- 语音识别在**你自己的电脑上**完成，声音不出本机。
- 发给 DeepSeek 的只有：**识别出来的文字** + **你放进资料文件夹的内容**。所以不想外发的东西（身份证号、公司内部数据等）不要放进资料文件夹。
- 设置、API Key、资料都存在 `C:\Users\ \AppData\Local\WasapiParaformerOverlay\` ，删掉这个文件夹就清空了。

---

8. 还有一些进阶功能

默认关着，需要时在设置页打开，用法见 使用说明.md：

- **屏幕题目自动解答**：屏幕上出现编程题时自动给出思路和代码。

- **我的麦克风**：面试官的问题识别错了，你自己补一句，AI 按你说的重新回答。
- **手机远程查看**：把字幕和回答同步到手机上看（需要自己有一台服务器，一般用不上）。

第二部分 功能详解

BigoLab 研发

实时识别电脑播放的声音（对方说的话），显示字幕，并由 AI 给出回答；可同步到手机查看。 只采集电脑播放的声音，不采集麦克风。

识别和 AI 回答都在后台服务里运行，字幕窗和手机页面只是显示端：隐藏字幕窗不影响提问、回答和手机同步。

1. 两种模式

启动时会先问一句「这次要做什么」：

- **面试模式** —— 听电脑或麦克风里的对话，实时字幕，替你组织回答。
- **笔试模式** —— 不采集任何声音，反复读屏幕，把看到的试题逐道作答并记录下来。

选好立即开始。勾选「记住选择，下次不再询问」后就不再弹这个窗口，改用托盘菜单的「切换到...模式」随时切换，或在设置页「工作模式」里改。

两种模式互斥：切到笔试模式时，录音和识别引擎会完全停下（不占显卡、不读麦克风）；切回面试模式再重新启动。

笔试模式怎么用

1. 切到笔试模式，打开笔试页面。
2. 每 2.5 秒看一次屏幕，画面有变化才读。读到像试题的内容，先把整页拆成一道道题（顺手改正识别错误），再逐题作答，**最多三道题同时进行。**
    - 支持的题型：单选、多选、判断、填空、简答、**算法编程题。**
    - 编程题如果页面上给了函数签名或 `class Solution` 模板，会按它的语言和签名作答；没给模板时用设置页「屏幕题目自动解答」里的代码语言（默认 C++）。
    - 什么样的屏幕才会触发：出现"单选题/判断题"这类字样，或"编号 + ABCD 选项"，或"编号 + 分值"，或算法题的"输入/输出/示例/提示"组合。普通网页、聊天、代码编辑器不会触发。
3. **电脑上什么都不显示**：字幕窗自动隐藏，不弹任何提醒。题目和答案只推送到手机远程控制台（需要先配好第 8 节的手机查看）。
4. **只答完整的题**：一屏里有多道题时，被截断的那道会跳过，等你滚动到它完整露出来再答；一屏里只有一道题时，不管完不完整都会作答。
5. **不做重复判断**：同一道题滚回去再看还会重新作答一次。这是有意的——宁可多答一遍，也不要因为"看着像答过"而漏掉。只有画面和文字都没变的同一屏不会反复提问。
6. 设置页「工作模式」卡片里可以：换作答模型（默认 `deepseek-v4-pro` ，更准）、写附加要求（例如「用 Java 作答」）、**导出记录**、**清空记录。**
7. 导出会在桌面生成一个 `- .md` ，含每道题的题目、答案和分析。导出不会清空，清空前请先导出。

手机上看什么：每道题一张卡片，题号、题型、标题在上，【答案】和【分析】在下，原题默认折起来（电脑屏幕上本来就有）。中途打开手机也能看到前面已经答完的题。

答案由 AI 生成，不保证全对。遇到拿不准的题，AI 会在【分析】里说明不确定的原因。

2. 启动与退出

- **启动**：双击桌面的 **聆答** 快捷方式，等约 30 秒，看到"聆答 已就绪"后字幕窗出现在屏幕底部。
- **退出**：右键系统托盘图标 → **退出 聆答**，会同时停止后台识别和 AI。
- 字幕窗上的"×"只是隐藏窗口，后台照常工作。
- 设置页：http://127.0.0.1:8765/settings（改完点"保存并应用"）

自动更新

启动完成后，聆答会悄悄问一次服务器有没有新版本（没网就跳过，不耽误启动）。 有新版本时弹一个框，写明这次改了什么，装不装由你定；选"否"之后这个版本不会再问第二遍。

更新只下载代码，通常几百 KB、十几秒。Python 依赖和语音模型装在 `%LOCALAPPDATA%\LingDa` 下（安装时若该盘空间不足会自动换盘，实际位置记在 `%LOCALAPPDATA%\LingDa\home.txt` ；模型跟着一起走，老机器上已经下好的模型留在原处不动）， 不在程序目录里，**换版本不会重新下载它们**； 个人设置、API Key、背景资料在 `%LOCALAPPDATA%\WasapiParaformerOverlay` ，更新同样不碰。 旧版本会备份到 `%LOCALAPPDATA%\LingDa\backup` ，万一新版有问题可以换回去。

也可以自己查：

```
powershell -ExecutionPolicy Bypass -File .\update.ps1 -Check
```

3. 字幕窗操作

操作	方法
移动 / 缩放	未锁定时直接拖动窗口或边角
锁定位置	鼠标移到字幕上，点右上角中间的锁图标
清空对话	右上角左边的图标（保留资料和设置，手机端同步清空）
隐藏 / 显示字幕窗	`Ctrl+Alt+H` ，或点右上角的"×"隐藏
进入编辑模式	`Ctrl+Alt+O`

字幕不见了，先按一次 `Ctrl+Alt+H` 。隐藏期间识别和回答照常进行，重新显示时内容就是最新的；托盘图标的提示文字会标明"窗口已隐藏"。

想停止 AI 回答（不再向 DeepSeek 发送内容），在设置页取消"启用 AI 助手"；想全部停止，用托盘菜单退出。

4. AI 回答

在设置页的"DeepSeek AI 助手"里：

- 填入 API Key，勾选"启用 AI 助手"，点"测试 DeepSeek 连接"。

- **处理模式**用默认的"面试问答"：AI 知道这是一场面试、你是候选人，只在面试官发问或提要求时替你直接作答；寒暄、点评、铺垫和你自己的发言不会触发回答。
- **附加提示词**可以写岗位和偏好，例如"我面的是 C++ 后端岗位，回答控制在 100 字以内"。
- **final 后等待**：一句话定稿后再等多久才提问，当前为 0（定稿即提问）；想等对方多说几句再答可调大。

对方说完约 1.5 秒后开始出字（判定说完约 1 秒，DeepSeek 出首字 0.3–0.9 秒）。

技术题会直接给出具体内容；经历类问题在没有背景资料时只给回答框架，用【】标出需要换成你自己情况的地方——所以资料文件夹填得越全，这类回答越能直接用。

对方停顿后很快又接着说时，还没输出完的回答会被撤回，合并成完整问题重新回答；已经输出完的回答保留，后半句作为追问接着回答。

只有存在显示端（字幕窗在运行、本机状态页开着，或手机正在查看）时才会向 AI 提问。

回答区怎么看

- 每轮问答之间有分隔线，左侧问题前有序号，答案和问题一一对应。
- 最新一轮最亮，左边有一条蓝色竖线；更早的几轮会变暗，便于一眼找到当前该看的那条。
- 新回答出现时，视图会把它的第一行顶到最上面，从头开始读；回答比窗口还长时才跟着往下滚。
- 回答的第一句（结论）是加粗的，赶时间时只看这一句即可。
- 往上翻看旧回答时，底部会出现"↓ 最新回答"按钮，点一下回到最新。

5. 背景资料

AI 回答项目和经历类问题的依据。程序每次启动时自动读取并建立索引，不需要手动导入；文件改动后几秒内自动重建，不用重启。

资料从哪里读（设置页"背景资料"里查看和修改）

- 默认资料文件夹：点"打开默认资料文件夹"，把 `.md` / `.txt` 文件放进去即可。
- 其他位置：在"其他资料位置"填文件夹或单个文件的完整路径，点"添加资料位置"。程序直接读原文件，不复制，所以你在原处修改就会生效。
- 只读取文件夹第一层的 `.md` / `.txt`（编码 UTF-8），子文件夹不读；"其中不读取"可以填文件名通配符排除个别文件。

资料可以放在默认资料文件夹里，也可以用「添加文件夹」指向你平时写资料的目录（原地读取，不复制）。

资料怎么用

- 合计不超过 15 万字：全部资料每次提问都带上。
- 超过 15 万字（现在是这种情况）：自动切换为检索。
  - 8000 字以内的短文件（简历、简历口径核对）每次都带上。
  - 长文档只把"摘要"每次带上：开头说明、章节目录，以及标题含"一分钟"的那一节。
  - 长文档的正文切成约一千字的片段，每个问题只附上最相关的十来段。
- 追问时默认沿用上一个问题所在的项目，不必每句都说项目名。

- 资料里标〔待核实〕的内容，AI 不会当成确定的事实说出来。

识别听错时

- 术语被听成近音的字词时，会按读音对回资料里的术语：英文术语写成了汉字或别的单词（Sinkhorn → "think好恩""辛克霍恩""sink iron"），中文术语写成了同音字（算子 → "蒜子"，访存 → "仿存"）。找回的术语用于检索，并作为"可能是××"提示给 AI。
- 只对资料里反复出现的术语有效：英文术语 6 个字母以上、出现 3 处以上；中文术语 2–4 个字、出现 2 处以上。
- **术语表**：资料文件夹里名为 `.txt`（或 `terms.txt`）的文件按术语表读取，每行 `:` `1, 2`，例如 `Kubernetes: ,` 。按读音匹配，列一两种常见写法即可（"积木ruai思""基母瑞思"也会对上）；保存后几秒内生效。文件就放在资料文件夹里（设置页「背景资料」可一键打开）。它只影响 AI 的理解和回答，字幕上的原文不会被改写。
- 问题没点名是哪个项目、上文也看不出来时，回答开头会带"（按 XX 项目）"，表示项目是 AI 猜的。

写资料的建议

- 每个项目一份文档，用 `#` `##` `###` 标题分节——检索按小节切片，标题写得具体，命中就准。
- 准备好的问答写成 **Q1 问题?** 开头的段落，每个问答会单独成片，并在检索时优先。
- 给每个项目写一节"一分钟项目介绍"，它会常驻，用来回答"介绍一下这个项目"这类笼统问题。

资料会随请求发送给 DeepSeek（简历里的手机号和邮箱也在其中），涉及保密或不想发送的内容请自行删改或排除。

6. 声音来源（手机面试 / 线下面试）

设置页「声音来源」两个选项：

- **电脑播放的声音**（默认）：腾讯会议、Zoom 这类在电脑上进行的面试。
- **麦克风**：手机面试、线下面试——凡是对方的声音不经过电脑的场合。选它之后再选一个输入麦克风。

手机面试怎么摆（按效果从好到差）：

1. **音频线**：手机耳机孔 → 电脑耳机孔（双公头 3.5mm 线），电脑把它当麦克风读。杂音最少，但手机和电脑都要有 3.5mm 孔。
2. **手机外放 + 电脑麦克风**：手机开免提放在电脑旁 30 厘米内，声音来源选「麦克风」。最省事，不用任何配件。
3. **安卓投屏**：用 scrcpy 把手机音频转到电脑播放，这样仍然走「电脑播放的声音」模式。

用外放时你自己说的话也会被一起识别进去。这不影响使用——AI 本来就按「转写里混有两边说话、没有说话人标记」来理解。

> 改声音来源会重新加载识别引擎，约 20 秒。改完先说一句话，确认字幕出得来再开始面试。

7. 识别语言

设置页"识别语言"：**中英混合**（当前）、中文、英文。切换后识别引擎会重新加载。

8. 手机远程查看

> 这一节需要你自己有一台可以公网访问的服务器（运行 RDViewer 中继）。没有的话跳过即可，不影响其它功能。

首次配对

1. 手机浏览器打开你的中继服务器地址（形如 `https://<  IP>:8443` ），提示证书不受信任时选择继续，登录后点"添加设备"，得到 8 位配对码（5 分钟有效）。
2. 电脑设置页"远程控制台"：填同一个中继地址和配对码，点"配对并连接"。

日常使用：手机登录后点设备"… · 聆答"，即可看到实时字幕和回答。右上角 `A-` / `A+` 调字号。

- 字幕窗隐藏时，手机照常收到字幕和回答。
- 手机页面切到后台或锁屏会暂停同步，回到前台自动补齐。
- 有设备在看时，电脑托盘会提示"远程查看中"。没人看时不上传任何内容；服务器只转发，不保存。

9. 屏幕题目自动解答（面试模式下）

设置页"屏幕题目自动解答"勾选启用后，屏幕上出现编程题 / 算法题时会自动求解，思路和完整代码显示在手机远程控制台的浮窗里。

- **只在有设备正在看远程控制台时才读屏幕**：每 2 秒截取主显示器，画面有变化才做文字识别（Windows 自带，在本机完成）。
- 识别出像题目的内容、并且连续两次看到同一屏，才把屏幕文字发给 DeepSeek；从题目出现到开始出字约 4–6 秒。
- **题目不变不会重复求解**：滚动页面、在编辑器里写代码、切到别的窗口再切回来都不会重新请求。换了一道题会自动解新题，浮窗只保留最新一题。
- 代码语言默认 C++，可在设置页改；页面上给了函数签名（如 `class Solution` ）时按它写，否则写带输入输出的完整程序。
- 浮窗可拖动标题栏移动，"—"收起成小标签，"复制代码"复制代码块。
- 想让当前题目重新解一遍：设置页点"重新解答当前题目"。
- 截图和识别出的文字不保存；屏幕文字（可能含屏幕上的其他内容）会发送给 DeepSeek。

10. 用自己的话纠正 AI（麦克风输入）

设置页"我的麦克风（补充输入）"勾选启用，并在下拉框里选你实际说话用的麦克风。

- **用途**：面试官的问题识别错了、AI 答偏了，你把问题复述或补充一遍，AI 按你说的重新回答，正在显示的错误回答会被替换。
- **一定会触发重新回答的说法**：`……` 、`……` 、`……` 、`……` 、`……` 、`/ /` `……` 、`……` ，或句尾带"对吧 / 对吗"。
- **你正常答题时 AI 不作答**：面试官说完后你的第一句话会交给 AI 判断（一般不回应），之后的话只作为背景保留、不再发请求，也不会打断屏幕上的回答，直到面试官再次开口。
- **和在线会议不冲突**：Windows 允许多个程序同时读同一个麦克风，不影响会议里对方听到的声音；在会议软件里静音不影响这里的识别（系统级或硬件静音会）。
- 对方说话时麦克风输入会被忽略，所以用外放也不会把对方的声音当成你说的。

- 你说的话不显示在字幕窗，会以"【我】"开头显示在手机远程控制台里。
- 仅"中英混合"识别模式可用；开关或换麦克风后识别引擎会重新加载（约 20 秒）。

11. 常见问题

现象	处理
没有字幕	确认电脑在播放声音，且用的是系统默认播放设备；按 `Ctrl+Alt+H` 试试
字幕有、AI 不回答	检查 API Key、余额，以及"启用 AI 助手"是否勾选
手机上有字幕但没有回答	同上；日志里 `assistant:` 开头的行会写明原因
专有名词识别错	把正确写法写进背景资料，关键术语多出现几次；听成近音字时会按读音自动对回去
回答开头是"（按 XX 项目）"	问题没点名，项目是 AI 猜的；猜错时等面试官下一句带上项目名，或请对方重复
AI 对无关声音作答	关掉电脑上其他正在播放声音的程序
手机显示"电脑端未在线"	电脑上 聆答 未启动，或设置页"远程控制台"显示未连接
手机提示会话过期	重新登录即可，无需重新配对
双击快捷方式后一闪而过	字幕窗若是被藏起来了，再双击一次就会让它重新出现；真的启动失败会弹框说明原因
手机面试听不到对方	设置页「声音来源」改成「麦克风」，并把手机开外放放在电脑旁
笔试模式不出题	确认已切到笔试模式；题目要完整露在屏幕上，只露一半不会作答
笔试模式下没有字幕	正常，这个模式不采集声音
更新失败	旧版本备份在 `%LOCALAPPDATA%\LingDa\backup\<  >-<  >` ，把里面的文件拷回程序目录即可
启动失败	查看项目目录下 `.runtime\service.err.log`

排查用的日志：

- `.runtime\service.err.log` ：识别服务与 AI。每轮提问的耗时、首字延迟、缓存命中都在 `assistant:` 开头的行里。
- `overlay_cs\bin\overlay.runtime.log` ：字幕窗。

新版本的这两份日志都不再记录字幕和回答的正文。