技術說明文件・TECHNICAL NOTE

# 菜根機器人語音互動系統：使用說明與設計概要

*A User Guide to the Voice Interaction System of the Caigen Discord Bot*

菜根機器人開發團隊

世界計畫車隊管理機器人 ・ 版本 b2026-11-09l ・ 2026 年 11 月

**摘 要**　本文說明菜根機器人在 Discord 語音頻道中的互動功能。系統將成員的語音即時轉成文字，經過斷句、語音辨識與三層自動校正後，判斷為指令、跑榜查詢或閒聊，再以語音回覆。文中介紹三種互動模式、可查詢的跑榜數據（時速、周回、場均、樓上樓下分差與班表）、車隊術語的校正機制、音樂播放，以及提升辨識率的使用建議與疑難排解方式。

**關鍵詞：**語音辨識；Discord；車隊管理；跑榜助理；自動校正；世界計畫

## 1　緒論

跑榜時雙手多半在操作遊戲，打字查排名或報班並不方便。本系統讓成員直接對機器人說話：問數字就查即時資料回答，講指令就協助執行，其餘則以對話方式回應。所有功能都在既有的語音頻道內完成，不需要另外安裝程式。

## 2　快速開始

1. 先進入要使用的語音頻道。
2. 輸入 `/語音 加入`，機器人進入頻道並開始收音。
3. 直接說話。文字頻道出現「你的名字（語音）」的字幕，即表示已收到。

結束時輸入 `/語音 離開`；若正在播放音樂，機器人不會斷線。

## 3　互動模式

系統提供三種互動模式，差別在觸發方式與回應速度，比較如表 1。

**表 1**　三種互動模式比較

| 模式 | 開始 | 結束 | 特性 |
|---|---|---|---|
| 直接講（預設） | 直接說話 | — | 每句都回 |
| 對話模式 | 喊「菜根」 | 說「掰掰」 | 30 秒內免再喊 |
| 即時對話 | `/語音 即時對話` | 說「結束對話」 | 延遲 <1 秒，可插話 |

## 4　系統流程

一句話從說出到得到回覆，會經過圖 1 的六個步驟。其中斷句與校正兩步驟對辨識率影響最大。

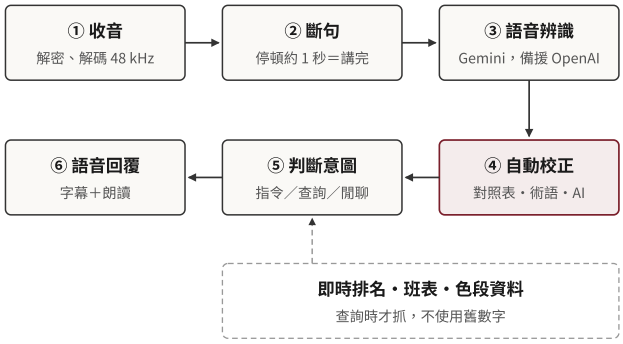



**圖 1**　語音互動的處理流程

即時對話模式則把聲音直接串流給 Gemini Live，不經過步驟 ② 至 ④。[1]

## 5　跑榜助理

### 5.1　可查詢的數據

凡是提到時速、場均、樓下、榜線、誰上車等字眼，系統都會先抓取當下的排名與班表，再依資料回答，不會自行推測數字。可查詢項目見表 2。

**表 2**　跑榜助理可回答的項目

| 項目 | 內容 |
|---|---|
| 名次與總分 | 跑者、以及說話者自己綁定的帳號 |
| 時速、周回 | 近 1 小時每小時分數與場數；另附 3 小時數據 |
| 場均 | 見式 (1)；另附色段監控的近 10 場平均 |

| 項目 | 內容 |
|---|---|
| 樓上、樓下 | 對方的時速、周回、場均與分差 |
| 追趕估算 | 照目前時速幾小時追上或被追上 |
| 榜線目標 | 下一個榜線（如 T30）還差幾分 |
| 預估終分 | 照目前時速到活動結束的總分 |
| 班表 | 這小時、下小時的跑者與 P2～P5、空位、待審 |
| 色段監控 | Auto／多人模式、閒置時間、異常警告 |

$$\text{場均} = \text{近 1 小時時速} \div \text{近 1 小時周回} \tag{1}$$

3 小時與 24 小時的場均刻意不計算。[2]

### 5.2 問法範例

- 「跑者時速多少？」「我的周回跟場均」
- 「樓下場均多少？差幾分？」
- 「照這速度幾小時會被追上？」
- 「下小時誰要上車？」「T100 現在幾分？」
- 「幫我總結一下戰況」：依表 2 順序完整說明

### 5.3 前置設定

跑者 ID 用 `/跑者id` 綁定；要問「我的」數字，說話者需以 `/成員 綁定id` 綁定自己的遊戲 ID；色段數據需先 `/色段 開始`。

## 6 語音指令

報班、取消、查班表、問房號等都可直接說，例如「幫我報班 20 到 24」、「取消 21 點的班」。查詢類會立即回答；會修改資料的操作會先跳出確認按鈕，按下後才生效。

## 7 語音自動校正

### 7.1 三層校正

辨識結果在送出判斷前，依序經過表 3 的三層校正。前兩層不需呼叫 AI，可立即完成；第三層若 4 秒內未回應，就採用前兩層的結果。

**表 3** 自動校正的三個層級

| 層級 | 範例 |
|---|---|
| 常見聽錯對照 | 報斑→報班、蔡根→菜根、常均→場均 |
| 字母＋數字術語 | 艾斯六→S6、屁二→P2、踢一百→T100 |
| AI 依名單校正 | 成員名、車隊術語、「二十點」→「20點」 |

為避免誤改，AI 若改動過多、改成回答問題或輸出多行說明，系統一律保留原句。有修改的字幕下方會標示「已自動校正・原句：…」，供成員核對。

### 7.2 自訂與自動學習

管理員可用 `/語音 校正詞`：

- **新增**：填「聽成」與「應該是」，適合固定被聽錯的詞。
- **新增術語**：只填「應該是」，加入車名、暱稱等車隊用語。
- **清掉自動學到的**：撤銷系統自行學到的對照。

AI 在三句不同的話中把同一組詞改對，系統會自動記成校正詞；色段暱稱與班表中的自訂跑者名也會自動納入術語。

## 8 音樂播放

**表 4** 音樂相關指令

| 指令 | 說明 |
|---|---|
| `/語音 播放` | 以網址或關鍵字點歌 |
| `/語音 跳過`、`/語音 暫停` | 下一首；暫停／繼續 |
| `/語音 音量` | 5～200%，建議不超過 100% |
| `/語音 混音` | 說話時音樂壓低而不暫停 |

## 9 使用建議

**提升辨識率的五個做法**

1. 說完停頓約 1 秒，系統才判定句子結束；句中換氣不會被切斷。
2. 一次一人說話；多人同時說，系統會依序逐句回覆。
3. Discord「語音與視訊」設定中，輸入模式選「語音活動」，靈敏度勿調過高。
4. 使用按鍵說話時，說完再放開按鍵。
5. 常被聽錯的詞，加成校正詞或術語（見 7.2）。

## 10 管理與疑難排解

**表 5** 常見問題與對應指令

| 狀況 | 處理方式 |
|---|---|
| 說話後沒有字幕 | `/語音 狀態`，查看「最近幾句」 |
| 聽不到機器人 | `/診斷 語音播報` |
| 音樂斷斷續續 | 發生當下執行 `/診斷 語音品質` |
| 調整回覆時機、字幕 | `/語音 辨識設定`（管理員） |
| 更換說話聲音 | `/語音 音色` |
| 自動念排名 | `/語音 播報`（每 10 分鐘） |

回報問題時，附上上述指令的畫面截圖最有助於判斷原因。

[1] 即時對話直接聽取聲音，因此不顯示自動校正標示；車隊術語表仍會提供給模型參考。

[2] 排名資料中 3 小時與 24 小時的「時速」為每小時平均，「場數」則為整段累計，兩者相除會差數倍，故僅以近 1 小時計算場均。