RecordStudio v1.0 產品白皮書

第一章　產品整體概述

本系統採用 **Client-Server（C/S）架構設計**，整體分為兩個核心組件：

- **S-Server（服務端）**：部署於 Mac mini M4 Pro，負責所有核心業務邏輯、錄製引擎、AI 推理、串流中轉、數據存儲。作為全系統的「調度大腦」，S-Server 管理所有 C-Client Agent 的註冊、心跳追蹤、指令下發與狀態彙總。
- **C-Client（客戶端）**：Windows 平台，運行 Recording Agent（Python 後台服務）+ Web 管理界面。作為分佈式「錄製工人」，C-Client 接收 S-Server 指令，執行各房間的視頻錄製作業（支援 NVR 模式與直連攝像頭模式），支援離線錄製（斷連後本地持續錄製，恢復後補報狀態）。

RecordStudio v1.0（以下簡稱「本系統」）是一款專為司法機構、仲裁中心、法務部門及大型企業會議室設計的商用級多媒體錄製管理平台。系統支援最多 10 個房間並行錄製，每房間最多 12 路攝像頭同步錄製（全系統最高 120 路並發），4 至 64 聲道 Dante 數字音頻採集（可配置），集即時串流、AI 轉寫翻譯、智能降噪、全文檢索於一體，實現從會議錄製到內容歸檔的全流程自動化管控。

本系統採用離線優先架構設計，所有核心功能（包括錄製、回放、語音轉寫、全文搜索、系統監控）均可在完全脫離互聯網的內網環境中獨立運行。AI 推理引擎部署於本地設備，確保會議音視頻數據全程不經過外部服務器，滿足司法錄取場景對數據主權與隱私安全的嚴格要求。系統原生支持普通話、粵語（廣東話）及繁體中文，三種中文變體均經過專項優化訓練。

RecordStudio v1.0 基於 Python 高性能異步後端框架構建，前端採用現代 Web 前端框架，通過模塊化 API 設計實現高內聚低耦合的系統架構。流媒體傳輸採用零轉碼串流引擎，實現 RTSP 至 WebRTC 的零轉碼串流，端到端延遲控制在 50-200 毫秒區間，確保會議現場的即時可視化監看體驗。

本系統提供單機版、專業版、企業版三層產品矩陣，分別針對單房間基礎錄製、多房間 AI 增強錄製、以及大型司法機構全功能合規錄製等差異化需求，提供精準匹配的軟件授權方案。

第二章　軟件核心定位與適用場景

2.1 核心產品定位

RecordStudio v1.0 定位為「專業級多媒體錄製與內容管理平台」，專注於為法律程序的錄製取證、會議內容的規範存檔、以及多語言環境下的信息精準傳遞提供技術支撐。系統以「安全可靠、智能高效、合規可審」為核心價值主張，助力各類機構實現錄製工作的標準化、數字化、智能化轉型。

在技術定位層面，本系統填補了傳統錄播設備與通用會議軟件之間的市場空白。傳統錄播設備功能單一、擴展性差；通用會議軟件無法滿足司法錄取的時間戳精準對齊、審計合規、多語言翻譯等專業需求。本系統則在兼顧易用性的同時，全面適配司法、法務、商務談判等高價值會議場景的專業要求。

2.2 行業痛點解決

傳統會議錄製工作長期面臨以下核心痛點：

多設備協同困難：會議現場通常部署多路攝像頭與專業音頻設備，各設備獨立運作，缺乏統一管控界面，操作人員需要在多個設備之間來回切換，不僅效率低下，更易因人為操作失誤導致錄製失敗或素材缺失。本系統採用 C/S 架構，C-Client 客戶端負責錄製執行，S-Server 服務端負責統一調度，提供最多 10 個房間、每房間最多 12 路攝像頭（全系統最高 120 路並發）、全系統 64 聲道 Dante 音頻的集中管控，一個界面完成所有設備的狀態監控與錄製操作。

錄製內容檢索不便：傳統方式錄製的會議素材以離散視頻文件存儲，查找特定發言段落需要逐幀播放定位，耗費大量時間。本系統集成全文搜索引擎，支持對 AI 轉寫結果進行秒級全文檢索，點擊搜索結果即可跳轉至對應時間點。

多語言環境溝通障礙：國際仲裁、涉外交涉等場景中，與會人員使用不同語言，傳統方式需要專人即時翻譯記錄。本系統內置離線翻譯引擎，支援中文、英文、葡文等 33 種語言的即時翻譯，轉寫結果與翻譯內容自動時間軸對齊。特別強調中文語境下的獨特需求——普通話與粵語的雙語切換、繁簡體差異、中文法律術語的精準識別，本系統均提供專項優化支持。

數據安全與合規風險：會議錄製內容往往涉及機密信息，傳統方案依賴外部雲服務進行 AI 處理，存在數據外泄風險。本系統所有 AI 模型本地部署，錄製數據全程在機構內網流轉，符合司法場景的數據主權要求。

2.3 適用場景

本系統適用於以下核心場景：

司法庭審錄製：各級法院、仲裁機構的庭審過程全程錄製，支持多機位同步取證、庭審錄音轉寫、庭審記錄歸檔查詢。完整的審計日誌功能滿足司法程序對錄製過程可追溯性的法定要求。

國際會議與談判：涉及多語言使用者的國際會議、商務談判、外交接洽等場景，提供即時翻譯與多語言字幕疊加功能，確保各方信息暢通。

大型研討會與論壇：容納數十人至數百人的大型會議場所，多房間並行錄製、統一後台管理，滿足大型活動的集中管控需求。

企業內部會議：上市公司董事會、監事會、重大決策會議的規範錄製與安全存檔，支持審計合規與事後查證。

2.4 適用人群

本系統的主要目標用戶群體包括：

司法從業人員：法官、仲裁員、書記員、法院技術管理人員，需要精準記錄庭審過程、規範管理庭審檔案。

企業法務部門：公司法務總監、合規專員、董秘辦公室工作人員，負責重大會議錄製與法律文書存檔管理。

行政運維團隊：會議中心技術運維人員、系統管理員、IT 支持工程師，負責會議系統的部署運維與日常管理。

翻譯與速錄服務機構：為國際會議提供同聲傳譯、速錄服務的專業機構，可利用本系統的 AI 轉寫翻譯功能提升工作效率。

第三章　功能模塊詳細介紹

3.1 多路同步錄製模組

本模組為系統核心錄製引擎，支援多路視頻音頻信號的並髮採集與統一管理。

每房間 12 路攝像頭同步錄製：系統採用 C/S 架構，錄製作業由 C-Client 客戶端執行，每個房間支援最多 12 路 IP Camera 或 NVR 設備的同步連接與錄製。全系統最多 10 個房間可同時並行錄製，最高支援 120 路攝像頭並發錄製。攝像頭數量可根據實際場地規模靈活配置（如小型會議室 2-4 路、標準會議室 6 路、大型會議廳 8-12 路），不受固定路數限制。

Recording Agent 雙模式架構：C-Client 內建的 Recording Agent 支援兩種錄製模式，可根據房間設備配置靈活切換：

模式	說明	技術實現	硬體需求
NVR 模式	Recording Agent 通過 AXIS ACS Pro API 控制 NVR 進行錄製，自身不拉流	AXIS ACS Pro API（端口 29204）深度集成，S-Server 從 NAS 讀取 NVR 錄製文件	極低，普通辦公電腦即可
攝像頭模式	Recording Agent 通過視頻處理引擎直接拉取 RTSP 流錄製	Python + 視頻處理引擎 + Windows 服務管理器 + MP4 容器封裝	需要 GPU 硬解支援

設備接入採用標準 RTSP 協議，兼容 AXIS、大華、海康威視、宇視等市場主流品牌設備。錄製過程中，各路視頻流獨立編碼存儲，同時保留精確時間戳元數據，確保後續多機位同步回放時的畫面對齊精度達到毫秒級別。

4-64 聲道 Dante 數字音頻採集：系統深度整合 Dante 數字音頻網絡協議，全系統提供 64 個聲道槽位，支援從 4 聲道基礎配置到 64 聲道專業配置的全方位音頻錄製需求。每個房間可獨立分配 4-8 聲道（如現場聲、翻譯通道等），聲道按房間順序分配，不交叉映射。Dante 作為 AoIP（Audio over Internet Protocol）領域的領先標準，提供低延遲、高保真、網絡化佈線的音頻傳輸體驗。系統內置 64 通道聲道分配器，支持靈活的音軌映射配置，滿足最多 10 個房間並行錄製的音頻輸入組合需求。

音頻處理管線（Audio Processing Pipeline）：

系統音頻從採集到存儲採用三路並行管線架構：



DSP 處理鏈：每個音頻通道獨立配置 DSP 參數，包含前增益（Pre-Gain）、噪聲門（Noise Gate）、低切濾波器（Low-Cut Filter）、壓縮器（Compressor）、後增益（Post-Gain）五級處理。DSP 處理為品質保障性操作，確保

錄製品質；處理參數隨錄製元數據一併歸檔，支持事後審計追溯。

三路輸出說明：

輸出	格式	用途	說明
WAV 歸檔	48kHz/24bit/多聲道	取證備份	原始品質，10 分鐘分段
AAC 播放	48kHz/多聲道	StudioPlayer 回放	瀏覽器原生支持，可切聲道
ASR 轉寫	16kHz/16bit/mono	離線語音轉寫	按需觸發，降採樣至引擎輸入規格

視頻音頻分離存儲：音頻為房間級資源，不與單一攝像頭綁定。存儲結構如下：

```
recordings/room_01/2026-05-14/session_001/
├── cam_01.mp4      ← 視頻 only
├── cam_02.mp4      ← 視頻 only
├── ...
├── audio_4ch.m4a   ← 4ch AAC（播放用，可切聲道）
└── audio_4ch.wav   ← 4ch WAV（歸檔取證用）
```

StudioPlayer 回放時同時載入各路視頻與音頻，通過時間戳精確對齊，音頻通道可自由切換監聽。

音視頻精確同步：採用統一的 NTP 時間源同步機制，確保所有接入設備的時間基準一致。視頻流與音頻流在錄製階段即完成時間戳標記，後續回放時通過時間軸對齊算法實現毫秒級精度的聲畫同步。

多房間並行錄製：系統支援多個會議房間的獨立錄製作業同時進行，各房間之間完全隔離、互不干擾。集中管控界面提供所有房間的錄製狀態總覽，運維人員可在單一 Dashboard 完成全場景的錄製調度與監控。

3.2 即時串流預覽模組

本模組提供會議現場的實時視頻監看能力，支援運維人員與授權人員對會議進程的即時掌握。

零轉碼串流引擎：串流傳輸採用零轉碼串流引擎作為核心引擎，該引擎實現 RTSP 協議到 WebRTC 協議的無縫轉換，全程無需視頻編碼轉碼操作。零轉碼架構帶來兩項顯著優勢：其一，端到端延遲控制在 50-200 毫秒區間，接近實時的監看體驗；其二，極低的 CPU/GPU 資源佔用，單台標準服務器即可支撐多路並發串流。

動態網格佈局：預覽界面採用響應式網格設計，根據當前接入的攝像頭數量自動調整佈局模式。支援從 1x1 單畫面到 4x3 十二宮格的自適應排列組合，用戶可根據監看需求靈活切換。在畫面數量較少時，系統自動放大單個畫面尺寸，確保監看清晰度。

主畫面切換與全屏模式：用戶可通過點擊操作指定任意一路畫面為主監視窗口，雙擊任意畫面則進入全屏模式。全屏模式下保留系統控制面板入口，可隨時退出返回標準視圖。

自動降級機制：當零轉碼串流引擎因網絡環境或設備兼容性問題無法正常運作時，系統自動降級為截圖模式，週期性抓取視頻幀作為靜態圖像展示，確保監看功能的可用性不受單點故障影響。

3.3 StudioPlayer 多路回放模組

本模組為系統核心回放工具，支援多機位同步審查與多聲道音頻回放。

12 路視頻同步回放：支援同時回放最多 12 路會議錄製視頻，畫面嚴格時間同步，播放控制（播放、暫停、停止、拖動）對所有通道即時生效。同步回放機制確保審查人員能夠從多個視角完整還原會議場景。

多聲道音頻同步播放：支援 1-8 聲道音頻同步播放，與房間 Dante 聲道配置對應，從單聲道訪談室到 8 聲道多語言會議室均可覆蓋。音頻輸出支援現場原聲與翻譯通道的自由切換，系統自動識別翻譯發言時的語言種類，匹配對應翻譯版本時自動開通直通播放。播放器內建聲道路由引擎，通過 ChannelSplitter 將多聲道音頻拆分為獨立通道，每個通道配備獨立 GainNode。用戶切換聲道時即時調整對應通道增益值，無延遲、無需重新載入。支持單聲道監聽與多聲道混音監聽兩種模式。

預渲染音頻波形顯示：播放界麵包含音頻波形可視化面板，採用 Peak Hold 峰值保持技術，配合 dB 歸一化處理與漸變色柱狀圖呈現。波形圖像預先渲染，確保播放過程中的圖形渲染零延遲。用戶可通過波形視圖直觀感知發言節奏與音量變化。

即時字幕疊加（SubtitleDisplay）：AI 轉寫結果以滾動字幕形式實時疊加於視頻畫面之上，字幕位置與時間軸嚴格對齊，支持字體大小、顏色、背景透明度等樣式自定義配置。字幕內容隨播放進度自動推進，暫停時字幕同步停止。

PTZ 雲台控制面板：集成對支援 PTZ（Pan-Tilt-Zoom）雲台功能的網絡攝像頭進行遠程控制。控制面板提供方向搖杆、變焦按鈕、預設位調用等常用功能，支援最多 16 個雲台預設位的保存與一鍵調用。回放過程中如需調整監看角度，可直接通過控制面板操作，無需切換至獨立設備管理界面。

3.4 AI 轉寫與翻譯模組

本模組為系統核心智能引擎，實現會議語音的自動轉寫與多語言翻譯。

本地部署離線語音轉寫引擎：語音轉寫功能基於離線語音轉寫引擎構建，模型部署於本地服務器，完全離線運行。會議音頻無需上傳至任何外部服務器，從根本上杜絕數據外泄風險。轉寫引擎原生支持普通話（Mandarin）、粵語（Cantonese）的語音識別，兩種中文變體均經過大規模語料訓練，識別準確率在安靜環境下達到行業領先水平。粵語轉寫支持粵語特有詞彙、語氣助詞、中英夾雜的精準識別；繁體中文輸出，符合港澳地區及台灣用戶使用習慣；支持中文法律術語、司法專有名詞的精準識別（如「無罪推定」、「上訴」、「抗辯」等）。轉寫引擎接收 16kHz/16bit/mono 格式的音頻輸入。系統在轉寫排程階段自動從多聲道音頻中提取目標聲道，執行降採樣與單聲道轉換後提交至轉寫引擎。每個聲道獨立轉寫，互不影響。

轉寫結果時間軸對齊：轉寫文本自動分割為獨立語句段落，每個段落精確標註對應的起始時間與結束時間。時間軸元數據支持與視頻播放進度的無縫對接，用戶點擊任意轉寫段落即可跳轉至對應播放位置。

離線翻譯引擎：翻譯功能採用離線翻譯引擎，支援中文、英文、葡文在內的 33 種語言對翻譯。中文（普通話/粵語）↔ 英文、中文（普通話/粵語）↔ 葡文 的雙向翻譯為核心語言對，優先優化。粵語口語到書面中文的自動轉換能力，使粵語發言轉換為更易閱讀的書面格式。模型內存佔用控制在 1GB 以內，單台標準服務器即可承載翻譯任務。翻譯引擎同樣部署於本地，確保翻譯過程的數據安全。

翻譯通道自動直通：系統自動識別各翻譯通道的發言語言，當檢測到翻譯發言的語言與當前監聽用戶的目標語言匹配時，自動開通翻譯音頻直通，無需手動切換聲道。

3.5 AI 降噪模組

本模組提供兩種降噪處理模式，適配不同作業階段的噪聲控制需求。

即時 AI 降噪引擎：即時 AI 降噪引擎為 AI 驅動的語音噪聲抑制技術，專為即時語音處理場景優化。即時降噪作為 DSP Chain 中 Noise Gate 環節的一部分嵌入音頻管線，實現實時監聽降噪，處理延遲低至 10 毫秒。降噪處理在音頻信號進入編碼器之前完成，確保錄製素材的聲音品質。

離線深度降噪引擎：對於已錄製的音頻素材，系統提供離線深度降噪處理能力。離線降噪對 WAV 歸檔檔案進行深度處理，離線深度降噪引擎基於 AI 驅動技術，在保持語音清晰度的前提下實現高達 39.5dB 的噪聲衰減。離線處理模式支援批量任務排程，可在非工作時段對歷史素材進行集中降噪優化。

標準編程接口：降噪模組提供標準編程接口，開發人員可通過 API 調用方式將降噪能力集成至自有系統。接口設計遵循簡潔原則，三行程式碼即可完成單音頻文件的降噪處理。

3.6 全文搜索模組

本模組提供對轉寫文本的強大檢索能力，實現會議內容的精準定位。

內建全文搜索引擎：系統後端集成內建全文搜索引擎，對所有轉寫文本建立倒排索引。搜索引擎採用相關性排序算法，根據關鍵詞匹配度與文檔權重計算相關性得分，返回結果按相關度降序排列。

時間軸定位跳轉：搜索結果不僅呈現匹配的文本片段，同時標註該片段在原始錄製中的精確時間戳。用戶點擊任意搜索結果，系統自動將播放器跳轉至對應時間點，無需手動拖動進度條。

多維度篩選功能：搜索界面支援按房間名稱、錄製日期範圍等條件進行結果篩選。對於多房間並行部署的機構，可快速定位特定會議室的歷史錄製內容。

3.7 系統監控模組

本模組為運維人員提供全場景運行狀態的可視化監控。

房間狀態總覽 Dashboard：主監控界面以卡片式佈局呈現所有會議房間的即時狀態，包括當前錄製進度、已錄製時長、即時幀率、在線設備數量等核心指標。狀態卡片支持自定義排序與分組，運維人員可根據機構實際佈局調整展示邏輯。

存儲空間監控：系統實時追蹤錄製存儲空間的使用情況，提供已用容量、剩餘容量、預估剩餘錄製時長等數據。當存儲空間低於預設閾值時，系統自動觸發告警通知。

設備分類監控：監控面板按照設備類型進行分類展示，包括網絡攝像頭狀態矩陣、音頻系統通道狀態、Dante 設備連接狀態、NAS 存儲設備健康度等。各設備支持一鍵詳情報告調取，快速定位故障節點。

看門狗多渠道告警：系統內建看門狗（Watchdog）監控引擎，支援郵件、Webhook、飛書等多種告警通知渠道。運維人員可配置差異化的告警規則，例如：設備離線觸發郵件通知、存儲空間不足觸發飛書消息、錄製異常中斷觸發 Webhook 回調。告警記錄完整存檔，支持事後追溯分析。

3.8 配置中心模組

本模組提供系統全局配置的集中管理與版本控制能力。

全局配置管理：系統所有可調整參數（錄製參數、串流參數、AI 模型參數、UI 樣式參數等）集中收斂於配置中心。管理員在統一界面完成所有配置操作，告別分散在各模組深層菜單的傳統模式。

版本控制機制：配置中心內建版本控制功能，每次配置變更自動生成歷史快照。管理員可隨時回溯至任意歷史版本，比較不同版本的配置差異。一鍵恢復功能支援將系統配置還原至指定歷史狀態。

Agent 配置管理：系統支援對各房間獨立部署的邊緣 Agent 進行遠程配置管理。配置變更後自動推送至目標 Agent，無需逐台設備手動操作。配置版本與房間 Agent 版本自動關聯記錄。

配置推送與追溯：配置下發支援灰度發布策略，可先在測試環境驗證再批量推送至生產環境。所有配置操作記錄完整審計日誌，精確記錄操作人、操作時間、變更內容。

3.9 審計日誌模組

本模組為系統安全合規提供完整的操作痕跡記錄能力。

全量操作日誌記錄：系統自動記錄所有用戶操作行為，包括但不限於：用戶登入與登出、錄製任務啟動與停止、配置文件修改、設備連接與斷開、搜索查詢操作、導出與下載請求等。日誌內容涵蓋操作人身份、操作時間、操作類型、操作對象、操作結果等完整要素。

分頁篩選查詢：日誌查詢界面支援按時間範圍、操作類型、操作人等條件進行高級篩選。查詢結果支援分頁展示，單頁顯示條目數可自定義調整。

日誌導出功能：審計日誌支援導出為標準格式文件，便於歸檔保存或移交第三方審計機構。導出操作本身亦被記錄，確保審計鏈條的完整性。

司法合規支撐：日誌模組的設計充分借鑒司法取證標準，所有日誌採用防篡改存儲機制，時間戳採用權威時間源同步。記錄內容可直接作為電子證據提交，滿足司法程序對電子數據真實性、完整性、合法性的取證要求。

3.10 設備管理模組

本模組提供對系統所接入各類設備的統一發現、配置與運維管理。

網絡攝像頭統一管理：系統自動發現並列舉接入網絡的 IP Camera 與 NVR 設備，支援 ONVIF 標準協議及主流廠商私有協議。設備管理界面呈現設備的基本信息、網絡狀態、視頻參數、碼流配置等，並支援遠程修改設備參數。已驗證兼容大華、海康威視、宇視等市場主流品牌型號。

NVR 設備整合：對於已部署 NVR（Network Video Recorder）存儲設備的機構，系統支援直接接入 NVR 進行錄製管理。接入後，NVR 設備中的歷史錄製文件可在系統內統一檢索與播放，無需切換至原生 NVR 管理界面。
Recording Agent 支援以下 NVR/設備控制协议：

設備類型	通信協議	端口	集成深度
AXIS 安訊士	ACS Pro API（服務器級）	29204	✅ 深度集成
AXIS 安訊士	VAPIX API（設備級）	80/443	✅ 深度集成
海康威視	ISAPI（HTTP REST）	80/443	🔄 規劃支持
大華	NetSDK（C DLL）	37777	🔄 規劃支持
通用設備	ONVIF Profile G	554	🔄 規劃支持

Dante 音頻設備管理：系統提供 Dante 音頻網絡的設備發現與拓撲呈現功能。管理員可查看 Dante 網絡中所有音頻設備的連接狀態、通道映射、采樣率配置等參數，並支援遠程調整通道分配方案。

NAS 存儲設備管理：錄製文件的集中存儲採用群暉（Synology）NAS 方案，配備 2×10GbE 網口鏈路聚合，提供全系統統一存儲工作目錄。NVR 錄製的視頻文件直接寫入 NAS，S-Server 的 Dante 音頻錄製、AI 處理產物、合成打包結果均直接讀寫 NAS，消除中間文件傳輸環節。NAS 提供設備的添加嚮導與健康監控功能，自動追蹤容量使用、讀寫性能、RAID 狀態等指標。

第四章　核心優勢與行業亮點

4.1 全離線 AI 處理架構

本系統最顯著的技術差異化優勢在於完整的本地化 AI 處理能力。語音轉寫採用離線語音轉寫引擎、翻譯採用離線翻譯引擎、兩項核心 AI 能力均部署於客戶自有設備運行。相比依賴雲端 AI 服務的競品方案，本系統在數據安全層面構建了根本性壁壘。

司法機構、大型企業對會議內容的保密性要求遠高於一般商業場景。會議錄製往往涉及商業機密、案件細節、個人隱私等敏感信息，任何環節的數據外泄都可能造成嚴重後果。本系統的離線 AI 架構確保錄製素材從採集、編碼、存儲到 AI 處理的完整生命週期均在機構內網完成，真正意義上實現「數據不出門」。

4.2 司法合規深度適配

系統的設計之初即以司法錄取場景為核心標杆，在功能設計與技術實現上全面對標司法合規要求。

時間戳精度的毫秒級對齊確保錄製素材的法庭證據效力。司法程序對錄製文件的時間標注有嚴格要求，任何時間偏差都可能成為程序瑕疵。本系統採用 NTP 統一授時機制，所有視頻音頻流的時間戳溯源至同一時間基準，為錄製素材的時間真實性提供技術保障。

完整的審計日誌體系覆蓋系統運維的全操作維度。從用戶登入認證到配置變更、從錄製任務管理到文件導出，每一步操作均有精確到秒的時間記錄與操作人身份標注。這種不可否認的操作痕跡記錄體系，完全滿足司法機構對信息系統安全審計的法定要求。

4.3 多語言即時翻譯能力

系統內置的離線翻譯引擎支援 33 種語言對的即時翻譯，這一能力在同類司法錄製軟件中處於領先地位。特別針對涉及中文、英文、葡文使用者的仲裁與外交場景，系統提供翻譯通道自動識別與直通功能，大幅降低會議期間的語言障礙。

系統對中文（含粵語）的處理深度遠超同類產品。粵語作為漢語的重要變體，擁有獨特的發音體系、詞彙系統和語法結構，對語音識別技術提出更高要求。本系統針對粵語開發專項優化方案，支持粵語特有詞彙（如「我哋」、「幾時」、「唔該」）、語氣助詞（如「喇」、「啰」、「啩」）以及中英夾雜的精準識別，有效解決其他系統在粵語識別上的普遍痛點。

繁體中文界面的原生支持確保港澳地區及台灣用戶的使用體驗。系統 UI、轉寫報告、字幕疊加等所有文本輸出均為繁體中文，符合目標市場的語言習慣，避免簡繁轉換帶來的用詞不當問題。

翻譯結果與時間軸的自動對齊使得譯文可與原始錄製精確關聯。審查人員在回放錄製時，轉寫原文與翻譯文本同步呈現，點擊任一語種的轉寫段落即可定位至原始發言時間點。這種雙語對照檢索能力極大提升了多語言會議內容的查閱效率。

4.4 零轉碼串流延遲優勢

零轉碼串流引擎的採用使本系統在即時串流場景擁有顯著延遲優勢。50-200 毫秒的端到端延遲相比傳統 RTSP→HLS→RTMP 的級聯轉碼方案（通常 3-10 秒延遲）提升了一到兩個數量級。

低延遲串流在會議監控場景中具有重要實用價值。運維人員在控制室即可近乎實時地掌握會場狀態，一旦發現異常情況可在第一時間響應處置。對於庭審錄製場景，低延遲監控使書記員能夠同步確認錄製質量，發現問題即時糾正，避免事後補救。

4.5 模塊化架構的靈活擴展

本系統採用高內聚低耦合的模塊化 API 架構設計，各功能模組之間通過標準化接口通信。這種架構帶來三項顯著優勢：

按需啟用功能：機構可根據實際需求選擇啟用特定模組，無需為閒置功能支付額外成本。單機版用戶從基礎錄製功能起步，隨業務髮展再平滑升級至專業版或企業版。

定製化開發便捷：模組化架構便於進行功能定製開發。機構可基于系統開放的 API 接口接入自有設備、整合現有 IT 系統、或者開發專屬增值功能。

運維故障隔離：單一模組的故障不會蔓延至整個系統。降級機制（如零轉碼串流引擎不可用時的截圖模式）確保系統在部分模組異常時仍維持核心功能可用。

4.6 一站式全流程覆蓋

從設備接入、錄製執行、即時監控、AI 處理、內容檢索到合規歸檔，本系統覆蓋會議錄製工作流的全環節。這種一站式設計避免了機構採用多套獨立系統拼接方案所帶來的集成複雜性、數據一致性風險與運維成本。

單一系統意味著單一接口、單一培訓、單一支持。機構的 IT 團隊與業務用戶僅需熟悉一套系統即可完成全部相關工作。這種簡化對人員編制精簡、培訓成本控制、運維效率提升均有直接貢獻。

第五章　典型應用場景示例

系統面向粵港澳大灣區及海內外華人市場設計，港式粵語、普通話、繁體中文為原生支持的核心語言組合。以下六個通用場景類型，分別代表不同的核心需求——司法存證、三語仲裁、機密會議、法律談判、學術論壇、行政聽證——展示系統如何針對差異化痛點提供針對性方案。

場景一：庭審——港式粵語司法存證

核心痛點：鏈式存證不可篡改 + 港式粵語精準轉寫 + 法定歸檔合規

場景描述：庭審過程以港式粵語為主，法官、控辯雙方、證人均以粵語交流，偶有翻譯人員以英文發言。庭審記錄須具備完整證據鏈——從錄製到歸檔，每個環節不可篡改、全程可追溯。外籍陪審員需通過即時字幕理解庭審進程。

港式粵語特色處理： - 原生支持港式粵語口語轉寫，精準識別「咁」「唔係嘛」「即係」等口語助詞，自動轉換為規範繁體中文書面語 - 支持法例引用格式的語音識別（如「第 X 章」「第 X 條」），法律術語準確率遠超通用語音引擎 - 粵語語者識別：自動區分法官、控方、辯方、證人發言，轉寫文本按發言人分段標記 - 繁體中文原生輸出，無需簡繁轉換，避免轉換誤差

系統配置： - 1 個庭室 - 6 路攝像頭（法官席、控方席、辯方席、證人席、全景、備用） - 8 聲道 Dante 音頻（法官話筒、控方話筒、辯方話筒、翻譯話筒、證人話筒、現場環境、備用 × 2）

工作流程：庭審開始後，系統以港式粵語模式即時轉寫，口語自動規範化為書面語；翻譯人員英文發言時同步中英翻譯；全程錄製文件附帶加密哈希校驗，生成不可篡改的證據鏈；最終輸出帶雙語字幕的庭審錄製文件及結構化轉寫文檔。

客戶價值： - 書記員工作量降低 70%，港式粵語轉寫精度滿足司法歸檔標準 - 鏈式存證確保庭審記錄法律效力，全程可追溯 - 雙語字幕輔助外籍陪審員，提升司法透明度 - 完全離線架構，庭審數據不出網絡，確保數據主權

場景二：仲裁——大灣區三語跨境

核心痛點：普通話-港式粵語-英文三語無縫切換 + 跨法域合規 + 裁決文書速出

場景描述：跨境商事仲裁中，普通話、港式粵語、英文三語頻繁切換——同一場仲裁內，內地律師講普通話、港澳律師講粵語、海外仲裁員講英文，三方需實時理解彼此陳述。裁決文書須同時出具三語版本，滿足不同法域的歸檔要求。

大灣區三語特色處理： - AI 自動識別發言語言，普通話、港式粵語、英文分通道獨立轉寫，無需手動切換 - 粵語通道專項優化：識別「入面」「而家」「咁樣」等大灣區粵語用詞，區別於書面中文 - 轉寫結果實時翻譯為其餘兩種語言，三語字幕同步疊加 - 仲裁文書自動生成三語對照版本，符合跨境仲裁國際慣例

系統配置： - 3 個仲裁庭，每庭獨立配置 - 每庭 4 路攝像頭（仲裁員席、申請人席、被申請人席、全景） - 每庭 8 聲道 Dante 音頻

工作流程：AI 即時識別語種，三語分通道轉寫並實時互譯；發言人自動標記（內地/港澳/海外）；仲裁結束後一鍵生成三語裁決文書草稿。

客戶價值： - 三語無縫切換，無需專人語言轉換，仲裁效率提升 - 三語裁決文書自動生成，準備週期縮短 50% - 完整三語記錄滿足跨境仲裁國際慣例，避免翻譯爭議 - 大灣區粵語專項優化，港式口語不漏字、不誤識

場景三：會議——機密分級與紀要速出

核心痛點：商業機密分級管控 + 監管合規存檔 + 會後紀要速出

場景描述：董事會或高管會議涉及重大商業決策、關聯交易審議、財務報告審批等敏感議程。會議內容機密等級極高，錄製文件須分級管控、權限嚴格；會議記錄需符合上市規則合規存檔要求，具有法律證據效力。會後董秘團隊需在最短時間內出紀要。

機密會議特色處理： - 錄製按機密等級分級存儲，不同權限人員僅可訪問授權內容 - 完整審計日誌記錄每一次訪問、下載、播放操作，滿足監管審查 - AI 自動生成結構化紀要模板：發言人標記 + 時間軸 + 議題分段 - 粵語/普通話混合會議自動語種切換，董秘無需事後逐句翻聽

系統配置： - 1 個會議室 - 4 路攝像頭（主席席、CEO 席、秘書席、全景） - 4 聲道 Dante 音頻（現場擴音拾取、獨立發言拾音 × 3）

工作流程：全程錄製，AI 以粵語或普通話即時轉寫；自動標註發言人與時間，生成結構化紀要模板；審計日誌全程記錄操作痕跡；會議結束即出紀要草稿。

客戶價值： - 機密分級管控，錄製內容不外洩，滿足合規要求 - 會議記錄不可篡改，提供可信賴的電子證據 - 董秘團隊工作效率提升 60%，會後紀要當日出 - 監管審查或法律訴訟中有完整證據鏈

場景四：談判——法律條文逐字核對

核心痛點：法律術語雙語逐字對照 + 條款時間軸標記 + 爭議證據留存

場景描述：中英雙方法務團隊進行重大商務談判，涉及合同條款的逐字確認。法律條文精準表述至關重要——「不可抗力」與「Force Majeure」的對應、「違約責任」與「Breach of Contract」的範圍，任何翻譯誤差都可能導致巨額損失。會後需完整保留談判記錄作為合同爭議證據。

法律精準特色處理： - 中英雙語同步轉寫，逐句對照，條文級別對齊 - 法律術語庫精準識別（如「標的」「違約責任」「不可抗力」與「Indemnity」「Breach of Contract」「Force Majeure」自動配對） - 時間軸自動標記重要條款段落，一鍵定位關鍵條文 - 談判全程錄製附哈希校驗，具備法律證據效力

系統配置： - 1 個會議室 - 4 路攝像頭（中方代表席、外方代表席、翻譯席、全景） - 4 聲道 Dante 音頻（中文原聲、英文原聲、翻譯通道、備用）

工作流程：中英文同時轉寫，逐句對照；法律術語自動配對標記；時間軸標記重要條款；最終輸出雙語對照文本 + 帶時間軸標記的錄製文件。

客戶價值： - 法律條文逐字可查，雙語對照避免翻譯誤差 - 時間軸標記一鍵定位，合同審查效率倍增 - 全程錄製作為中立證據，潛在爭議中有據可查 - 節省專業法律翻譯費用，降低談判成本

場景五：論壇——三語字幕即時分發

核心痛點：多會場並行 + 三語即時字幕分發 + 引用歸檔

場景描述：國際學術或行業論壇中，演講者使用港式粵語、普通話或英文輪流發言，聽眾來自世界各地，需實時看到自己語言的字幕。論壇分多個會場並行，統一管理。內容需結構化歸檔，支持後續引用。

論壇場景特色處理： - 多會場並行錄製，統一管理平台集中監控 - AI 自動識別演講者語種，港式粵語/普通話/英文分通道轉寫並實時互譯 - 三語字幕即時分發，聽眾按需切換顯示語言 - 歸檔含時間軸 + 發言人 + 三語對照文本，支持引用格式導出

系統配置： - 2 個會場並行錄製 - 每會場 4 路攝像頭（主講人席、主持人席、提問席、全景） - 每會場 4 聲道 Dante 音頻

工作流程：AI 自動識別語種並分通道轉寫；三語字幕即時生成並分發至會場顯示屏；論壇結束後自動生成結構化三語存檔，含發言人標記與時間軸。

客戶價值： - 三語即時字幕消滅語言障礙，聽眾體驗大幅提升 - 多會場統一管理，主辦方一站式監控 - 三語存檔支持學術引用，論壇內容資產化 - 港式粵語演講精準轉寫，內容不因語言流失

場景六：聽證——中葡雙語法定合規

核心痛點：中葡雙語法定要求 + 行政程序全程可追溯 + 多語種合規歸檔

場景描述：行政聽證會涉及執照續期、違規處罰等事項，參與者使用港式粵語、普通話、葡文三種語言。記錄須符合法律對中葡雙語記錄的法定要求——這不是可選項，而是法律強制規定。聽證記錄需全程可追溯，支持行政覆議和司法救濟程序。

中葡雙語特色處理： - 三語同步轉寫（港式粵語、普通話、葡文），中葡互譯為核心語言對優先處理 - 粵語口語自動轉換為規範繁體中文，與葡文對照輸出，符合法定雙語格式 - 合規歸檔：附完整審計日誌與哈希校驗，滿足行政程序法律要求 - 聽證記錄支持行政覆議調取，全程可追溯

系統配置： - 1 個聽證室 - 6 路攝像頭（主持人席、申請人席、代理人席、翻譯席、見證席、全景） - 8 聲道 Dante 音頻（粵語發言、普通話發言、葡文翻譯、現場拾音、備用 × 4）

工作流程：三語同步轉寫，港式粵語自動規範化為書面中文；中葡互譯優先處理；翻譯結果以繁體中文和葡文同步呈現；合規歸檔，附完整審計日誌與加密校驗。

客戶價值： - 完全滿足法律對中葡雙語記錄的法定強制要求 - 全程可追溯的聽證記錄，支持行政覆議和司法救濟 - 監管機構工作效率提升，避免語言障礙導致的聽證延誤 - 為當事人提供公正、透明的行政程序記錄

第六章　系統兼容性與運行環境

5.1 架構概述

本系統採用 C/S（Client-Server）架構設計，支援三種部署模式以滿足不同規模機構的需求：

- **S-Server（服務端）**：macOS（Apple Silicon），Mac mini M4 Pro 為指定硬件平台，作為「調度大腦」管理所有 C-Client Agent 的註冊、心跳、指令下發與狀態彙總
- **C-Client（客戶端）**：Windows 平台，內建 Recording Agent（Python 後台服務），通過 Windows 服務管理器註冊為 Windows 服務開機自啟；提供 Web 管理界面供操作員訪問

Agent 通信架構： - **HTTP REST 通道**：Agent 註冊、心跳、狀態上報，支援指數退避重連與離線緩衝 - **WebSocket 通道**：S-Server 即時指令推送（錄製開始/停止等），斷線自動重連 - 操作員通過 Web 瀏覽器連接 S-Server 統一管理所有 C-Client

離線容忍（Offline Tolerance）：Agent 與 S-Server 斷連後，本地持續錄製不受影響；連接恢復後自動補報緩衝期間的錄製狀態。

三種部署模式

部署模式	說明	適用場景
最簡模式	單台 Mac mini 完成全部工作，無需 C-Client、NAS、NVR	單房間小型法庭、個人律師事務所、功能評估演示、開發測試
NVR 模式	S-Server + C-Client + NVR + NAS	已有 NVR 基礎設施的機構
攝像頭模式	S-Server + n台 C-Client + NAS（無 NVR）	大型多房間部署

最簡模式架構特點

最簡模式是本系統的單機部署方案，專為資源需求較小的場景設計：

- **單機全功能**：1 台 Mac mini M4 Pro 完成錄製、串流、AI 推理、回放全部功能
- **本地存儲**：直接使用本機 NVMe SSD 存儲（1-4 TB），無需 NAS
- **S-Server 自錄**：S-Server 內建錄製引擎，直接拉取 RTSP 流錄製，無需 C-Client Agent
- **最大容量**：支援 1-2 個房間、2-6 路攝像頭並發
- **靈活操作**：可在同一台 Mac mini 上開瀏覽器操作，或從其他電腦遠程訪問

最簡模式是評估系統能力、展示產品功能的理想方案，也是個人用戶和小型機構的經濟之選。

5.2 S-Server 服務端硬件要求

S-Server 作為全系統調度大腦，管理所有 C-Client Agent 的註冊、心跳、指令下發與狀態彙總。其硬件需求隨部署規模遞增，以下按典型場景與極限場景分別給出配置：

典型場景（6 房間 × 6 路，部分 NVR）

項目	配置
硬件平台	Mac mini M4 Pro
CPU	Apple M4 Pro（10 核心）
統一記憶體	**24 GB**
存儲	512 GB SSD + 1GbE NAS
網卡	千兆網卡
GPU	內建 Metal GPU（M4 Pro）

極限場景（10 房間 × 12 路 = 120 路並發）

項目	配置
硬件平台	Mac mini M4 Pro

項目	配置
CPU	Apple M4 Pro（14 核心）
統一記憶體	**48-64 GB**
存儲	1 TB SSD + 10GbE NAS（32TB+ RAID 6）
網卡	**萬兆網卡**（Mac mini 選配或 Thunderbolt 轉接）
GPU	內建 Metal GPU（M4 Pro）

說明：

- Mac mini M4 Pro 是 S-Server 的指定部署平台，利用 Apple Silicon 的硬件加速推理框架實現 AI 推理性能優化
- **典型場景下 24 GB 記憶體即可滿足**：AI 轉寫引擎 + 翻譯引擎 + 降噪引擎 + 零轉碼串流引擎同時運行約佔 17 GB，留有充足餘量
- **極限場景需 48-64 GB 記憶體**：120 路串流中轉 + 多實例 AI 推理需約 30+ GB
- **網絡是關鍵瓶頸**：典型場景 1GbE 即可（約 208 Mbps，佔用 21%）；極限場景 1GbE 不足（約 640 Mbps），必須升級 10GbE
- 存儲空間需根據錄製量估算：典型場景每日約 518 GB，極限場景每日約 1.7 TB，建議配置外接 NAS 擴展
- 詳細容量分析見《CAPACITY_ANALYSIS.md》

5.3 C-Client 客戶端硬件要求

C-Client 負責各房間的攝像頭錄製執行，內建 Recording Agent（Python 後台服務）。技術棧說明：

組件	技術選型	說明
運行環境	Python 3.11+	Agent 核心代碼執行環境
視頻處理	視頻處理引擎	RTSP 流拉取、編碼封裝
Windows 服務	Windows 服務管理器	將 Python 進程註冊為 Windows 服務，開機自啟
容器格式	MP4	錄製文件封裝格式
通信模組	HTTP REST + WebSocket	與 S-Server 雙通道通信

C-Client 支援離線錄製（Offline Recording）：當 S-Server 不可達時，Agent 本地持續錄製；連接恢復後自動補報緩衝期間的錄製狀態。

部署方式取決於房間是否配備 NVR：

直連錄製模式（無 NVR，C-Client 拉流錄製）

項目	最低配置	推薦配置
操作系統	Windows 10 64-bit	Windows 11 64-bit
CPU	Intel Core i5 第 10 代 / AMD Ryzen 5 5600	Intel Core i7 第 12 代 / AMD Ryzen 7 5800

項目	最低配置	推薦配置
內存	8 GB	16 GB 或以上
顯卡	NVIDIA GTX 1650（每房間 ≤6 路）	NVIDIA RTX 4060（每房間 ≤12 路）
存儲	500 GB NVMe SSD	1 TB NVMe SSD 或以上
瀏覽器	Google Chrome 90+ / Microsoft Edge 90+	Google Chrome 最新版
網卡	千兆網卡	千兆網卡
顯示器	1920×1080	2560×1440 或以上（多路預覽場景）

NVR 託管模式（有 NVR，C-Client 僅做管理調度）

項目	配置
操作系統	Windows 10 64-bit 或以上
CPU	Intel Core i5 第 8 代或同等
內存	8 GB
顯卡	集成顯卡即可（無需硬解）
存儲	10 GB 可用空間
瀏覽器	Google Chrome 90+
網卡	千兆網卡

說明：

- **每房間建議部署 1 台 C-Client**（分佈式架構），故障隔離、負載分散
- **有 NVR 的房間**，C-Client 不拉流錄製，硬件需求大幅降低，普通辦公電腦即可
- **無 NVR 的房間**，C-Client 需 GPU 硬解支援多路 RTSP 錄製：RTX 4060 可承擔 12 路 1080p，RTX 3060 可承擔 6 路
- 內存需求取決於同時錄製/預覽的視頻路數，6 路以下 8 GB 可滿足，12 路建議 16 GB

5.4 設備兼容性

IP Camera / NVR：系統通過標準 RTSP 協議接入網絡攝像頭與 NVR 設備。已驗證兼容設備品牌包括但不限於：**AXIS（安訊士）**、大華（Dahua）、海康威視（Hikvision）、宇視（Uniview）、華為（Huawei）、科達（Kedacom）等主流廠商的全線產品。對於採用私有協議的設備，系統提供靈活的插件擴展機制。

AXIS 攝像頭深度集成說明：

- 系統支持 AXIS VAPIX API 直連控制，實現 PTZ 雲台操控、預設位管理、流配置檔動態調整等高級功能
- 支持 AXIS Camera Station Pro（ACS Pro）服務器級 API（端口 29204），可通過 ACS NVR 進行錄像管理、歷史回放、遠程配置

- AXIS 設備自動發現（ONVIF + AXIS 發現服務）、邊緣存儲錄製管理（SD 卡/NAS 故障轉移）
- RTSP 流 URL 格式：`rtsp://<servername>/axis-media/media.amp?videocodec=h264&resolution=1920x1080`
- 支持認證方式：Digest Access Authentication（HTTP/RTSP）、Basic Access Authentication（HTTPS）

Dante 音頻設備：系統相容 Audinate Dante 生態下的全系列設備，包括 Dante 麥克風、Dante 音頻接口、Dante 網絡交換機等。Dante 網絡需與系統所在主機處於同一局域網段。

NAS 存儲設備：系統以群暉（Synology）NAS 為標準存儲平台，支援 2×10GbE 鏈路聚合。同時兼容威聯通（QNAP）、華為 OceanStor、NetApp 等品牌。NAS 通過標準 SMB/CIFS 協議掛載至全系統各設備。

5.5 網絡環境要求

內網優先設計：系統核心功能（錄製、回放、AI 轉寫翻譯、搜索、監控）在純內網環境下完全可用，無需互聯網連接。這一設計充分滿足司法機構、大型企業等對網絡安全有嚴格要求的客戶群體。

外部連接按需開通：以下功能需要網絡連接，機構可根據安全策略選擇性啟用：NAS 同步備份（需連接備份網絡）、外部告警通知（需連接互聯網或專線）、License 激活與驗證（僅在激活時需要一次性互聯網連接）。

防火牆配置指引：如需穿越防火牆部署，系統僅需要對外開放有限端口。詳細的端口配置清單與防火牆配置示例見《部署指南》文檔。

5.6 離線與內網支持

本系統的離線優先架構是其核心設計特色之一。系統提供完整的功能離線可用性聲明：

完全離線可用功能：錄製模組、即時串流預覽、播放控制、AI 轉寫、AI 翻譯、AI 降噪、全文搜索、系統監控、配置管理、審計日誌、設備管理等全部核心模組均可在完全脫離互聯網的環境中正常運作。

離線條件下的性能說明：AI 轉寫與翻譯模組在離線環境下的處理性能與聯網環境完全一致，因為 AI 模型運行於本地硬體資源。離線部署不存在因網絡帶寬或延遲造成的性能瓶頸。

數據安全保障：離線架構從根本上杜絕了錄製數據通過網絡傳輸外泄的可能性。所有敏感素材的處理、存儲、檢索均在機構自有設備上完成，符合最嚴格的數據主權要求。

5.7 容量場景分級

本系統的硬件配置應根據實際部署規模分級規劃，避免過度投資或容量不足。

場景分級對照表

場景等級	房間數	每房間路數	總並發路數	NVR 分擔	S-Server 記憶體	網絡
典型場景	6	6	36	部分房間有 NVR	24 GB	1GbE
中等場景	8	8	64	約半數 NVR	32 GB	1GbE（建議 10GbE）
極限場景	10	12	120	少量 NVR	48-64 GB	10GbE 必須

NVR 分擔效應說明

當房間配備 NVR 時，Recording Agent 不直接拉 RTSP 流進行錄製，而是由 NVR 負責錄製存儲，Agent 僅做管理調度。這大幅降低 C-Client 的硬體需求：

部署模式	C-Client 職責	硬體需求	適用場景
NVR 模式	管理調度（NVR 控制）	極低，普通辦公電腦	已有 NVR 的房間
攝像頭模式	直連錄製（視頻處理引擎拉流）	需要 GPU 硬解	無 NVR 的房間

有 NVR 的房間，C-Client 不拉流錄製，CPU/GPU/記憶體需求大幅降低，普通辦公電腦即可勝任。建議採用混合模式：優先利用現有 NVR 設施，降低總體部署成本。

各場景配置建議

場景	C-Client 配置	說明
典型場景	6 台（有/無 NVR 混合）	每台管理 1 個房間
中等場景	8 台	每台管理 1 個房間
極限場景	10 台（直連錄製）	RTX 4060 以上 GPU

詳細容量分析見《CAPACITY_ANALYSIS.md》。

第七章　系統架構與運行邏輯

6.1 整體技術架構

本系統採用 **C/S（Client-Server）架構**設計，整體分為三個核心層級：

第一層：S-Server 服務端（Mac mini M4 Pro，macOS）

- 硬體抽象層：負責對攝像頭、音頻設備、存儲設備等物理硬體的統一抽象與驅動管理。通過標準化接口屏蔽設備差異，使上層邏輯無需關心設備品牌與型號的具體實現細節
- 服務端業務邏輯層：採用 Python 高性能異步後端框架構建，提供系統核心業務邏輯的處理能力。高性能異步後端框架的非同步架構確保系統在高並發場景下的響應穩定性。服務端以模組化方式組織，每個功能模組獨立為一個子服務，子服務之間通過內部消息隊列通信
- API 層：基於 RESTful 設計原則構建標準化 API 接口。接口設計遵循統一規範，包括標準化錯誤碼、統一響應格式、規範的版本控制策略等。API 文檔通過 OpenAPI 規範自動生成，支持在線調試與客戶端代碼自動生成

第二層：C-Client Recording Agent 層（Windows PC）

- Recording Agent 為 Python 後台服務，通過 Windows 服務管理器註冊為 Windows 服務，開機自動啟動

- 支援雙模式錄製：
- **NVR 模式**：通過 AXIS ACS Pro API（端口 29204）控制 NVR 錄製，自身不拉流
- **攝像頭模式**：通過視頻處理引擎直接拉取 RTSP 流，使用 GPU 硬解編碼封裝為 MP4
- 技術棧：Python 3.11+ / 視頻處理引擎 / Windows 服務管理器 / MP4 容器

第三層：Web 前端層（瀏覽器）

- Web 前端層：採用現代 Web 前端框架構建響應式 Web 管理界面。整個前端為單頁應用（SPA），頁面切換無刷新，確保流暢的用戶體驗

Agent 通信架構：

通道	用途	特性
HTTP REST	註冊/心跳/狀態上報	支援指數退避重連、離線緩衝
WebSocket	實時指令接收	斷線自動重連，S-Server 即時下發錄製指令

離線容忍（Offline Tolerance）：Agent 與 S-Server 斷連後，本地持續錄製不受影響；連接恢復後自動補報緩衝期間的錄製狀態。

串流通過 WebRTC 直連：C-Client 直接與 S-Server 上的零轉碼串流引擎建立 WebRTC 連接，無需經過 API 層中轉，大幅降低服務端負載

6.2 核心模組運行邏輯

錄製流程（C/S 分工）：

步驟	執行方	操作描述
1	操作員	在 Web UI 下發錄製指令 → S-Server API
2	S-Server	通過 WebSocket 將指令推送至目標房間的 C-Client Agent
3	Agent	根據模式執行：NVR 模式調用 ACS Pro API 觸發 NVR 錄製；攝像頭模式視頻處理引擎拉流錄製。S-Server 同時進行 Dante 音頻採集 → DSP 處理 → 三路輸出（WAV/AAC/ASR 排程）
4	Agent	錄製狀態定時通過 HTTP 心跳上報至 S-Server
5	-	視頻錄製文件寫入 NAS（per-cam MP4），音頻錄製文件寫入 NAS（room 級 WAV + AAC）

串流流程：即時串流採用零轉碼串流引擎作為中轉。零轉碼串流引擎訂閱各路 RTSP 流，實時轉換為 WebRTC 協議輸出。前端瀏覽器通過 WebRTC 信令建立連接後，直接與零轉碼串流引擎交換媒體數據，無需經過後端服務。這種架構將媒體傳輸路徑與業務邏輯路徑分離，大幅降低服務端負載。

AI 處理流程：AI 轉寫與翻譯為離線批量處理任務。管理員在系統界面觸發轉寫任務後，任務進入排程隊列。排程服務根據伺服器負載情況分發任務至 AI 處理 Worker。Worker 加載離線語音轉寫引擎對音頻素材進行轉寫，轉

寫結果存入內建全文搜索引擎索引。翻譯任務類似，翻譯模型加載後對轉寫文本進行批量翻譯，翻譯結果與原文本建立時序關聯。

回放流程：用戶打開播放界面後，前端請求播放列表元數據。後端返回指定時間範圍內的視頻文件清單與對應的轉寫/翻譯數據。前端播放器組件同時拉取多路視頻流進行解碼渲染，音頻波形組件請求音頻波形預渲染數據。即時字幕組件訂閱播放進度事件，實時匹配並疊加字幕。整個回放過程無需後端轉碼，採用流式傳輸減少帶寬佔用。

6.3 數據存儲架構

嵌入式數據庫：系統採用嵌入式數據庫作為主要關係型數據庫，用於存儲元數據、配置數據、審計日誌等結構化數據。嵌入式數據庫的單文件特性簡化了系統部署與備份操作。數據庫集成全文搜索引擎功能，支撐全文搜索功能。

文件存儲：錄製文件按「房間/日期/錄製場次」三級目錄結構組織，便於管理員按物理位置定位素材。文件命名遵循「設備ID_起始時間_結束時間」規則，文件名即包含完整查詢信息。

NAS 統一存儲架構：系統採用群暉（Synology）NAS 作為統一存儲工作目錄，配備 2×10GbE 網口鏈路聚合（20 Gbps 總帶寬）。

NAS 工作流程細化：

階段	參與方	操作
1. 視頻錄製	NVR / C-Client Agent	NVR 錄製 MP4 直寫 NAS；Agent 直連錄製寫入 NAS
2. 音頻錄製	S-Server	Dante 音頻直接寫入 NAS
3. AI 處理	S-Server	從 NAS 讀取音頻，處理結果（轉寫/翻譯）寫回 NAS
4. 合成打包	S-Server	從 NAS 讀取音頻+MP4，合成打包後寫回 NAS

消除中間傳輸：所有參與方（NVR、C-Client Agent、S-Server）均直讀直寫 NAS，無需經過任何中間服務器轉發。NAS 以 SMB/CIFS 協議掛載至各設備，對上層應用透明如本地目錄。

6.4 高可用與故障隔離

進程隔離：各功能模組運行於獨立進程空間，單一模組崩潰不會導致整個系統不可用。崩潰的模組由看門狗進程自動重啟恢復。

降級策略：當非核心模組（如零轉碼串流引擎）不可用時，系統自動啟用降級方案。降級期間系統維持錄製、回放等核心功能，運維人員收到告警通知後介入處理。

數據完整性保障：錄製過程採用預寫式日誌（WAL）機制，確保進程異常中斷時數據不損壞。文件寫入採用原子操作，避免部分寫入造成文件損壞。

第八章　版本劃分與授權模式

7.1 版本矩陣

本系統提供單機版、專業版、企業版三個產品版本，各版本功能配置與授權範圍如下：

功能項	單機版	專業版	企業版
部署模式			
最簡模式（單機）	✅	✅	✅
NVR 模式	✅	✅	✅
攝像頭模式	❌	✅	✅
支持房間數	單房間	多房間	多房間
同步錄製路數	每房間 2 路攝像頭	每房間 6 路攝像頭	每房間 12 路攝像頭
AI 功能			
AI 語音轉寫	❌	支持	支持
AI 離線翻譯	❌	支持	支持
AI 降噪	❌	支持	支持
音頻聲道數	4 聲道	16 聲道	64 聲道
核心功能			
即時串流預覽	❌	支持	支持
全文搜索	❌	支持	支持
系統監控	基礎監控	完整監控	完整監控
審計日誌	❌	基礎日誌	完整審計合規日誌
配置版本控制	❌	支持	支持
設備管理	基礎管理	完整管理	完整管理
支援			
技術支持	郵件支持	郵件支持	優先響應支持
最大席位數	1 席位	5 席位	不限席位

7.2 適用版本建議

單機版適用於以下場景：個人用戶或小型團隊進行功能評估與學習研究；經費預算有限但對錄製功能有基礎需求的個人律師事務所；需要記錄單一會議室的中小型商務場合。單機版支援最簡模式和 NVR 模式，AI 功能可作為進階選配。

專業版適用於以下場景：多房間部署的中小型仲裁機構或法務部門；需要 AI 轉寫提升工作效率的會議記錄需求；對錄製內容有檢索需求的培訓機構或研究場所；有一定 IT 運維能力的企業內部部署。

企業版適用於以下場景：各級法院、司法機關的庭審錄製系統建設；大型仲裁中心、國際會議中心的多廳室集中管理；對數據安全與審計合規有嚴格要求的上市公司、金融機構；需要定製化開發與深度技術支持的政府事業單位。

7.3 授權模式

本系統採用「永久授權 + 年費維護」的混合授權模式：

永久授權（License Fee）：用戶一次性支付軟件授權費用，獲得軟件的永久使用權。授權費用按版本與部署規模分級定價，具體報價請聯繫銷售團隊獲取定製方案。

年費維護（Maintenance Fee）：用戶按年支付維護服務費用，享有的服務包括：軟件版本定期更新與功能升級、安全補丁推送、技術問題遠程支持、優先新功能體驗等。年費維護為可選服務，用戶可在首年優惠期後決定是否續費。選擇不續費維護的用戶可繼續使用當前版本，但不再獲得更新與支持。

7.4 後續更新規劃

本系統遵循持續迭代的產品策略，主要更新方向包括：

功能增強：規劃中的功能包括更多 AI 模型的本地化適配（如 Qwen3.6-35B 大語言模型用於智能摘要）、支持更多音頻格式與視頻編碼格式、增強的數據分析報表功能等。

平台擴展：規劃支援 Windows 操作系統部署，擴展移動端（iOS/Android）遠程管理 App，進一步完善 Docker/Kubernetes 容器化部署方案。

安全加固：定期發布安全補丁，修復已知漏洞。重大版本更新將引入更嚴格的安全機制，包括強化身份認證、敏感操作二次確認、數據加密增強等。

7.5 技術支持服務

郵件支持（全版本可用）：所有用戶均可通過郵件提交技術問題與功能建議，瀏覽使用文檔、FAQ、教程視頻。

優先支持（企業版）：企業版用戶享有優先響應支持通道，工作時間內 4 小時響應承諾。支援遠程桌面協助、優先問題排查、定制化指導。

現場支持（可選增值服務）：企業版用戶可另行選購現場技術支持服務，由工程師到場提供系統部署、培訓、故障處理等服務。具體報價與行程安排另行商議。

第九章　總結

RecordStudio v1.0 是專為司法機構、仲裁中心及大型企業會議管理場景打造的商用級多媒體錄製與內容管理平台。系統以「安全可靠、智能高效、合規可審」為核心設計理念，採用 C/S 架構實現錄製與管理的分工協作——C-Client Recording Agent（Windows 平台，Python 後台服務）作為分佈式「錄製工人」執行各房間的視頻錄製

作業，支援 NVR 模式（通過 AXIS ACS Pro API 控制 NVR）與攝像頭模式（視頻處理引擎直接拉流），支援離線錄製；S-Server 服務端（Mac mini M4 Pro）作為全系統「調度大腦」，負責 AI 推理、串流中轉、統一調度與狀態彙總。

Recording Agent 核心價值：

特性	說明
分佈式錄製	每房間部署一台 C-Client，故障隔離、負載分散
離線容忍（Offline Tolerance）	Agent 斷連後本地持續錄製，恢復後自動補報狀態
雙模式適配	NVR 模式（輕量無 GPU 需求）+ 攝像頭模式（視頻處理引擎拉流）
統一調度	通過 HTTP/WebSocket 雙通道與 S-Server 即時通信

系統全面覆蓋從多路設備同步錄製、即時串流監控、AI 語音轉寫翻譯、智能降噪處理、到全文檢索回放、系統運維監控的完整業務流程。

在技術層面，本系統採用離線優先架構，所有核心 AI 能力（離線語音轉寫引擎、離線翻譯引擎）本地部署，確保會議數據全程在機構內網流轉，真正做到「數據不出門」。零轉碼串流引擎帶來 50-200 毫秒的業界領先延遲性能。毫秒級精度的時間戳對齊機制為司法取證提供無可爭議的時間真實性保障。完整的審計日誌體系滿足司法程序對電子數據的法定取證要求。系統對中文（普通話、粵語）的原生深度支持，配合繁體中文界面與中文法律術語的專項優化，是面向大中華區及海外華人社群的核心競爭力。

在產品層面，本系統提供從單機版到企業版的三層版本矩陣，精準匹配從個人用戶到大型機構的差異化需求。永久授權加年費維護的靈活商業模式降低用戶初始投入門檻。模塊化 API 架構支撐定製化開發與第三方系統整合。

RecordStudio v1.0 致力於成為司法法務機構、企業合規部門信賴的會議記錄管理工具。我們將持續傾聽用戶聲音，不斷迭代產品功能，為更多機構的數字化轉型提供有力支撐。

如需了解更多產品詳情、索取定製方案或安排產品演示，請聯繫我們的銷售與技術團隊。

文件版本：v1.0 最後更新日期：2026 年 5 月 開發者：Bob