AI 老闆助理 + Clawdbot 資安防護設計

> 這份文件同時給老闆與技術人看：
>
> - 老闆：知道這套系統怎麼保護自己的帳號與資料
> - 技術：知道在安裝與部署時要注意哪些資安措施

---

1. 權限最小化（Least Privilege）

1. **Google** 授權只開必要的 **scopes**
    - 方案設計時要分級：
        - 行程助理 → calendar
        - 行程＋寄信 → calendar,gmail
        - 再加文件處理 → calendar,gmail,drive,docs,sheets
    - 不預設全開所有 service。每多一個 service，都要在文件裡說明用途。
2. 每個帳號的授權獨立管理
    - 為不同用途使用不同帳號：
        - 例如：aiagent...（測試）、jacky...（主帳）、step1...（實驗）
    - gog auth add 時，清楚記錄每個帳號的 services。
3. 撤銷權限要簡單
    - 在 Google 帳號 → 安全性 → 第三方存取中，可以隨時撤銷 Gog/Clawdbot 的存取。
    - 技術上也可以透過刪除本機 token 檔、重新執行 gog auth add 來重設授權。

---

2. 帳號與密碼處理

1. 永遠不要把密碼交給 **AI**
    - Gmail / KKTIX / 其他服務的密碼，不應以文字形式交給 AI。
    - 登入流程：
        - 由老闆在瀏覽器或 agent-browser 的 headed 模式中親自輸入。
        - 之後可以用 agent-browser state save auth.json 保存 cookies，AI 只用 state，而不是密碼本身。
2. 優先使用 **OAuth / token**，而不是密碼重複使用
    - Google 授權使用 OAuth client + token。
    - Slack / Notion / Trello 等使用專用 token，而非個人密碼。
3. 本機環境保護

- 開啟 FileVault 磁碟加密。
- 一定要有登入密碼與螢幕鎖。
- 不要把 credentials.json 與 token 檔同步到公開雲端硬碟（如無加密的 Dropbox、公開 GitHub）。

---

3. 憑證與敏感檔案存放

3.1 Gog 憑證與 token

- 路徑（以 macOS 為例）：
  - OAuth client：~/Library/Application Support/gogcli/credentials.json
  - token 檔：~/Library/Application Support/gogcli/ 底下由 gog 產生的檔案
- 保護方式：
  - 檔案只存在本機，受 OS 權限管理。
  - 建議定期檢查這個資料夾，確保沒有多餘的測試 token。

3.2 Clawdbot 專案中的敏感內容

- docs/ 與 可複製模板/ 主要是規則與模板，一般不放敏感資料。
- 若有存放 API key / secrets：
  - 應放在 .env 或 Gateway 環境設定中，而不是 md 檔。
  - .env 檔應加入 .gitignore，避免被推到遠端 repo。

---

4. 日誌與可追溯性（Auditability）

1. 每日 **AI** 操作 **log**
   - 檔案：logs/ai-actions-YYYY-MM-DD.log。
   - 內容：
     - 建立 / 修改了哪些檔案（路徑＋檔名）
     - 轉成了哪些 PDF
     - 傳到聊天的檔案清單
     - 新增 / 調整了哪些資料夾結構
   - 更新方式：
     - 首次建立可用 >，之後一律使用 >> 追加。
   - 老闆可以用：

     ```
     cd /Users/user/clawd
     tail -f logs/ai-actions-YYYY-MM-DD.log
     ```

     即時看到 AI 正在做什麼。
2. 每日會報 **PDF**
   - 每天收尾時，將「今日摘要＋詳細 log」組成一份報告：

        - 路徑：每日會報/YYYY-MM-DD_AI助理每日工作報告.pdf
    - 作用：
        - 給老闆看當天 AI 做了什麼
        - 也作為日後查證的依據

---

5. 對外連線與網站自動化的風險控制

1. **web_search（Brave API）**
    - 只用來做「查資料＋摘要」，不要用來查敏感個資。
    - Brave API key 不應硬寫在程式碼或 md 中，應透過環境變數 / Clawdbot config 設定。
2. **agent-browser**（網站自動化）
    - 主要用在：
        - 登入老闆指定的後台（在老闆完成首次登入 / 驗證後）
        - 自動化重複操作（抓報表、填表單）
    - 避免用在：
        - 大量對外發送表單 / 留言（容易被視為濫用）
        - 規模化爬取對方網站所有內容（可能違反服務條款）
    - 防止誤操作：
        - 訂票 / 訂位流程只做到「確認頁」，最後送出由老闆按。
        - 涉及金錢交易的操作，一律要求老闆最後人工確認。
3. **Google Search** 與防機器人
    - 已知：直接用 agent-browser 操控 Google 搜尋頁面，可能被導向防機器人頁面（"Why did this happen?"）。
    - 建議：
        - 用 web_search（API）或讓老闆在普通瀏覽器先搜尋並提供 URL。
        - Browser Automation 專注在「進入特定 URL 後的操作」。

---

6. 多老闆、多 **workspace** 的資料隔離

1. 檔案層級隔離
    - 每位老闆使用獨立資料夾（例）：
        - /Users/user/clawd/customers/bossA/
        - /Users/user/clawd/customers/bossB/
    - 各自有：
        - docs/（客製說明）
        - 可複製模板/（針對該老闆調整的規則與範例）
        - logs/（操作記錄）

2. **gog** 授權隔離
    - 不同老闆用不同 Google 帳號，避免混用同一組憑證。
    - 若同一主機要服務多位老闆，需明確記錄每個帳號對應的客戶，並且在指令中總是加上 --account。
3. **Clawdbot workspace** 隔離
    - 可以考慮為不同老闆啟用不同的 session / 標籤，避免訊息混在一起。
    - 在 AI 端要避免在錯誤的對話裡引用不該引用的檔案或 log。

---

7. 模型授權（OAuth）更新策略

1. **Access token** 自動刷新
    - 多數模型（例如 google-antigravity）在取得 OAuth 授權後，會持有：access token（短效）＋ refresh token（長效）。
    - 在 access token 過期前後，底層 SDK 或 Clawdbot 會自動使用 refresh token 取得新的 access token，無須老闆介入。
2. 整體授權 / **profile** 可能需要重新登入
    - 當 doctor 顯示類似：google-antigravity:xxx@gmail.com: expiring (XXm) 時，代表整體授權（或 refresh token）接近失效。
    - 出於安全與 Google 政策考量，這一類「重新授權」通常需要老闆在瀏覽器中親自點選同意，不應完全自動化。
3. 建議作法
    - 由 AI 週期性檢查 doctor 狀態，發現模型授權即將到期時：
        - 在聊天裡主動提醒老闆：「模型 X 的授權快到期了，我可以告訴你怎麼三步完成重新登入。」
        - 提供簡短步驟（例如：clawdbot configure → 選 provider → 瀏覽器點同意）。
    - 不應在老闆完全不知情的情況下，持續自動重新取得長期授權。

8. 商品化時可以寫進合約 / 說明書的保障點

1. 資料所有權
    - 所有 Gmail / Calendar / Drive / Notion / Slack 的資料，仍屬老闆所有。
    - AI 助理只在授權範圍內讀寫，不會另存一份到供應商自己的伺服器。
2. 可撤銷性
    - 老闆可以隨時撤銷 OAuth 授權，或要求刪除本機 token / 設定。
    - 可提供簡單步驟（例如 Google 第三方存取頁面連結）。
3. 透明度
    - 透過每日 log + 每日會報，老闆可以具體看到 AI 做了什麼。
    - 如有需要，可提供特定日期的 log 作為檢查依據。
4. 風險範圍說明

- 明確寫出這套系統會 / 不會做的事情，例如：
    - 會：在指定範圍內幫忙排會、回信、整理資料、填寫表單前半段。
    - 不會：擅自代表老闆做金錢交易的最後確認、發送不經老闆允許的訊息。

這份檔案建議也放一份在 可複製模板/security/AI 老闆助理資安防護說明.md，
讓任何新的部署都能從相同的資安標準出發。