

# PowerAutomation Manus 功能測試案例規格書

作者: Manus AI

版本: 1.0

日期: 2025年6月23日

測試目標: 驗證PowerAutomation系統中Manus自動化功能的完整性和可靠性

## 測試概述

本測試規格書詳細定義了PowerAutomation系統中Manus自動化功能的完整測試案例。測試將使用錄製MCP（Model Context Protocol）和Replay分析功能來記錄、分析和評估每個功能模組的性能表現。測試範圍涵蓋Manus平台的核心交互功能，包括信息發送、對話歷史管理、任務列表操作和文件檔案處理等關鍵業務流程。

測試將採用自動化瀏覽器控制技術，結合Playwright框架實現精確的頁面操作模擬。每個測試案例都將被完整記錄，包括操作步驟、響應時間、成功率和錯誤處理機制的驗證。通過MCP錄製功能，我們能夠捕獲完整的交互序列，並使用Replay分析功能對測試結果進行深度分析，識別性能瓶頸和優化機會。

## 測試目標與範圍

### 主要測試目標

本次測試的主要目標是全面驗證PowerAutomation系統與Manus平台的集成效果，確保所有自動化功能能夠穩定、準確地執行預期操作。測試將重點關注系統的可靠性、性能表現和用戶體驗質量。通過系統性的測試執行，我們期望達成以下具體目標：

首先，驗證Manus登錄和身份驗證機制的穩定性。系統必須能夠使用提供的憑證（chuang.hsiaoyen@gmail.com / silentfleet#1234）成功登錄Manus平台，並維持穩定的會話狀態。登錄過程應該在合理的時間範圍內完成，並能夠處理各種網絡條件和頁面載入狀況。

其次，測試信息發送功能的準確性和效率。系統應該能夠準確識別輸入框元素，正確輸入測試信息，並成功提交。發送過程中的每個步驟都需要被精確記錄，包括元素定位時間、輸入響應時間和提交確認狀態。

第三，驗證對話歷史獲取和分類功能的完整性。系統需要能夠遍歷完整的對話記錄，包括通過滾動載入的歷史消息，並能夠根據預定義的分類標準對對話內容進行智能分類。分類結果應該準確反映對話的主題、重要性和處理狀態。

## 測試範圍界定

測試範圍涵蓋PowerAutomation系統中所有與Manus平台交互的核心功能模組。具體包括瀏覽器自動化控制、頁面元素識別與操作、數據提取與處理、以及智能分析與分類等關鍵組件。

在功能層面，測試將覆蓋用戶界面交互的各個方面，從基本的頁面導航到複雜的數據操作流程。每個功能模組都將接受獨立測試和集成測試，確保單個組件的正確性以及組件間協作的有效性。

在技術層面，測試將驗證Playwright瀏覽器控制框架的穩定性、CSS選擇器的準確性、JavaScript執行的可靠性，以及異步操作的處理能力。特別關注在不同網絡條件和頁面載入狀態下的系統表現。

在數據層面，測試將驗證數據提取的完整性、分類算法的準確性、以及數據存儲和檢索的效率。確保系統能夠處理各種格式和大小的數據，並維持數據的一致性和完整性。



## 測試環境配置

---

### 硬體與軟體環境

測試環境基於Ubuntu 22.04 LTS操作系統，運行在具備充足計算資源的虛擬化環境中。系統配置包括多核心處理器、充足的內存空間和高速網絡連接，確保測試過程中不會因為資源限制而影響結果的準確性。

軟體環境包括Python 3.11運行時、Playwright瀏覽器自動化框架、Flask Web應用框架，以及相關的依賴庫。所有軟體組件都已更新到最新穩定版本，確保測試環境的現代性和兼容性。

瀏覽器環境使用Chromium引擎，配置為無頭模式以提高執行效率，同時保留完整的JavaScript執行能力和DOM操作支持。瀏覽器設置包括適當的超時配置、用戶代理字符串和

視窗大小設定。

## 網絡與安全配置

測試環境配置了穩定的網絡連接，支持HTTPS協議和現代Web標準。網絡配置包括適當的代理設置、DNS解析配置和防火牆規則，確保能夠正常訪問Manus平台的所有功能。

安全配置方面，測試環境實施了適當的訪問控制和數據保護措施。測試用的登錄憑證將被安全存儲和使用，所有敏感數據都將在測試完成後被適當清理。

為了確保測試結果的可重現性，環境配置將被完整記錄和版本控制。任何配置變更都將被追蹤，確保測試環境的一致性和穩定性。



## 詳細測試案例設計

---

### TC001: Manus 登錄驗證測試

**測試目標:** 驗證PowerAutomation系統能夠成功登錄Manus平台並維持穩定會話

**前置條件:** - PowerAutomation系統已正確安裝和配置 - 網絡連接正常，能夠訪問 <https://manus.chat> - 測試憑證有效 (chuang.hsiaoyen@gmail.com / silentfleet#1234)

**測試步驟:** 1. 啟動PowerAutomation Manus瀏覽器控制器 2. 導航到Manus登錄頁面 3. 識別並填寫電子郵件輸入框 4. 識別並填寫密碼輸入框 5. 點擊登錄按鈕 6. 等待頁面重定向並驗證登錄成功 7. 檢查會話狀態和用戶界面元素

**預期結果:** - 成功導航到登錄頁面 (響應時間 < 5秒) - 正確識別和填寫登錄表單元素 - 登錄過程順利完成 (總時間 < 15秒) - 成功重定向到主界面 - 會話狀態保持穩定

**成功標準:** - 登錄成功率  $\geq 95\%$  - 平均登錄時間  $\leq 10$ 秒 - 會話維持時間  $\geq 30$ 分鐘 - 錯誤處理機制正常運作





**MCP錄製要點:** - 記錄完整的登錄流程操作序列 - 捕獲頁面載入時間和響應延遲 - 記錄任何錯誤或異常情況 - 分析登錄成功的關鍵因素

### TC002: 信息發送功能測試

**測試目標:** 驗證系統能夠準確識別輸入框並成功發送測試信息

**前置條件:** - 已成功登錄Manus平台 - 位於可以發送信息的對話界面 - 輸入框元素可見且可操作

**測試數據準備:** 測試將使用多種類型的信息內容來驗證系統的處理能力:

1. **基本文本信息:** " PowerAutomation 測試信息 - 基本文本發送測試"
2. **包含特殊字符的信息:** "測試特殊字符: @#\$%^&\*()\_+=[{}]|;'\",./<>?"
3. **多行文本信息:** 包含換行符和段落結構的長文本
4. **包含Emoji的信息:** "   PowerAutomation 功能測試    "
5. **技術內容信息:** 包含代碼片段和技術術語的信息

**測試步驟:** 1. 識別當前頁面的信息輸入框元素 2. 清空輸入框中的現有內容 3. 輸入第一條測試信息（基本文本） 4. 驗證輸入內容的正確性 5. 點擊發送按鈕或按下Enter鍵 6. 等待信息發送完成並確認 7. 重複步驟2-6，測試其他類型的信息 8. 驗證所有發送的信息都正確顯示在對話中

**預期結果:** - 準確識別輸入框元素（識別時間 < 2秒） - 正確輸入各種類型的測試信息 - 成功發送所有測試信息（發送成功率 100%） - 發送的信息在對話中正確顯示 - 特殊字符和Emoji正確處理和顯示

**性能指標:** - 單條信息發送時間 ≤ 3秒 - 輸入框識別準確率 ≥ 98% - 信息內容完整性 100% - 特殊字符處理準確率 ≥ 95%

**錯誤處理驗證:** - 網絡中斷時的重試機制 - 輸入框不可用時的處理 - 信息發送失敗時的錯誤報告 - 超長信息的截斷或分段處理

## TC003: 對話歷史獲取測試

**測試目標:** 驗證系統能夠完整獲取對話歷史，包括需要滾動載入的舊消息

**前置條件:** - 已成功登錄並位於包含歷史對話的界面 - 對話中存在足夠的歷史消息（建議 > 50條） - 頁面滾動功能正常

**測試策略:** 對話歷史獲取測試將採用漸進式滾動策略，從最新的消息開始，逐步向上滾動載入更早的歷史記錄。系統需要能夠識別何時到達對話的開始，並確保沒有遺漏任何消息。

**測試步驟:** 1. 記錄當前可見的最新消息作為起始點 2. 執行向上滾動操作，觸發歷史消息載入 3. 等待新消息載入完成（監控DOM變化） 4. 提取新載入的消息內容和元數據 5. 重複步驟2-4，直到無法載入更多歷史消息 6. 驗證消息的時間順序和完整性 7. 統計總共獲取的消息數量 8. 檢查是否存在重複或遺漏的消息

**數據提取要求:** 對於每條獲取的消息，系統需要提取以下信息： - 消息內容（文本、圖片、文件等） - 發送者信息（用戶名、角色） - 時間戳（精確到秒） - 消息類型（文本、系統通知、文件分享等） - 消息狀態（已讀、未讀、編輯等） - 相關附件或媒體文件信息

**預期結果:** - 成功載入所有可用的歷史消息 - 消息順序正確（按時間排序） - 沒有重複或遺漏的消息 - 所有消息元數據完整準確 - 滾動操作流暢，無卡頓或錯誤

**性能要求:** - 單次滾動響應時間  $\leq 2$  秒 - 消息載入完成時間  $\leq 5$  秒 - 總體獲取效率  $\geq 10$  條消息/秒 - 內存使用量保持在合理範圍內

## TC004: 對話內容智能分類測試

**測試目標:** 驗證系統能夠對獲取的對話內容進行準確的智能分類

**分類標準定義:** 系統將根據以下標準對對話內容進行多維度分類：

**按內容類型分類:** - 技術討論：包含代碼、技術術語、問題解決等內容 - 項目管理：涉及任務分配、進度報告、計劃討論等 - 一般交流：日常溝通、問候、非正式討論等 - 文件分享：包含文件上傳、鏈接分享、資源共享等 - 系統通知：自動生成的狀態更新、提醒等

**按重要性分類:** - 高優先級：緊急問題、重要決策、關鍵信息等 - 中優先級：常規工作討論、一般性問題等 - 低優先級：閒聊、非關鍵信息等

**按處理狀態分類:** - 需要回應：包含問題或請求的消息 - 僅供參考：信息性消息，無需特定行動 - 已處理：已經得到回應或解決的問題

**測試步驟:** 1. 使用TC003獲取的對話歷史數據作為輸入 2. 對每條消息應用智能分類算法 3. 生成分類結果和置信度分數 4. 驗證分類結果的準確性 5. 統計各類別的消息分布 6. 識別分類困難或模糊的消息 7. 生成分類報告和建議

**預期結果:** - 分類準確率  $\geq 85\%$  - 高置信度分類 ( $>0.8$ ) 比例  $\geq 70\%$  - 分類處理速度  $\geq 100$  條消息/秒 - 分類結果一致性和可重現性

## TC005: 任務列表遍歷測試

**測試目標:** 驗證系統能夠遍歷左側任務列表，類似TRAE倉庫的組織方式

**前置條件:** - 已成功登錄Manus平台 - 左側任務列表可見且包含多個任務項目 - 任務列表支持點擊和導航操作

**測試步驟:** 1. 識別左側任務列表容器元素 2. 獲取所有可見的任務項目 3. 提取每個任務的基本信息（標題、狀態、時間等） 4. 依次點擊每個任務項目 5. 驗證任務切換是否成功 6. 記錄每個任務的詳細信息 7. 檢查任務之間的關聯關係 8. 驗證任務列表的完整性

**數據提取要求:** - 任務標題和描述 - 任務狀態（進行中、已完成、待處理等） - 創建時間和最後更新時間 - 參與者信息 - 任務優先級 - 相關文件和附件數量

**預期結果:** - 成功識別和遍歷所有任務項目 - 任務信息提取完整準確 - 任務切換操作流暢無誤 - 任務列表結構清晰有序

## TC006: 任務文件檔案獲取測試

**測試目標:** 驗證系統能夠獲取特定任務中的所有文件檔案，包括分類和批量下載功能

**前置條件:** - 已成功遍歷任務列表並選定目標任務 - 目標任務包含多種類型的文件 - 文件分類和下載功能可用

**文件分類測試:** 根據之前分析的Manus界面，系統需要測試以下四種文件分類：

1. **Documents（文檔）**：PDF、Word、Excel、PowerPoint等辦公文檔
2. **Images（圖片）**：JPG、PNG、GIF、SVG等圖像文件
3. **Code files（代碼文件）**：Python、JavaScript、HTML、CSS等代碼文件
4. **Links（鏈接）**：外部鏈接、參考資料、網頁書籤等

**測試步驟:** 1. 進入目標任務的詳細頁面 2. 定位"View all files in this task"按鈕 3. 點擊按鈕打開文件列表界面 4. 識別文件分類標籤（紅色圈圈標示的分類） 5. 依次點擊每個分類標籤 6. 記錄每個分類下的文件列表 7. 測試批量下載功能（綠色圈圈標示的下載按鈕） 8. 驗證下載的文件完整性和正確性

**預期結果:** - 成功打開文件列表界面 - 正確識別和操作分類標籤 - 準確獲取每個分類的文件列表 - 批量下載功能正常工作 - 下載的文件完整且可用

**性能指標:** - 文件列表載入時間  $\leq 5$ 秒 - 分類切換響應時間  $\leq 2$ 秒 - 文件下載速度  $\geq 1$ MB/秒 - 下載成功率  $\geq 98\%$

# MCP錄製和Replay分析規格

## MCP錄製配置

**錄製範圍:** 每個測試案例都將使用MCP（Model Context Protocol）進行完整錄製，捕獲以下關鍵信息：

**操作層面錄製:** - 瀏覽器操作序列（點擊、輸入、滾動、導航等） - 頁面元素識別過程（CSS選擇器、XPath、元素屬性） - 時間戳記錄（操作開始時間、完成時間、響應延遲） - 錯誤和異常處理過程

**數據層面錄製:** - 提取的數據內容和結構 - 數據處理和轉換過程 - 分類算法的決策過程 - 結果驗證和質量檢查

**性能層面錄製:** - 內存使用情況 - CPU使用率 - 網絡請求和響應時間 - 頁面載入和渲染時間

## Replay分析標準

**成功率分析:** - 操作成功率統計 - 失敗原因分類和分析 - 重試機制效果評估 - 穩定性指標計算

**性能分析:** - 響應時間分布分析 - 性能瓶頸識別 - 資源使用效率評估 - 優化建議生成

**質量分析:** - 數據準確性驗證 - 分類結果質量評估 - 用戶體驗指標測量 - 系統可靠性評估

## 測試數據管理

### 測試數據準備

**真實數據使用:** 測試將使用真實的Manus平台數據，包括實際的對話記錄、任務列表和文件檔案。這確保測試結果能夠真實反映系統在生產環境中的表現。

**數據隱私保護:** 所有測試過程中獲取的敏感數據都將被適當處理和保護。個人信息將被匿名化或脫敏處理，確保符合數據保護要求。

**測試數據分類:** - 基準數據：用於驗證基本功能的標準數據集 - 邊界數據：用於測試系統極限和邊界條件的數據 - 異常數據：用於驗證錯誤處理機制的異常情況數據 - 性能數據：用於性能測試的大量數據集

## 結果數據存儲

**錄製數據存儲:** 所有MCP錄製數據將被結構化存儲，包括操作序列、時間戳、結果數據等。存儲格式將支持後續的分析和重播。

**分析結果存儲:** Replay分析的結果將被保存為結構化報告，包括統計數據、圖表、建議等。結果格式將支持多種輸出方式，包括PDF報告、JSON數據和視覺化圖表。

## 測試執行策略

---

### 測試順序安排

測試將按照邏輯依賴關係有序執行：

- 基礎功能測試:** 首先執行登錄驗證，確保後續測試的前置條件
- 核心功能測試:** 依次測試信息發送、對話歷史獲取等核心功能
- 高級功能測試:** 測試分類、任務管理等高級功能
- 集成測試:** 驗證各功能模組之間的協作效果
- 性能測試:** 在功能驗證完成後進行性能評估

### 錯誤處理策略

**自動重試機制:** 對於網絡相關的臨時性錯誤，系統將自動重試最多3次，每次重試間隔遞增。

**錯誤分類處理:** 不同類型的錯誤將採用不同的處理策略： - 網絡錯誤：重試機制 - 元素識別錯誤：嘗試替代選擇器 - 權限錯誤：記錄並跳過相關測試 - 系統錯誤：立即停止並報告

**測試中斷條件:** 以下情況將導致測試中斷： - 無法連接到Manus平台 - 登錄憑證失效或被拒絕 - 系統資源不足（內存、磁盤空間等） - 關鍵依賴服務不可用

## 成功標準定義

---

### 功能性成功標準

**基本功能要求:** - 登錄成功率  $\geq 95\%$  - 信息發送成功率  $\geq 98\%$  - 數據獲取完整性  $\geq 95\%$  - 分類準確率  $\geq 85\%$



**性能要求:** - 平均響應時間  $\leq 5$ 秒 - 頁面載入時間  $\leq 10$ 秒 - 數據處理速度  $\geq 100$ 條記錄/秒 - 系統資源使用率  $\leq 80\%$

## 質量標準

**可靠性要求:** - 系統穩定運行時間  $\geq 2$ 小時 - 錯誤恢復成功率  $\geq 90\%$  - 數據一致性 100% - 重複測試結果一致性  $\geq 95\%$

**用戶體驗要求:** - 操作流暢性評分  $\geq 4.0/5.0$  - 錯誤信息清晰度評分  $\geq 4.0/5.0$  - 整體滿意度評分  $\geq 4.0/5.0$

## 測試完成標準

---

測試將在滿足以下條件時被認為完成：

- 所有測試案例執行完畢:** 六個主要測試案例都已執行並記錄結果
- MCP錄製數據完整:** 所有操作過程都被完整錄製並可重播
- Replay分析完成:** 所有錄製數據都經過分析並生成報告
- 成功標準達成:** 關鍵性能指標達到預定標準
- 文檔和錄屏完成:** 完整的測試報告和錄屏文件已生成

## 風險評估與應對

---

### 技術風險

**網絡連接風險:** Manus平台可能因為網絡問題無法訪問 - 應對措施：準備備用網絡連接，實施重試機制

**瀏覽器兼容性風險:** 頁面結構變化可能導致元素識別失敗 - 應對措施：使用多種元素識別策略，準備備用選擇器

**性能風險:** 大量數據處理可能導致系統資源不足 - 應對措施：實施資源監控，必要時分批處理數據

## 業務風險

**數據隱私風險:** 測試過程可能涉及敏感數據 - 應對措施：實施數據脫敏，嚴格控制數據訪問

**服務中斷風險:** Manus平台維護可能影響測試 - 應對措施：選擇合適的測試時間窗口，準備應急計劃

本測試規格書為PowerAutomation系統的Manus功能測試提供了全面的指導框架。通過系統性的測試執行和分析，我們將能夠全面評估系統的功能性、性能和可靠性，為後續的優化和部署提供重要依據。