COUNTRY GARDEN SERVICES HOLDINGS COMPANY LIMITED

碧桂園服務控股有限公司

(於開曼群島註冊成立之有限公司)

(股份代號：6098.HK)

人工智能與自動化決策管理政策

1. 總則

1.1 碧桂園服務控股有限公司（簡稱「碧桂園服務」或「集團」）為規範集團人工智能、自動化決策工具全生命週期管理，落實數據隱私保護、網絡安全防護、算法公平、人工複核、可解釋性、權責劃分、能力邊界管控、降低生態足跡、禁用違規人工智能系統九大核心要求，防範算法偏見、隱私洩露、網絡入侵、決策失控、違規監控、操縱行為、算力資源過度消耗等風險，依據《網絡安全法》《數據安全法》《個人信息保護法》《生成式人工智能服務管理暫行辦法》《人臉識別技術應用安全管理規定》、網絡安全等級保護標準、上市公司內控及《歐盟人工智能法案》合規原則制定本辦法，保障業主、員工合法權益及集團運營安全。

1.2 集團最高人工智能審批決策機構為集團執行管理層，涉及高風險、重大新型人工智能系統上線須經執行管理層審議批准。

2. 適用範圍

集團及所有下屬子公司、各地項目分支機構的全部運營活動下的所有用工均須遵從本辦法；覆蓋自研、外購全部人工智能模型與自動化決策系統，涵蓋需求設計、開發訓練、上線運行、迭代更新、下線銷燬全流程。嚴禁部署、接入、使用符合《歐盟人工智能法案》禁止類型的人工智能系統。

3. 術語和定義

3.1 **人工智能系統：**依託機器學習、計算機視覺、大模型等技術自動完成數據識別、分析、判定的軟硬件，包含自研人工智能、第三方人工智能。

3.2 **自動化決策：**人工智能自主輸出業務結論並自動執行操作，如業主準入、費用催收、人臉通行、設備派單等。

3.3 **人在迴路（human-in-the-loop）：**高風險人工智能流程設置人工複核節點，人工智能僅提供參考結論，關鍵操作須經人工確認執行。

3.4 **算法偏見：**模型因訓練數據、設計缺陷，對不同業主、群體產生差異化、不公平判定結果。

3.5 **人工智能可解釋性：**人工智能決策具備完整溯源依據，可清晰說明判定邏輯。

3.6 **人工智能能力邊界：**明確人工智能允許處理的業務與數據範圍，劃定禁止操作事項。

3.7 **生態足跡：**人工智能模型及配套算力數據中心全生命週期綜合環境佔用，涵蓋耗電、耗水、硬件資源、碳排放、電子廢棄物等環境影響，應儘量降低整體環境負荷。

3.8 **違規人工智能系統：**指從事操縱行為、利用系統漏洞、實施社會評分、開展未經授權生物識別監控，以及《歐盟人工智能法案》明令禁止的人工智能系統。

3.9 **人工智能安全事件：**包含隱私洩露、網絡入侵、算法歧視、決策失控、超範圍使用人工智能模型、高生態足跡運行等情形。

4. 核心管理原則

4.1 **隱私優先準則**：遵循最小必要原則採集數據，嚴格保護個人及生物敏感信息，杜絕無授權、超範圍使用數據。

4.2 **網絡安全準則**：人工智能模型、數據存儲、調用鏈路全程做網絡防護，封堵入侵、數據竊取通道，嚴控內外網數據交互風險。

4.3 **公平無偏見準則**：校驗訓練數據與算法，消除歧視性偏差，保障算法判定公平公正。

4.4 **人在迴路準則**：高風險業務保留完整人工干預權限，杜絕人工智能全權自主執行關鍵業務。

4.5 **透明可解釋準則**：留存人工智能決策推理鏈路，保證判定可追溯、可解釋，禁止黑盒判定。

4.6 **透明可識別準則**：人工智能生成內容及決策結論須顯式與隱式標識，確保用戶識別人工智能屬性，交互服務首次互動須告知。

4.7 **權責清晰準則**：遵循誰使用、誰負責，誰開發、誰擔責，明確人工智能產出結果責任主體。

4.8 **邊界管控準則**：上線前明確人工智能能力邊界，嚴禁超出既定業務範圍運行模型。

4.9 **綠色低碳準則**：選用輕量化模型、綠色算力架構，優化算力調度，降低人工智能整體生態足跡，減少環境負擔。

4.10 **合規禁限準則**：嚴禁部署、啟用操縱行為、漏洞利用、社會評分、非授權生物識別監控及其他受法律禁止的人工智能系統。

5. 職責與分工

5.1 **業務使用部門**：統籌人工智能業務需求，明確人工智能能力邊界，開展合規培訓與人工智能結果複核，受理業主異議並完成人工重審，及時上報安全隱患、異常高能耗運行問題，承擔本場景人工智能決策的主要業務責任。

5.2 **數字化能力中心（人工智能開發實施團隊）**：負責自研人工智能技術管控，落實數據脫敏與隱私保護、網絡防護、算法糾偏、決策溯源，選用輕量化模型降低生態足跡，開展合規核查，杜絕違規人工智能模型，完成上線及迭代風險評估、全流程審計日誌留存、下線數據銷燬，承擔模型技術相關責任。重大高風險人工智能上線方案須報送集團執行管理層審批。

5.3 **信息安全團隊**：統籌人工智能合規管控與外網模型訪問管控，定期開展安全、算法及算力能耗審計，核查生態足跡管控情況，處置人工智能安全事件，配合等保及第三方審計工作。

6. 人工智能全流程管控要求

6.1 **上線準入與邊界管控**：人工智能上線前完成整體合規風險評估並留存記錄，明確能力邊界及算力能耗管控方案，核查是否屬於違規人工智能系統；人臉、財務等高敏感場景及重大新型人工智能系統須經集團執行管理層審議批准，不得越界執行審批、扣款、權限開通等高風險操作，嚴控整體生態足跡。嚴禁部署操縱行為、漏洞利用、社會評分、未經授權生物識別監控及合規禁限類人工智能系統。

6.2 **數據隱私保護**：人臉等敏感信息須取得單獨書面授權、本地內網隔離存儲、最短週期留存；人工智能訓練數據必須脫敏；嚴禁將敏感數據傳入外部人工智能，內部數據導出執行審批，高敏感數據原則禁止落地導出。

6.3 **網絡安全防護**：人工智能數據庫及模型部署隔離安全區域，配置防火牆、訪問管控、接口鑑權，定期漏洞巡檢修復；僅允許通過內網合規渠道使用人工智能，禁止私傳敏感數據至外部模型。

6.4 **算法偏見防控**：均衡訓練樣本、剔除歧視特徵，定期開展算法公平性檢測，及時修正存在偏見的模型，杜絕差異化不公判定。

6.5 **人在迴路管控：**租戶準入、大額費用調整、批量催收、人臉權限開通等高風險場景，人工智能僅提供參考意見，強制人工複核確認後執行；系統支持隨時人工駁回、終止人工智能自動流程；低風險場景定期抽樣覆盤。

6.6 **可解釋性管理：**人工智能決策留存完整溯源日誌，可向業主、審計說明判定依據；對外服務採用易懂話術做好解釋。

6.7 **運維與下線管理：**定期開展漏洞、算法偏差及算力能耗巡檢；重大版本迭代重新完成合規風險評估；持續優化模型架構以降低生態足跡；下線時徹底銷燬訓練數據與模型；第三方模型到期完成數據刪除確認；敏感人工智能輸出文件添加水印，嚴控傳輸渠道。

7. 人工智能安全事件處置與審計

7.1 人工智能安全事件分級處置

事件等級	典型情形	處置時限	處置要求
一級重大事件	人臉批量洩露、大額損失、模型被盜、大規模算法歧視、長期超高生態足跡運行	1 小時內	關停涉事人工智能、上報管理層、溯源排查、告知受影響業主、留存審計日誌並依規報備
二級一般事件	局部算法偏見、不合規外部人工智能接入、第三方模型生態足跡超標	24 小時內	全面整改、限制相關人員權限、開展合規培訓、優化算力調度
三級輕微事件	輕微越界使用、無敏感信息違規調用外部人工智能、短期能耗異常	當天	完成整改、內部通報、完善邊界管控及輕量化優化方案

7.2 **常態化審計：**信息安全團隊定期開展人工智能專項審計，核查隱私保護、網絡安全、算法公平、人在迴路落地、可解釋性、能力邊界、生態足跡管控執行情況，形成整改閉環；年度等保及第三方審計留存全套評估、日誌及整改資料。

7.3 人工智能全流程審計日誌留存不少於 3 年，滿足合規及追溯要求。

8. 政策審視與修訂

本政策由集團董事會審議批准，並由集團執行管理層及信息安全團隊監督其有效實施，適時對本政策的適用性及有效性進行審視及檢討，並結合檢討結果、法律法規更新、行業事件發展及公司實際運營情況評估考量修訂本政策，以促進公司持续提升可持续發展水平。