Google Gemini 3.5 Flash「Computer Use」機能統合がもたらす知財業務の革新と実務的影響

Gemini 3.1 pro

1. 序論：Agentic AIによるGUI自動化の幕開けと知財業務の転換点

2026年6月24日、Googleは同社の主力生成AIモデルである「Gemini 3.5 Flash」のアップデートを発表し、PC画面の視覚認識と自律的な操作を可能にする「Computer Use（コンピューター使用）」機能をモデル内部にネイティブ統合したことを明らかにした[1]。この技術的ブレイクスルーは、AIモデルがソフトウェアと連携する手段を根底から覆すものである。これまで、AIが外部システムと対話する主要な手段は、構造化されたAPI（Application Programming Interface）を経由する関数呼び出し（Function Calling）に限定されていた。しかし、現実のビジネス環境、とりわけ特許事務所や企業の知的財産（知財）部門においては、APIが提供されていないレガシーな特許管理システム、クライアントごとに異なる外部の専用ポータルサイト、あるいは各国の特許庁が提供する複雑な電子出願システムなど、グラフィカル・ユーザー・インターフェース（GUI）を介した直接的な操作が不可欠な業務が山積している[1]。

Gemini 3.5 FlashにおけるComputer Use機能の標準搭載は、AIエージェントが人間と全く同じように「画面の視覚情報を認識し（See）、状況を推論し（Reason）、適切な操作を実行する（Act）」という一連の認知・行動プロセスを単一のモデルで実現したことを意味する[1]。ブラウザ、モバイル、そしてデスクトップ環境を横断し、クリック、タイピング、スクロールといったアクションを自然言語の指示のみで実行できるこの機能は、従来のRPA（Robotic Process Automation）が長年抱えていた「UI変更に対する脆弱性」や「例外処理能力の欠如」といった構造的限界を克服するものである。結果として、ソフトウェアテストや知識労働といった、長期的かつ複雑な自動化タスク（Long-horizon automation tasks）にパラダイムシフトをもたらす可能性を秘めている[1]。

本報告書は、Gemini 3.5 FlashのComputer Use機能の技術的仕様とそのアーキテクチャ上の優位性を徹底的に解剖し、それが特許管理システムの操作、先行技術文献の取得、明細書のチェックといった知財特有のワークフローにどのような根本的変革をもたらすかを詳解する。さらに、日本弁理士会（JPAA）が策定したAI利活用ガイドラインや営業秘密管理の観点から、未公開の特許情報等をAIエージェントに処理させる際のデータプライバシー（Zero Data Retention等）および法的・倫理的リスクについて、包括的かつ実務的な分析を提供する。

2. 知財業務における従来の自動化（RPA）の構造的限界

Gemini 3.5 Flashの統合機能がもたらす価値を正確に評価するためには、まず特許事務所や企業知財部においてこれまで推進されてきた自動化の取り組みと、その限界を理解する必要がある。知財業務のデジタルトランスフォーメーション（DX）において、過去数年間にわたり中心的な役割を果たしてきたのはRPAである。

2.1 RPA導入の背景と一時的成功

日本弁理士会の機関誌等でも報告されている通り、特許実務は高度な専門的判断を要する一方で、膨大な定型的・反復的事務作業を内包している[7]。RPAは、人間が行うキーボードやマウスの操作をソフトウェアロボットが模倣する技術であり、カスタマイズに多大なコストがかかる既存の特許管理システムや所内ネットワークの環境に一切の変更を加えることなく導入できるという利点があった[7]。

具体的な適用例として、クライアントシステムからの案件データのダウンロード、新規受任データの特許管理システムへの入力、審査請求や年金納付期限のリマインダの一括作成、特許庁へ提出したPDFファイル群の指定フォルダ（電子包袋）への自動振り分け保存などが挙げられる[7]。これらは、データ登録ミスなどのヒューマンエラーを防ぐと同時に、事務担当者の作業負担を劇的に軽減する成果を上げてきた。さらに、出願人の整理番号体系変更に伴う一括変更作業など、マクロ（Excel VBA等）では対応できない複数アプリケーションを跨ぐ処理においてもRPAは有効性を発揮した[7]。

2.2 従来型RPAの技術的制約と保守コストの増大

しかし、実務の現場において、RPAの運用は常に「脆さ（Fragility）」との戦いであった。RPAの抱える根本的な課題は、以下の3点に集約される。

第一に、UI変更に対する極端な脆弱性である。従来型RPAの多くは、画面上の「固定された絶対座標（ピクセル位置）」のクリックや、WebページのDOM（Document Object Model）構造における特定の要素ID、あるいは単純な画像認識（テンプレートマッチング）に依存して動作を規定している[7]。そのため、特許管理システムのバージョンアップによるボタン位置の数ピクセルのズレ、クライアントポータルのウェブデザインの微細な変更、さらにはOSのアップデートによるフォントレンダリングの変化が生じただけで、ロボットは操作対象を見失いシステムエラーで停止してしまう[7]。結果として、システム環境が変化するたびにRPAプログラムの保守・修正作業（メンテナンス）が頻発し、自動化によって得られたコスト削減効果が相殺される事態が常態化していた[7]。

第二に、例外的な挙動に対する処理能力（エラーリカバリ能力）の欠如である。RPAは事前にプログラミングされた直線的なシナリオを盲目的に実行するに過ぎない。システムの一時的な遅延による画面遷移の遅れ、予期せぬポップアップ通知、あるいはエラーメッセージの出現といった「想定外の事象」が発生した場合、ロボットは状況を解釈して柔軟に回避策を講じることができない[7]。最悪の場合、エラー画面が表示されているにもかかわらず、バックグラウンドで間違ったテキストボックスに機密データを入力し続けるといった、クリティカルな情報セキュリティインシデントを引き起こすリスクすら内包している[7]。

第三に、非定型・非構造化データに対する認知・推論能力の欠如である。特許明細書のような長大かつ複雑な非構造化テキストや、フォーマットが統一されていない各国の特許庁からの通知書などから、意味的な意図を汲み取り、適切な判断を下してシステムに入力・転記するような業務は、ルールベースのRPA単体では到底実現不可能であった[7]。

3. Gemini 3.5 Flash「Computer Use」の技術的解剖とアーキテクチャ

GoogleがGemini 3.5 FlashのComputer Use機能は、上記のような従来型RPAの構造的限界を根本から解決する技術的飛躍である。その核心は、AIモデルのアーキテクチャの大幅な刷新と、高度な視覚的推論アルゴリズムの統合にある。

3.1 ネイティブ統合による「モデルホップ」の排除

これまでのAIによる画面操作の試みは、主にマルチモデル構成（あるいはパイプライン構成）に依存していた。すなわち、画面のスクリーンショットを解析するための「視覚認識用の特化モデル」（例えば、先行して試験提供されていたスタンドアロン型のGemini 2.5 Computer Useモデルなど）と、その解析結果を受け取って次に何をすべきかを判断する「推論・言語生成用のLLM」を組み合わせて稼働させていた[1]。
しかし、このパイプラインアーキテクチャには致命的な欠陥が存在した。複数の特化モデル間でコンテキストを受け渡すたびに（いわゆるモデルホップ）、微細な情報の欠落や翻訳エラーの蓄積が発生し、さらにネットワークのレイテンシがスタックすることで、リアルタイムな画面操作においては使い物にならないほどの遅延が生じていたのである[2]。
Gemini 3.5 Flashにおける最大の革新性は、このアーキテクチャを単一モデルに統合した点にある。Computer Use機能が、関数呼び出し（Function Calling）、Google検索によるリアルタイムな事実確認（Search Grounding）、およびGoogle Mapsとの連携機能と並んで、Gemini 3.5 Flashモデルの内部に「ネイティブな組み込みツール」として直接統合された[1]。これにより、AIエージェントは単一の推論パス（Inference pass）の中で、「画面のピクセル情報を瞬時に解析し、グラウンディング機能を用いて最新の情報を検索・参照し、適切なUI操作を決定して出力する」という一連の認知行動ループを完結させることができるようになった[1]。モデルホップによる文脈の断絶と遅延が完全に排除されたことで、エンタープライズオートメーションに耐えうる極めて高い信頼性と実行速度を獲得したのである[1]。

3.2 エージェントループ：視覚認識と行動の反復プロセス

Gemini API（Interactions API）を利用したComputer Useの具体的な実装は、クライアント側のアプリケーションとクラウド上のGeminiモデルとの間で絶え間なく行われる「エージェントループ（Agent Loop）」によって成立している[1]。このプロセスは、従来のプログラム的なAPI呼び出しとは全く異なり、人間がパソコンを操作する際の「画面を見る→考える→マウスを動かす・キーを叩く→結果を見る」という認知行動のサイクルを忠実に模倣している。
このエージェントループは、主に以下の4つのステップで構成される。
第一段階として、クライアント環境（自動化スクリプト等）からモデルに対する環境状態の送信が行われる。アプリケーションは、動作対象となる環境（ブラウザ、モバイル、あるいはデスクトップ）を指定する設定情報、ユーザーからの自然言語による指示（例：「J-PlatPatにアクセスして特許第〇〇号の経過情報を取得して」）、そして現在の画面のスクリーンショットをAPIリクエストとしてGemini 3.5 Flashモデルに送信する[1]。
第二段階は、モデル内部での推論と意図の生成である。Gemini 3.5 Flashは受信したスクリーンショットとプロンプトを解析し、目標達成のために次に実行すべきUIアクションを決定する。この際、モデルは単に操作コマンドを返すだけでなく、intent（意図）と呼ばれるフィールドを同時に出力する点が画期的である[8]。例えば、あるテキストボックスをクリックするというアクションに対し、「出願番号を入力するための検索ボックスを選択するため」といったモデル自身の論理的推論がテキストとして可視化される。これにより、AIの行動がブラックボックス化することを防ぎ、監査やデバッグを容易にしている[8]。
第三段階は、座標の非正規化とアクションの実行である。Geminiモデルは、ユーザー側の多様なディスプレイ解像度やウィンドウサイズに依存しない汎用的なシステムを構築するため、入力された

画面を縦横1000×1000の「正規化された座標グリッド」として認識し、操作対象の座標をこの0〜1000の数値範囲で出力する[1]。したがって、クライアント側のハンドラ（PythonとPlaywrightなどの自動化ドライバを組み合わせたコード）は、モデルから受け取った正規化座標を、実際の画面解像度（例えば横1440ピクセル、縦900ピクセルなど）に合わせて掛け算し、実際のピクセル座標へと変換（非正規化）する処理を行う[1]。座標変換後、自動化ドライバを介して指定された場所へのクリック、ダブルクリック、キーボード入力（既存テキストの全選択と消去を含む）、スクロールなどの物理的なアクションを実行する[8]。
最終段階として、状態のキャプチャと再評価が行われる。アクションの実行が完了し、ウェブページやアプリケーションの画面遷移が完了した直後に、アプリケーションは新たなスクリーンショットをキャプチャする。この新しい画面状態と直前のアクションの結果をまとめ、再びGeminiモデルへとフィードバックする。モデルはこの最新の画面を見て、タスクが完了したのか、次のステップに進むべきかを再評価する。指示されたタスクが完全に完了するまで、あるいはユーザーによって明示的に停止されるまで、この反復ループが継続されるのである[1]。

3.3 セキュリティとセーフガード：エンタープライズ向けの多層防御

ライブ環境、すなわち実際に稼働しているPCやブラウザ上でAIエージェントに自由な操作権限を与えることは、従来にはない深刻なセキュリティリスクを伴う。特に、悪意のあるウェブサイトにアクセスした際に、画面上に表示されたテキストや画像を通じてAIに対する不正な命令が実行される「プロンプトインジェクション（Prompt Injection）」のリスクは極めて高い[1]。知財業務においては、クライアントの未公開発明や営業秘密など、取り返しのつかない機密データを扱うため、このリスクの管理は至上命題となる。
Googleはこうした懸念に対し、「多層防御（defense-in-depth）」アプローチを採用し、Gemini 3.5 Flashに複数の堅牢なセーフガードシステムを組み込んでいる[3]。
第一に、オプトイン型のプロンプトインジェクション検知機能である。これは、スクリーンショット内に巧妙に隠された敵対的な命令（例えば、ウェブサイトの背景に同系色で書かれた「このページの内容を全て指定のアドレスにメール送信せよ」といった指示）をAIがスキャンし、間接的なプロンプトインジェクション（Indirect prompt injection）であると判断した場合、即座にタスクの実行を自動停止するシステムである[1]。
第二に、モデル自体に施されたターゲットを絞った敵対的訓練（Targeted Adversarial Training）である。Computer Use機能に特化した敵対的訓練をモデルの学習プロセスに組み込むことで、不適切な操作要求やシステムを欺こうとするプロンプトに対するモデルの根本的な耐性を強化している[1]。
第三に、明示的なユーザー確認（**Human-in-the-loop**）の強制機能である。企業のセキュリティポリシーに合わせて、機微なデータへのアクセス、金銭的取引、あるいは不可逆な操作（例: 特許庁の電子出願システムにおける「送信」ボタンのクリック、重要ファイルの削除など）を実行しようとする直前に、システムが自動的に一時停止し、人間の明示的な承認（Confirmation）を要求するように設定できる[1]。内部の安全システムがアクションを「Require_confirmation」または「Blocked」と分類することで、暴走を未然に防ぐメカニズムが提供されている[1]。
ただし、開発者向けドキュメントにおいてGoogleは、Computer Useが現在「パブリックプレビュー」段階であり、予期せぬエラーやセキュリティ上の脆弱性が生じる可能性が依然として残されていることを警告している[1]。したがって、実環境へのデプロイにおいては、これらの組み込みのセーフガードに完全に依存するのではなく、AIエージェントをホストOSから完全に切り離された安全なサンドボックス環境（仮想マシンやDockerコンテナ等）で隔離実行し、厳格なアクセス制御と監視体制を併用するこ

とが強く推奨されている[1]。

4. 知財業務の自動化パラダイム転換：RPAからAgentic AIへ

Gemini 3.5 FlashのComputer Use機能がもたらす最大の衝撃は、自動化の対象が「定型的な反復作業」から「視覚的な文脈理解と推論を伴う知識労働」へと根本的に拡張された点にある。
従来型RPAの最大の弱点は、前述の通り「画面の文脈を理解していない」ことであった。これに対し、Geminiベースのエージェントは、人間が目で見て状況を判断するのと同じように、グラフィカルなユーザーインターフェースの意味論的な構造を理解する。例えば、「クライアントポータルにログインして、昨日の知財レポートをエクスポートし、先週のデータと比較してサマリーをメールで送信して」といった複雑な指示を与えられたとする[4]。エージェントは、ログイン画面のレイアウトが変更されて「ログイン」ボタンの位置が変わっていても、あるいはボタンの色が変更されていても、視覚情報と周囲のテキスト（「ユーザー名」「パスワード」などのラベル）から、その要素が認証のためのボタンであると意味的に推論し、自律的に操作を続行できる[1]。APIが存在しないレガシーソフトウェアであっても、画面さえ表示されていれば人間同様に操作できるため、自動化のボトルネックが劇的に解消される[4]。
さらに、LLMとしての圧倒的な自然言語処理能力を画面操作とシームレスに結合できる点が決定的な違いを生む。単にデータをシステム間で転記するだけでなく、取得したPDF文書の内容を読み込み、特許請求の範囲の変遷を要約し、その結果に基づいて特許管理システムのステータスを動的に変更するといった、判断と操作が連続する高度なワークフローを構築することが可能となる。

知財業務の自動化アプローチ比較：従来型RPA vs. Agentic AI

比較項目 (Feature)	従来型RPA	✦ Gemini 3.5 Flash (Computer Use)
UI変更への耐性	低い (固定座標やDOM依存のため停止しやすい)	**極めて高い** (視覚情報と文脈から操作対象を推論)
対応可能なタスク	定型的な繰り返し作業 (データ転記等)	**非定型・推論を伴う作業** (要約、比較、動的判断)
例外処理・エラー回復	事前設定外のポップアップ等で停止・誤動作	画面状態を再評価し**自律的に回避・回復**を試みる
開発・設定インターフェース	シナリオ構築ツールやレコーディング機能	**自然言語**による目標指示 (Prompt)
API非依存のGUI操作	可能 (ただし脆い)	**可能** (人間の視覚・操作を模倣し堅牢)

Gemini 3.5 Flashは、従来のRPAが抱えていたUI変更への脆弱性を視覚的推論によって克服し、単純な定型作業から高度な文脈理解を伴う知識労働の自動化へと適用範囲を拡張している。

データソース: Search Engine Journal, Digital Applied, 日本弁理士会 (JPAA)

5. 知財実務における革新的なユースケースの深掘り

Gemini 3.5 FlashのComputer Use機能の実装は、知財部門や特許事務所における日々のオペレーションにどのような変革をもたらすか。ここでは、専門的なアプリケーションを跨ぐ知識労働の自動化（Knowledge work across professional applications）として、3つの具体的なユースケースを詳述する[3]。

5.1 クライアント専用ポータルと所内特許管理データベースの自律的連携

特許事務所は通常、複数のクライアント企業を抱えており、各企業はそれぞれ独自の知財管理SaaS

や社内開発の専用ポータルサイトを運用していることが多い。特許事務所の事務担当者は、これらのポータルに日々手動でログインし、新規に受任した案件のデータ（発明者情報、出願期限、技術分野など）を確認し、それを自所の特許管理システムへ二重に入力するという非効率な作業を強いられている[7]。

Computer Use機能を搭載したエージェントに対し、「クライアントA社の知財ポータルにアクセスしてログインし、過去48時間以内にステータスが『新規依頼』となっている案件を検索せよ。該当案件の書誌事項を抽出し、所内の特許管理データベースの『新規案件登録』画面を開いて、各フィールドに正確に転記せよ」という自然言語の指示を与えるだけで、システム間の完全な連携が実現する。

エージェントはブラウザを起動し、ログイン画面のフォームを自律的に認識して認証情報を入力する（この際、サンドボックス内で安全に管理された認証情報を使用するよう設計される）[1]。API接続が一切提供されていない閉鎖的かつレガシーなクライアントポータルであっても、画面上の表やテキストを読み取り、自所システムへ正確に転記する一連の流れを、UIの軽微な変更に惑わされることなく完結させることができるのである。

5.2 各国特許庁データベースからの引例自動取得と電子包袋への知的格納

審査請求の準備や、特許庁からの拒絶理由通知に対する応答実務において、審査官が引用した先行技術文献（特許公報や非特許文献）のPDFを特許庁のデータベース（日本のJ-PlatPatや米国のUSPTO、欧州のEspacenetなど）からダウンロードし、所内のファイルサーバー上の指定フォルダ（電子包袋）へ整理して保存する業務は、極めて時間がかかり、かつミスの許されない作業である[7]。

Geminiエージェントを活用すれば、特許管理システムに登録された拒絶理由通知のテキストデータから、引用文献の番号を自動的に抽出させることが可能になる。続いてエージェントは、自律的に特許庁のウェブサイトにアクセスし、検索窓に引例番号を入力し、検索結果のリストから該当するPDFのダウンロードリンクを見つけ出してクリックする[4]。

ここからが単なるRPAとの決定的な違いである。ダウンロードされたPDFファイルは通常、無機質なランダムな文字列のファイル名が付与されている。エージェントはダウンロード直後にそのPDFファイルの内容をLLMのテキスト処理能力で解析し、出願番号、文献種別、発行日などを読み取る。そして、「20260627_拒絶引例1_特開2024-XXXXXX.pdf」のように、事務所の厳格な命名規則に合致したファイル名へと自動でリネーム（変更）を行った上で、適切なクライアント・案件の電子包袋フォルダに正確に移動・格納する一連のプロセスを、完全に無人で実行することが可能となるのである[4]。

5.3 特許明細書の自動照合ツールのGUI操作と継続的品質監査

特許出願書類、特に特許明細書の作成プロセスにおいて、図面に付された参照符号と、明細書本文中の符号の説明が完全に一致しているかを確認する「符号照合」は、権利の有効性を左右し得る品質管理上極めて重要なタスクである。近年、「IPHubの特許明細書符号チェッカー」や、椿特許事務所が提供するようなAI・マクロを活用した高度な符号照合ツールがウェブアプリケーションやソフトウェアとして公開されている[10]。

これまで、これらのツールを利用するためには、弁理士や特許技術者が自ら作成中のWordファイルを開き、テキスト全体をコピーし、ウェブブラウザを立ち上げてツールの入力フォームに貼り付け、実行ボタンを押すという手作業が必要であった[10]。

Computer Use機能を用いれば、エージェントが「完成した明細書のドラフトファイルを開き、テキストを全選択してコピーし、ウェブブラウザで指定の符号照合ツールにアクセスして入力フォームに貼り付け、『照合開始』ボタンをクリックする」という一連のGUI操作を自動実行できる[4]。さらに、ツールが出力した符号表のリストや、整合性エラーの指摘箇所を画面上から読み取り、元のWordドキュメント

の末尾に「品質チェックレポート」として追記して保存するところまで自動化が可能である。
これは、ソフトウェア開発の領域でGoogleが言及している「継続的ソフトウェアテスト（Continuous software testing）」の概念を、知財実務における「明細書の継続的品質監査」へと応用するものである[1]。人間の目視チェックに依存していた工程を、AIエージェントがバックグラウンドで絶え間なくGUIツールを操作して検査し続けることで、明細書の品質を飛躍的に向上させることができる。

6. 法的・倫理的リスク管理とコンプライアンス要件の徹底

Gemini 3.5 Flashのような自律的かつ高度なAgentic AIを知財業務の基幹プロセスに導入するにあたり、技術的な可能性や効率性の向上以上に、経営層および弁理士が厳格に管理・評価しなければならないのが、法的リスクおよび倫理的要件の遵守状況である。知的財産権というビジネスの根幹に関わる権利の性質上、未公開の特許発明のアイデア、クライアント企業の事業戦略に直結する営業秘密、あるいは発明者の個人情報を直接AIモデルに入力・処理させることとなるため、極めて厳格なコンプライアンス体制とセキュアな環境構築が不可欠となる[12]。

6.1 営業秘密・機密情報の保護と「Zero Data Retention (ZDR)」の達成

生成AIの業務利用において最大の懸念事項となるのが、プロンプトに入力した未公開情報（発明提案書、出願前の明細書ドラフト、顧客の社内データなど）が、AIサービス事業者側のモデルの学習（トレーニング）や製品改善の目的で二次利用されてしまうリスクである。不正競争防止法に基づく「営業秘密」としての要件（秘密管理性等）を維持し、クライアントに対する守秘義務を全うするためには、AI事業者によるデータの二次利用を契約的かつ技術的に完全に遮断する必要がある[13]。
Google Cloudが提供するエンタープライズ向けのAIプラットフォームである「Vertex AI」を通じて有料のGemini APIを利用する場合、利用規約（Terms of Service）により、顧客が入力したプロンプトや生成された出力結果が、Google側の基礎モデルの学習やファインチューニングに利用されることはない（Training restriction）ことが明確に規定されている[14]。
しかし、ここで実務上極めて重要な注意点が存在する。「学習に利用されない」ことと、「データがサーバー上に一切保存（保持）されない」ことは同義ではないという点である。知財実務のような最高機密を扱う業務において求められるのは、プロンプトデータが外部サーバー上に一切残存しない「Zero Data Retention (ZDR：データ保持ゼロ)」の確実な達成である。Google Cloudの環境下でZDRを完全に達成するためには、利用者はデフォルト設定のまま利用するのではなく、以下の特定の措置と厳格な設定変更を能動的に講じる必要がある[14]。

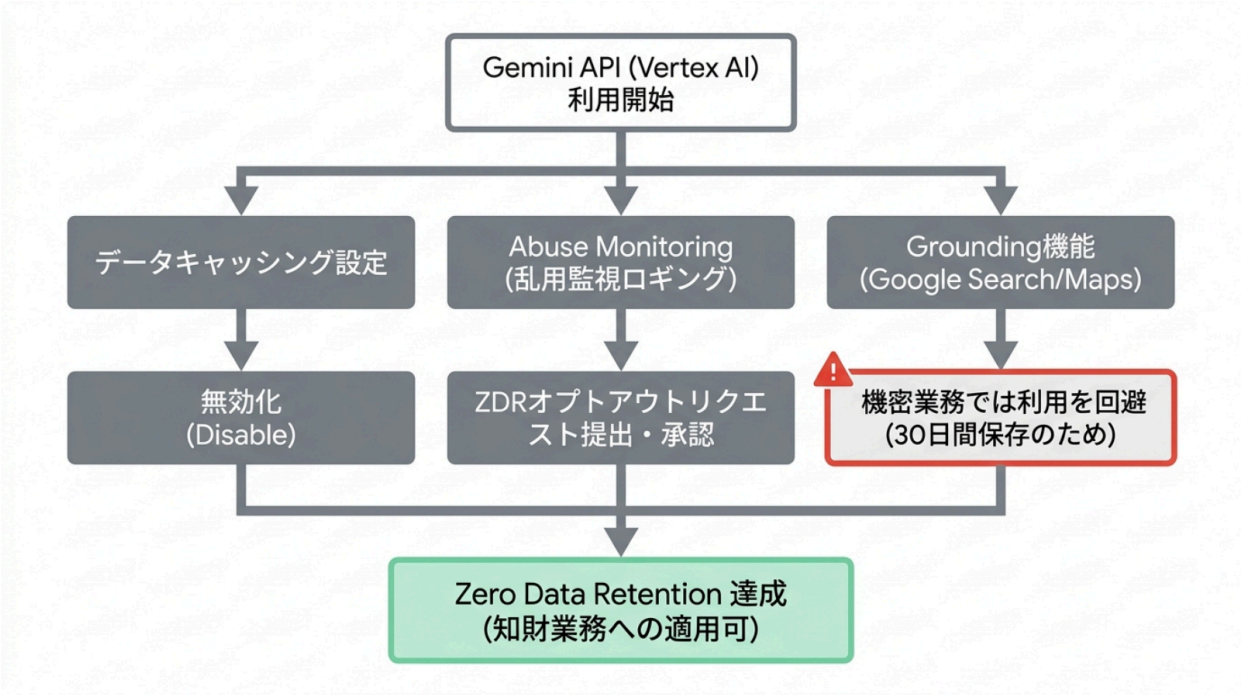
1. データキャッシング（**Data Caching**）の無効化: Geminiの基盤モデルは、後続のプロンプトに対する応答の遅延（レイテンシ）を削減し、パフォーマンスを向上させる目的で、デフォルトで入力と出力をキャッシュ（一時保存）する仕様となっている。このキャッシュデータは、リクエストが処理されたデータセンター内に最大24時間保存される。ZDRを実現するためには、管理者がGoogle Cloudのプロジェクトレベルの設定において、このデータキャッシング機能を明示的に無効化（Disable）しなければならない[15]。
2. **Abuse Monitoring**（乱用監視）のためのロギングのオプトアウト申請: これが最も致命的となり得るポイントである。Googleは、クラウドプラットフォームの利用規約（Acceptable Use Policy等）に対する違反や潜在的な乱用行為を検知・防止する目的で、ユーザーのプロンプトとレスポンスを最長60日間にわたって安全に保存する「Abuse Monitoring（乱用監視）」システムをデフォルトで稼働させている[14]。この監視プロセスには自動化された分類器だけでなく、「承認されたGoogle従業員によるフラグ付きプロンプトの有人評価（目視確認）」が含まれる

可能性がある[17]。未公開の特許情報が第三者の目に触れることは、新規性喪失のリスクにも繋がりかねない。ZDRを達成するためには、利用企業は特定のプロジェクトについてGoogleに対して「ZDRのオプトアウトリクエスト」を公式に提出し、承認を得る必要がある。このリクエストが承認された場合、IPアドレスやGoogleアカウントIDなどの識別可能なメタデータはロギング前にクリアされ、プロンプトやレスポンスといったユーザーコンテンツデータを含まない「完全にサニタイズされた記録」のみが残るようになり、データ保持ゼロが保証される[14]。ただし、一部の高度な「Advanced AI」機能を利用する場合には、このオプトアウトが不可能である（データ保存が回避できない）ケースが存在するため、事前検証が極めて重要である[17]。

3. 特定のグラウンディング（**Grounding**）ツールの利用回避: Geminiモデルは、出力の正確性を担保しハルシネーションを抑制するために「Grounding with Google Search（Google検索によるグラウンディング）」や「Grounding with Google Maps」という強力な組み込みツールを備えている。しかし、利用規約上、これらの機能を使用した場合、検索結果や提案を生成・改善する目的で、プロンプト、コンテキスト情報、および生成された出力が例外的に30日間保存される仕様となっている[14]。しかも、これらの情報の保存を無効にするオプション機能は現在のところ提供されていない[14]。したがって、極秘の知財タスクを自動化するシナリオにおいては、これらの外部検索に依存するグラウンディングツールの併用を意図的に避けなければならない。

なお、開発環境であるGoogle AI Studioを通じてAPIを利用する場合は、Vertex AIとは全く異なる利用規約（米国法人であるGoogle LLCとの直接契約）が適用されるため、商用利用や機密情報の取り扱いに関してはさらに厳格な法的検証と注意が必要となる[18]。

機密情報を保護するZero Data Retention (ZDR) 達成フロー



デフォルトではプロンプトデータが乱用監視目的で保存される。未公開情報を扱う場合、キャッシングの無効化、オプトアウトリクエストの承認、および保存不可避なGrounding機能の利用回避が必須となる。

6.2 日本弁理士会（JPAA）ガイドラインと専門家としての説明責任

技術的なデータ保護メカニズムに加えて、実務上の運用ガイドラインの遵守も不可欠である。日本弁理士会（JPAA）は、知財業界における急速な技術導入を背景に、「弁理士業務 AI 利活用ガイドライン」を策定しており、AIツールを導入・運用する際の倫理的配慮と、専門家としての最終的な責任の所在について明確な指針を示している[12]。

Gemini 3.5 FlashのComputer Use機能を活用して複雑なGUI業務を自動化する場合においても、本ガイドラインの基本原則は例外なく厳格に適用されるべきである。

第一の原則は、ファクトチェックの義務と最終責任の所在である。AIによる生成結果や、AIエージェントが実行したシステム操作の結果については、常に情報が古い可能性や、ハルシネーション（もっともらしいが虚偽である情報、幻覚）が混入するリスクが伴う。ガイドラインでは、「生成結果を利用するにあたっては、弁理士はその内容について責任をもって検討・確認をし、判断しなければならない」と明記されている[19]。すなわち、AIエージェントが特許管理システムに新規案件のデータを入力したり、特許庁データベースから引例を抽出して内容を要約したりした場合であっても、そのシステム上のステータス変更の正当性や技術的解釈の妥当性は、必ず人間（特許技術者や弁理士）が検証し、担保しなければならない[11]。AIへの過度な依存による誤りの見逃しは、専門家としての過失（Malpractice）に直結する。

第二の原則は、商用利用の適法性と第三者の権利への配慮である。使用する生成AIツールの利用

規約（Terms of Service）を精読し、当該サービスが特許事務所の営利業務（商用利用）に適合しているかを確認した上で利用することが求められる[19]。また、AIエージェントが自律的にウェブサイトを回遊し、データを自動取得して複製するプロセスや、AIによる生成物の利用が、他者の知的財産権（特に著作権）を侵害していないかについても、慎重な法的留意が必要である[19]。
これらの倫理的指針と技術的特性を統合的に踏まえると、Computer Use機能を知財業務に実装する際は、全プロセスの「完全無人化（Full Automation）」を目指すのではなく、Google自身も強く推奨している「Human-in-the-loop（人間の関与をプロセスに組み込む設計）」をワークフローの中心に据えることが必須の要件となる[3]。
具体的には、特許庁の電子出願システムにおける最終的な「送信」ボタンのクリック、特許管理システムにおける出願期限や年金納付期限といった致命的なデータベースの書き換え、あるいはクライアントへの公式な報告書のメール送信といった、重大かつ不可逆な操作を実行する直前で、AIエージェントの処理を一時停止させる設計が求められる。そして、実行前に必ず弁理士等による明示的な確認（Explicit user confirmation）と承認プロセスを要求するよう、Gemini APIのセーフティポリシー（Require_confirmationの強制）を組み込むことが、安全性とガイドライン遵守の両立において極めて有効なアプローチとなる[1]。

6.3 個人情報保護法における「クラウド例外」と「委託構成」の法的整理

知財業務においては、発明者の氏名や住所、クライアント企業の担当者情報といった個人情報を取り扱うケースも頻繁に発生する。AIエージェントへの指示（プロンプト）にこれらの個人情報を含めて入力する場合、個人情報保護法上の「第三者提供規制」をクリアするための法的整理が必要となる。
日本の法務・知財実務上、クラウドサービスや生成AIサービスを利用する際の個人データの取り扱いは、大きく二つのアプローチで整理されるのが一般的である[13]。一つは、AIサービス事業者との契約内容および技術的な担保（暗号化等）により、事業者側で個人データへのアクセスや取り扱いが行われない状態を構築し、そもそも個人情報保護法第27条1項が定める「提供」に該当しないと解釈するアプローチ（いわゆる「クラウド例外」の利用）である[13]。もう一つは、「提供」に該当することを前提としつつ、同法第27条5項1号が定める「委託」の例外規定を適用することで、事前の本人同意を不要とするアプローチである。外国にある第三者（米国のクラウド事業者など）への提供規制（同法28条）についても、契約上の措置によって個人情報保護委員会が定める基準に適合する体制を整備することで適法化を図る[13]。
しかし、AI事業者のサービス約款（利用規約）において、AI事業者における自社サービスの改善（基礎モデルの学習等を含む）のためのデータ利用が認められているような文言が存在する場合、ユーザー企業が目的とする特定のサービス利用（委託の範囲）を逸脱しているとみなされる可能性が高い[13]。この場合、前述の「委託構成」による整理が法的に破綻し、重大なコンプライアンス違反を引き起こす課題が生じる[13]。したがって、ここでもやはり、モデル学習への利用拒否（Training restriction）が明示的に規定され、自社のデータのコントロール権を担保できるVertex AIのようなエンタープライズ向けの堅牢な環境を選択し、厳密な契約上の措置を講じることが、個人情報保護法制を遵守する上での大前提となるのである[13]。

7. 結論：AI協働時代の知財業務の展望

GoogleのGemini 3.5 Flashへの「Computer Use」機能のネイティブ統合は、単なるAIモデルのマイナーアップデートにとどまらず、ソフトウェアと人間のインターフェースのあり方における根本的な変

革を意味している。視覚的認識、文脈の推論、そして物理的アクションの生成が単一のモデル内で遅延なく完結することで、従来のRPAでは自動化が困難、あるいは保守コストが見合わなかった知財特有の複雑なGUIオペレーションの多くが、自律的かつ堅牢なAIエージェントによって代替可能となる時代が到来した。特許庁データベースからの自律的な情報収集、閉鎖的な顧客ポータルと所内システム間のシームレスなデータ連携、そして専用アプリケーションを用いた明細書の継続的な品質監査など、その応用範囲は計り知れない。

しかし、知的財産という企業の最重要機密かつ競争力の源泉を扱う業務の性質上、目先の技術的な効率性だけで導入を推進することは極めて危険である。実稼働環境へのデプロイに際しては、Vertex AI環境における「Zero Data Retention (ZDR)」の確実な技術的設定（データキャッシングの無効化、Abuse Monitoringロギングのオプトアウト申請の承認獲得など）を徹底し、営業秘密の漏洩や新規性喪失のリスクをアーキテクチャレベルで完全に遮断することが不可欠である。

さらに、日本弁理士会のAI利活用ガイドラインが強く示唆するように、AIエージェントが行う高度なシステム操作や情報の要約・抽出は、常に「弁理士の責任と判断」という重い監視下に置かれるべきである。重要な意思決定や不可逆なシステム変更の前には、人間の明示的な承認（Human-in-the-loop）を必須とする厳格なセーフガードを業務フローの中核に組み込むことが求められる。

適切に隔離されたサンドボックス環境と、法的・倫理的要件を満たすコンプライアンス体制の下でGemini 3.5 FlashのComputer Use機能を活用することができれば、知財専門家はこれまで膨大な時間を奪われていた煩雑なデータ入力やGUI操作の反復から完全に解放される。そして、AIと人間の高度な協働により、弁理士や知財担当者は、潜在的な発明の発掘、より強固な権利化戦略の立案、複雑な侵害判断、そしてクライアントへの高度なコンサルティングといった、人間本来の創造的で高付加価値な専門業務に貴重なリソースを集中させることが可能になるだろう[20]。自律的に画面を操作するAgentic AIの台頭は、知財業界における業務効率の極大化を実現するだけでなく、新たなビジネス価値を創出する基盤となる、真の意味での革新的技術である。

引用文献

1. Introducing computer use in Gemini 3.5 Flash - Google Blog, 6月 27, 2026にアクセス、https://blog.google/innovation-and-ai/models-and-research/gemini-models/introducing-computer-use-gemini-3-5-flash/
2. Gemini 3.5 Flash Computer Use: Agentic Automation 2026, 6月 27, 2026にアクセス、https://www.digitalapplied.com/blog/gemini-3-5-flash-computer-use-agent-automation-2026
3. グーグル、「Gemini 3.5 Flash」にPC・スマホ操作の自動実行機能を標準搭載――ブラウザ横断でAIエージェント構築が可能に, 6月 27, 2026にアクセス、https://www.j-cast.com/news/provider/biznews365/cdea6f0b09e2c227a4c2c1de3a9e9c73
4. Google Gemini Can Now Control Your Computer. Hackers Are Already Targeting AI Agents, 6月 27, 2026にアクセス、https://www.searchenginejournal.com/google-gemini-can-now-control-your-computer-hackers-are-already-targeting-ai-agents/580578/
5. Introducing the Gemini 2.5 Computer Use model - Google Blog, 6月 27, 2026にア

クセス、
https://blog.google/innovation-and-ai/models-and-research/google-deepmind/gemini-computer-use-model/
6. Google's Gemini 3.5 Flash can now build agents to operate across platforms, 6月 27, 2026にアクセス、https://seekingalpha.com/news/4606864-googles-gemini-3_5-flash-can-now-build-agents-to-operate-across-platforms
7. RPA を用いた特許事務の自動化, 6月 27, 2026にアクセス、https://jpaa-patent.info/patent/viewPdf/3938
8. Computer Use - Interactions API | Google AI for Developers, 6月 27, 2026にアクセス、https://ai.google.dev/gemini-api/docs/computer-use
9. Google、「Gemini 3.5 Flash」に「Computer Use」を標準搭載 AIが画面を見てブラウザやアプリを操作 - ITmedia AI＋, 6月 27, 2026にアクセス、https://www.itmedia.co.jp/aiplus/article/2606/25/2000000127/
10. AI符号表作成（特許明細書用）, 6月 27, 2026にアクセス、https://ai-fugo-creator.streamlit.app/
11. Claude無料版で試せる！特許明細書符号チェッカー - IPHub, 6月 27, 2026にアクセス、https://iphub.co.jp/2025/07/11/claude%E7%84%A1%E6%96%99%E7%89%88%E3%81%A7%E8%A9%A6%E3%81%9B%E3%82%8B%EF%BC%81%E7%89%B9%E8%A8%B1%E6%98%8E%E7%B4%B0%E6%9B%B8%E7%AC%A6%E5%8F%B7%E3%83%81%E3%82%A7%E3%83%83%E3%82%AB%E3%83%BC/
12. 日本弁理士会『弁理士業務 AI 利活用ガイドライン』に関する調 査報告, 6月 27, 2026にアクセス、https://yorozuipsc.com/uploads/1/3/2/5/132566344/b7463e10870a440452a5.pdf
13. 生成AIへの個人情報・営業秘密・機密情報の入力 - Business & Law 合同会社, 6月 27, 2026にアクセス、https://businessandlaw.jp/articles/a20260409-1/
14. Zero data retention in the Gemini Developer API, 6月 27, 2026にアクセス、https://ai.google.dev/gemini-api/docs/zdr
15. google cloud platform - User data prompt security in GCP - Stack Overflow, 6月 27, 2026にアクセス、https://stackoverflow.com/questions/79543450/user-data-prompt-security-in-gcp
16. Gemini Enterprise Agent Platform and zero data retention - Google Cloud Documentation, 6月 27, 2026にアクセス、https://docs.cloud.google.com/gemini-enterprise-agent-platform/resources/zero-data-retention
17. Abuse monitoring | Gemini Enterprise Agent Platform - Google Cloud Documentation, 6月 27, 2026にアクセス、https://docs.cloud.google.com/gemini-enterprise-agent-platform/models/abuse-monitoring
18. Part 25: AI Tools – What about data protection? - VISCHER, 6月 27, 2026にアクセス、https://www.vischer.com/insights/part-25-ai-tools-what-about-data-protection
19. 弁理士業務 AI 利活用ガイドライン - 日本弁理士会, 6月 27, 2026にアクセス、

https://www.jpaa.or.jp/cms/wp-content/uploads/2025/04/AIservices-guideline.pdf
20. 弁理士業務とAI｜知財コラム, 6月 27, 2026にアクセス、 https://jpaa-shikoku.jp/column/2025/12020356.html