第 27 回日本オフィス学会大会 27th JOS National Convention
日本オフィス学会研究部会発表 A Report of the Study Group for JOS National Convention

AI 画像生成を活用したオフィス環境の人間工学的評価
想定環境の可視化による評価基準の統一化
オフィス人間工学研究部会

Ergonomic Evaluation of Office Environments Utilizing AI Image Generation
Standardization of Evaluation Criteria through the Visualization of Assumed Environments
Office Ergonomic Research Group

伴地 芳啓（早稲田大学）、三家 礼子（早稲田大学）、北島 洋樹（公益財団法人 大原記念労働科学研究所）、
石井 賢治（公益財団法人 大原記念労働科学研究所）、西谷 光雄（元業界紙編集者）、平尾 章成（芝浦工業大学）、
古阪 幸代（株式会社インデックスファシリティーズ・WFM・フルリエゾン）
Yoshihiro BANCHI, Reiko MITSUYA, Hiroki KITAJIMA, Kenji ISHII, Mitsuo NISHITANI,
Akinari HIRAO, Sachiyo FURUSAKA

1. はじめに

1.1 研究の背景

ハイブリッドワークなどの多様な働き方が定着する一方で、不十分な設備環境での業務がワーカーの身体に負担を蓄積させる要因として報告されている [1, 2]。働く場所の選択が生産性や健康に与える影響には個人差もあり、適切な環境選択は必ずしも容易ではない [3]。そのため、多様な働く環境に対して様々な観点から適合性を評価し、改善を促す仕組みが求められている。

適切なワークプレイス設計には、家具の寸法だけでなく、照明、自然要素、音環境など、心理や疲労回復に寄与する環境要因も考慮する必要がある [4, 5]。これらは作業時の疲労軽減だけでなく、ワーカー同士のコミュニケーションにも影響を及ぼす [6]。

当研究部会では前回大会において、大規模言語モデル（LLM）を環境の一次評価ツールとして活用する手法を検討した [7]。LLM に人間工学的な適合性評価を行わせることで、評価負荷を軽減し、改善の方向性を導き出すアプローチの可能性を示した。安全管理等における AI の活用は近年関心を集めており [8]、人間と AI の効果的な協働方法の模索が重要となっている。

1.2 評価前提におけるテキスト記述の限界と想定のブレ

しかし、LLM を利用した自動評価を実用化する上で、評価対象となる環境を言葉だけで表現した際に生じる、想定環境のズレが課題となった。

人間工学的な評価を正確に行うためには、机の広さや照度などの物理的な条件が重要な前提条件となる。しかし、「一般的な事務オフィス」といった言葉だけの提示では、テキストに書かれていない仕様について独自の解釈が入り込んでしまう。この課題は、大きく以下の 2 点にまとめられる。

- 仕様の認識のズレ：
  同じ言葉を提示されても、機器などの想定が人によって異なり、評価の前提条件が揃わなくなる。たとえば、広いデスクを想定する者と、機能が限られた机を想定する者といったズレである。
- コンテキストの抜け落ち：
  言葉だけでは周辺環境の情報が抜け落ちてしまう。人によって前提が異なることで、たとえば他人が頻繁に通行する不快感などが想像しにくくなる。

安全アセスメントなどの領域においても、現場の画像をインプットとすることで評価の信頼性を高められることが報告されている [9]。評価結果のばらつきを抑え、判断基準を一定に保つために、インプット情報を「画像」として視覚的に固定化する手法が考えられる。画像による可視化は、空間のコンテキストを一瞬で共有し、評価の境界条件を揃えることに有効である。一方で、評価が「画像に切り取られた特定の瞬間や局所的な状況」に過度に引っ張られ、個別的・表面的な課題抽出に偏ってしまうという課題もはらんでいる。実際に、製造現場の安全評価にマルチモーダル AI を適用した先行研究においても、2D の静止画像から安全距離などの空間的な関係性を正確に評価することには本質的な限界（視覚的な曖昧さ）があることが指摘されており [10]、画像という局所的な情報に依存することによる評価の偏りは、画像アセスメント特有の課題として認識されている。

1.3 研究の目的

本研究の目的は、画像生成技術を活用して環境条件を視覚的に可視化し、その画像を共通の評価対象として用いることで、人間工学的評価における前提条件の共有化が可能かを検討することである。

テキストベースの環境設定から脱却し、可視化された画像を共通の評価対象として提示することで、評価者間で生じていた解釈のブレを低減し、共通認識を形成する。具体的には、LLM にオフィス環境を想定させ、それを画像生成技術により可視化する。これらの共通画像をインプットとし、LLM に人間工学的な評価を行わせることで、人間と AI の協働評価手法の可能性を探る。

2. 研究手法

本研究では、オフィス環境評価における「前提条件の共有」という課題に対し、生成 AI を活用した以下の手順で検証を行った。

2.1 評価対象となるオフィス環境の想定

評価対象となるオフィス環境のバリエーションを整理するため、OpenAI ChatGPT および Google Gemini を用いた。両モデルに対し「人間工学専門家としてオフィスの評価を行っている」という役割（ペルソナ）を設定し、一般的なオフィス環境をいくつか想定させ、それぞれの詳細な仕様をテキストで記述させた。

2.2 画像生成 AI を用いた想定環境の可視化

次に、各モデルが出力した詳細なテキスト記述をそのまま各 AI の画像生成機能へのプロンプトとして入力し、オフィス空間の画像を生成させた。具体的には、ChatGPT の記述からは ChatGPT に統合されたネイティブな画像生成機能を、Gemini の記述からは Gemini に統合された画像生成モデルである Imagen 3 を用いた。言葉のみの条件提示で生じるイメージのズレを防ぎ、AI が想定した環境を画像として可視化・固定するアプローチを採用した。

2.3 可視化画像に対する人間工学的評価

可視化されたオフィス画像を対象として、改めて ChatGPT および Gemini を用いて人間工学的な評価を実施した。「画像のオフィスの、良い点と悪い点、改善案を列挙してください」という指示とともに各画像を入力し、視環境、作業姿勢、動線、心理的側面などの具体的な課題と改善案を抽出させた。

2.4 複数モデル間の評価特性の比較分析

同一の画像に対して各モデルがどのような評価を行うかを比較するため、本研究では Gemini によって想定・可視化されたオフィス画像を代表例として共通の評価対象に設定した。これを ChatGPT と Gemini の双方に入力し、両モデルの適合性・リスクの指摘や改善提案の着眼点の違いを定性的に比較分析した。

3. 想定オフィス環境の可視化と評価結果

3.1 テキスト提示時における環境想定と画像出力の差異

「一般的な事務オフィス」というテキスト条件から各モデルに環境を想定・可視化させた結果、前提とする物理条件に違いが確認された。

・Gemini の想定と可視化：

日本のオフィスの時代性や現場の空気感を重視した 3 タイプを提示した。画像にはブラウン管モニターや書類の山など、やや古い機器や雑然とした様子が反映された。

・ChatGPT の想定と可視化：

寸法や照度などの具体的な数値スペックに基づき、在宅勤務を含めた 5 タイプを提示した。画像には最新の機器や機能的な設備が整然と配置された。

言語による条件設定のみでは、モデル間で想定する前提条件に明らかな解釈のズレが生じることが確認された。

3.2 共通画像の提示による評価特性と観点の差異

前提条件のズレを解消するため、Gemini が生成した 3 タイプの画像を共通の評価対象として両モデルに評価を行わせた。対象空間の物理的ファクトが統一された状態においても、両モデルの評価アプローチには明確な違いが観察された。

・伝統的な島型オフィスの評価：

Gemini は「受話器を肩と耳で挟む」といった姿勢から生じる身体的リスクを抽出し、机上の整理などの現場改善を提案した。ChatGPT は通路幅の確保などの管理・統制上の利点を評価し、ルール策定など組織的なアプローチを提案した。

・ABW オフィスの評価：

Gemini は調整機能のない椅子やノート PC の直置きによる腰や首への局所負荷を警告し、PC スタンドの利用など個人の工夫を推奨した。ChatGPT は心理的ストレス軽減効果を評価する一方で、集中エリアと協働エリアのゾーニング分離など空間設計の更新を提案した。

・開発・オペレーションオフィスの評価：

Gemini は輝度対比による眼精疲労や姿勢の固定化といった生理的リスクを指摘し、ディスプレイ背面の照明追加などを提案した。ChatGPT は集中しやすい空間構造や管理者からの視認性を評価しつつ、長時間の複合的な作業負荷に対する設備や運用ルールの改善を提案した。

4. 考察

4.1 画像化によるコンテキスト固定と前提の共有

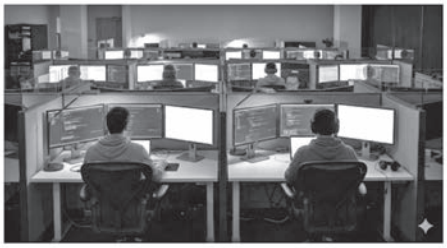
テキストによる条件提示では、モデルによって機器や設備



島型オフィス



フリーアドレス・ABW



開発・オペレーション

図 1 Gemini による生成画像



島型レイアウト



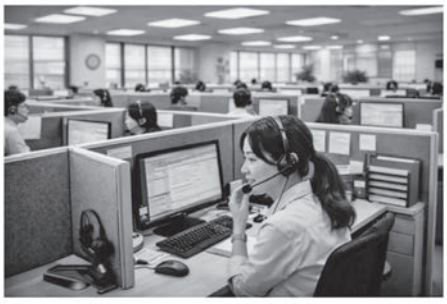
フリーアドレス型・オープン



研究職・設計職



在宅勤務・テレワーク



コールセンター

図 2 ChatGPT による生成画像

などの暗黙の前提条件に独自の解釈が入り込むことが確認された。一方で、画像を用いて前提条件を視覚的に固定することは、解釈のブレによる評価のばらつきを抑え、アセスメントの境界条件を一定に揃えるために有効な手法であると考えられる。

4.2 モデルの評価特性の差異と人間工学的評価の相補性

同一の画像を入力した場合でも、モデルによって評価アプローチに明確な違いが観察された。Gemini は、個人の身体感覚に主眼を置いたミクロな視点からのボトムアップ的なアプローチを示した。不自然な姿勢による局所的な負担などを抽出し、身の回りの工夫で即時対応できる具体的な対策を提案する傾向がある。一方、ChatGPT は、空間構成や管理体制に主眼を置いたトップダウン的なアプローチを示した。照明の均一性や通路幅などを管理者視点の安全性として捉え、規格に準拠した定量的な指標やルール策定など、仕組みによる解決を提案する傾向がある。このようなモデル間の評価特性の違いは、プロンプトで役割（ペルソナ）や出力形式を詳細に指定することである程度の調整や方向付けは可能であると考えられる。しかし、各モデルの学習データやアーキテクチャに根ざした本質的なアプローチの差異をプロンプトのみで完全に同一化することは難しい。したがって、これら「ミクロな身体エルゴノミクス」と「マクロなファシリティシステム」は補完的な役割を担っており、両者の視点を統合して評価の一次評価案を作成することは、評価漏れの少ない多角的なオフィス設計・運用に有用であることが示唆された。

4.3 AI 活用の限界と「AI と人間」の協働モデル

第一に、LLM は提示された画像データの範囲や与えられた役割に強く縛られる。そのため、画像に描かれていない暗黙の行動特性や人間関係、専門家が持つ暗黙知といった定性的なコンテキストを完全に理解･評価することは困難である。第二に、人間工学が重視する「人間中心設計」は、ワーカー個人の身体特性や感情などに寄り添う個別性を本質とする。

AIが出力する評価は体系的な「一般解」に留まるため、各人の個別性に対応した配慮や微調整には、人間の専門家によるヒアリングや介入が必要となる。したがって、人間工学評価における生成AIの役割は専門家の代替ではなく、幅広い代替案を提示して専門家のスコープを拡大する支援ツールと捉えるのが妥当である。AIが画像から客観的なリスクを網羅的に抽出し、それを基に人間の専門家が状況的解釈や個別対応を加える「ハイブリッドなアセスメント」が、多様化する現代の実務において実効性が高いと考えられる。このように、AIを意思決定の代替ではなく、専門家の判断を補完し議論を深めるためのツールとして位置づけるアプローチは、安全管理や事故調査におけるAI適用の先行研究[11]の知見とも合致している。

5. おわりに

本研究では、オフィス環境の人間工学的評価における前提条件のズレを解消するため、AIを用いて環境を視覚的に固定化し、複数のLLMによる適合性評価の有効性を検証した。

テキスト提示のみでは、LLMは過去の学習データに引きずられた独自の思い込みにより、解釈の乖離を生じやすい[12]。対して、画像を共通インプットとすることで、与えられた具体的な情報に基づいた客観的な分析が働き、不適切な推測を排除できるとされる。本研究でも、画像による前提条件の固定化が解釈のブレを効果的に抑制し、より客観的なアセスメントを可能にすることが確認された。また、同一画像に対するクロス評価を通じて、身体感覚や局所的な負荷に焦点を当てるGeminiの「ミクロな視点」と、規格適合や空間管理に優れるChatGPTの「マクロな視点」が、互いに補完的な関係にあることが明らかになった。

働き方や環境が多様化する現代の人間工学においては、ワーカーと個別の家具との関係性だけでなく、「オフィス空間全体」をシステムとして統合的に評価する視点も求められている。空間の全体像を画像で捉え、ミクロとマクロの両面からアプローチする本手法は、複雑化するワークプレイスの実践的なアセスメントにおいて非常に有効である。

本手法を用いた複数モデルによる評価案の生成は、専門家の評価スコープを拡張し、多角的なアプローチを提供する。しかし、AIは学習データに依存しており、組織特有の人間関係やワーカーの身体的既往など、画像に表れない定性的なコンテキストを理解することはできない。

今後は、本手法で得られたAIの評価結果と、人間の専門家によるアセスメント、ならびにワーカーの主観評価やウェアラブルデバイスを用いた生理学的計測データとの定量的な相関分析が期待される。

参考文献

[1] Bloom, N., Han, R. & Liang, J. Hybrid working from home improves retention without damaging performance. Nature 630, 920–925 (2024).

[2] 甲斐 裕子 他 , テレワークの常態化による労働者の筋骨格系への影響や生活習慣病との関連性を踏まえた具体的方策に資する研究 , 厚生労働科学研究成果データベース , 2023.

[3] 久米 功一 , 鶴 光太郎 , 川上 淳之 , 在宅勤務で個人の生産性は どう変わるか , RIETI Discussion Paper Series, 23-J-044, 2023.

[4] Qinghua Lei, Chao Yuan, Stephen Siu Yu Lau. A quantitative study for indoor workplace biophilic design to improve health and productivity performance, Journal of Cleaner Production, Volume 324, 2021, 129168.

[5] 石川 敦雄 , 視環境要因が対人認知・行動に及ぼす影響に関する文献レビュー：明るさ，色，そして奥行きに着目して , 日本建築学会環境系論文集 , 2022, 87 巻 , 802 号 , p. 797-808.

[6] 石川 敦雄 , 楠見 孝 , オフィス環境デザインがコミュニケーションに対する期待に及ぼす影響 , 日本感性工学会論文誌 , 2020, 19 巻 , 2 号 , p. 135-143.

[7] 日本オフィス学会オフィス人間工学研究部会 , 作業内容と作業環境の人間工学的検討 , 第 26 回日本オフィス学会大会梗概集 , 2025.

[8] Alireza Mortezapour. Ergonomic LLM or LLM for Ergonomics? Prompt engineering insights for an interventional case study, 22 April 2024, PREPRINT (Version 1) available at Research Square.

[9] Şekerci, Y., & Özkul, P. (2025). EVALUATING THE EFFECTIVENESS AND PROMPT ALIGNMENT OF TEXT-TO-IMAGE AI MODELS IN OFFICE LIGHTING DESIGN: A COMPARATIVE ANALYSIS. ArtGRID - Journal of Architecture Engineering and Fine Arts, 7(2), 233-256.

[10] Katsampiris-Salgado, K., Dimitropoulos, N., & Michalos, G. (2026). Generative AI-based risk assessment and hazard identification towards safety compliant workplace design. CIRP Annals.

[11] 中西 美和 , 特集：事故調査と AI, 人間工学 , 2025, 61 巻 , 3 号 , p.154-159.

[12] Lou, S., Chen, Y., Liang, X., Lin, L., & Zhang, Q. (2024). Quantifying In-Context Reasoning Effects and Memorization Effects in LLMs. arXiv preprint arXiv:2405.11880.