



AIによる人事評価はどのように 職場に受け入れられるか

杉谷 陽子

(上智大学教授)

1 はじめに

産業場面において、人工知能（artificial intelligence：AI）が活用されるようになって久しい。身近なところでは、「チャッピー」（OpenAI社が提供するChatGPTの愛称）が2025年の流行語大賞の候補となるほどに普及し、対話型生成AIは日常生活でもビジネスの現場でも、もはや欠かせない存在となっている。また、消費者から直接は目に見えない形でもAIは多方面で活用されている。例えば、需要に応じてチケットの価格を決定したり、閲覧者の関心に合わせた広告を表示するなど、ビジネスの効率化と品質向上に大きく貢献している。

本稿では、AIが組織マネジメント、より具体的には、人事評価に対してどのような影響をもたらしているのかについて考察していく。人事評価とは、組織内の従業員の業績や能力を評価し、人材の配置・報酬・育成に反映するプロセスである。人事評価のプロセスおよびその評価結果は、組織メンバーのモチベーションやウェルビーイング（幸福感）に大きな影響を与え、結果的に、組織パフォーマンスに直結する。

人事評価は組織マネジメントにおいて非常に重要なプロセスであるため、各組織において、多くの時間的、経済的コストを要するものである。したがって近年では、人事審査の効率化を目的として、大手企業を中心に人事評価の一端をAIによって代替しようとする動きが活発化している。一例として、IBM Watsonを活用した書類選考が挙げられる（ソフトバンク株式会社2017）。人間が手分けして応募書類を読むのに比べ、疲れ知らずのAIは、短時間で審査を終えることが可能である。また、新卒採用の現場では、1次面接をAI面接官が担当したという企業も出始めている（日本経済新聞2025）。

2 AIによる人事評価に対する懸念

AIのアシストによって、人事評価の質が保たれながらも効率化されるのであれば、多くの組織にとって歓迎すべき進展である。しかしながら、AIによる人事評価には、人間が行う従来型の人事審査と比べた時に、さまざまな懸念や課題があることが指摘されている。

その懸念の1つは、AIの判断は信頼に値するものであるのか現状では不明である、という点である。AIの判断は、学習データに基づくアルゴリズムの推定結果にすぎない。データ不足や問いが曖昧な時など、AIは事実と異なる内容を生成することもある（この現象は「ハルシネーション」と呼ばれている）。したがって、とりわけ重要な意思決定においては、人間の目でも確認する作業が求められることが多く、AI単独の判断には慎重さが求められるという指摘である。もう1つ、しばしば強い懸念が示されるのが、AIの判断は過去の慣例に基づいてバイアスがかかるため、偏見を助長・再生産する可能性がある、という指摘である（概説として、Hunkenschroer and Luetge 2022など）。先に述べた通り、AIは学習データに基づく予測を返すため、学習データに時代錯誤の慣習が多く含まれていると、それらを反映してしまい、現在では倫理的問題を含む判断を下すことがある。例えば、過去の人事データに、正規雇用が男性に偏る傾向が含まれていると、AIはその傾向を学習してしまう。結果として、書類審査で男性に有利な判断を下す可能性がある。ソーシャルメディアの広告配信はアルゴリズムで「最適化」される結果、中立的に作成した求人広告であっても配信先が特定の属性に偏るといった報告もある（Ali et al. 2019）。ヨーロッパではAI規制法（EU Artificial Intelligence Act; Regulation (EU) 2024/1689）によって、採用や人事評価等に用いられるAIシステムはハイリスクに分類され、人間による監督を

可能にする設計と運用の要件が課されている。

3 アルゴリズム忌避

人事評価に限らず、人は、これまで人間が担ってきたさまざまな活動を AI が代替することへ嫌悪感を示すことが知られる。これを「アルゴリズム忌避」(algorithm aversion; Dietvorst, Simmons and Massey 2015) という。アルゴリズム忌避は、その領域において AI が人間のパフォーマンスをすでに凌駕するレベルにあることを知ってもなお、「なんとなく AI は嫌だ」と感じることを指す。例えば、大変美しい絵画に出会って感動し、自宅のリビングに飾りたいと思った時に、この絵の作者は AI だと言われると高い金額に見合う価値はないと感じるようになる (Heimstad, Wien and Gaustad 2025)。

人事評価についても、アルゴリズム忌避が生じることが報告されており、従業員は一般的に、AI よりも人間 (例: 上司・同僚) に評価される方が望ましいと感じている (Bankins et al. 2024)。前述の通り、AI の判断には、ハルシネーションの問題や倫理的懸念も指摘されていることに鑑みれば、人事評価に AI を活用すること自体が妥当性に乏しいようにも見える。

しかし、これらの否定的意見を理由として、組織は AI アシスタントの活用を、将来にわたって断念すべきなのであろうか。おそらく多くの読者の意見は否であらう。世界の情勢に鑑みれば、むしろ、どうやって AI を活用していけばよいかを戦略的に検討すべき局面である。なぜならば、上述の問題の多くは、AI 技術の進展や、運用・活用の改善によって、長期的には解消されていくと思われるからである。

例えば、AI が過去の偏見を再生産するのではないかという懸念については、AI に学習させるデータを吟味すること、アルゴリズムのデザインの調整、ユーザー側が AI への指示 (プロンプト) を工夫することによって、一定程度は低減可能である。近年では、領域特定の AI (例えば、動画作成 AI、医療診断 AI など) の開発も急速に進んでおり、それらは従来の汎用的なモデルの問題を改善し、高いパフォーマンスを発揮している。問題の多くの部分は、人間が AI の特性をよく知り、AI を使いこなせるようになることで軽減される。したがって、AI の技術的な限界に立脚した議論は、AI による人事評価が持つ本質的問題とは言えないと筆者は考えている。

AI による人事評価を導入する際に経営者が最も考えなければならないのは、AI の技術的進展によっては解決されないところにある。それはすなわち、AI によって評価されることを従業員が「不公平だ」「理不尽だ」と感じたり、AI を選考に用いることで応募者が「あの企業は従業員に冷たい」という印象を抱いたりするような、主観的かつ感情的な問題である。

4 課題特性や目標に依存したアルゴリズム忌避

ここで有名な研究を紹介したい。Castelo, Bos and Lehmann (2019) による“Task-Dependent Algorithm Aversion” (タスク依存のアルゴリズム忌避) というタイトルの論文である。本論文はまず、人は、AI が得意な課題 (タスク) に関する素朴信念 (Lay belief) を持っている、と指摘する。素朴信念とは、一般の人が抱く先入観 (『○○は△△である』という思い込み) である。Castelo らは、人々は、「アルゴリズム (AI) は客観性の高いタスク (論理的でルールに基づく作業) には向いているが、主観的タスク (感情的で、直感が求められる作業) には弱い」という素朴信念を持っていると指摘した。これは AI に限らず、古くは、コンピューターや機械などに対して人間が持っていた素朴信念とも共通する感覚である。このような強い素朴信念の存在によって、例えば、天気予報や統計的な分析、株価の予測など、数理的でルールベースのタスクでは、AI が活用されることに拒否反応を示す人は少なく、むしろ賛成派が多くなる。一方で、「人柄」や「個別の事情」や「感情」をふまえた判断が求められる領域、例えば、医療診断、芸術活動、恋愛相談、人事採用などでは、AI 利用への拒否反応が出やすい。

この研究が示唆するのは、AI を円滑に活用するために、企業が人々の素朴信念を理解し、それに沿った活用方法を設計する必要があるという点である。例えば、近年では、オンラインで買い物をする際に、AI が「あなたへのおすすめの商品」を表示してることがある。消費者はこれらの AI レコメンデーションをどれくらい信用するのだろうか。Longoni and Cian (2022) は、素朴信念に基づき、AI レコメンデーションが受け入れられるかどうかは、消費者が製品を購入する目的が実用的 (utilitarian) か、快楽的 (hedonic) かに依存すると論じている。実験では、人間のパティシエがチョコレートを推薦するアプリと AI がチョコ

レートが推薦するアプリが比較された。チョコレートの味や香りを楽しむ快樂の文脈では「AIは感情・経験を理解できない」という素朴信念から、AIによるレコメンデーションは信頼されにくかった。しかし、栄養価や原材料など科学的情報を重視する実用的文脈ではこの傾向が逆転し、AIのレコメンデーションの方が信頼できると評価された。

さて、人事評価は、どのような「タスク」と位置づけられるであろうか。人事評価は決して、数理的でルールベースのタスクではなく、ロジックによって正解が導き出せるようなものでもない。先に挙げたCastelo, Bos and Lehmann (2019)の研究でも、人事採用や解雇判断は、「極めて客観性が低いタスク」に分類されており、アルゴリズム忌避が高かった。特に、評価される側の従業員の立場からは、「自分の努力や成長、困難を乗り越えた事情、チームへの貢献」といった“物語”を理解してほしいと願うものである。そこにAIが介入すると、人間性を無視して、結果だけを冷徹に点数化されて評価が決まってしまうように感じられ、強い反発を生むことが多くの研究で示されている(Newman, Fast and Harmon 2020)。実際には、AIに学習させるデータを工夫すること、プロンプトで適切な指示を送ることで、AIに情意面を考慮させることは一定程度可能だ。しかし、素朴信念の効果は大きく、「AIが使われている」と聞くだけで、従業員は大事なところを分かってもらえないという直感的な不安を感じる。

5 AIが人間の価値への「脅威」となる

人がアルゴリズム忌避を生じる原因の1つとして、AIは主観的・感情的なタスクが苦手という素朴信念が介在していることを論じた。しかし、それが原因のすべてではない。もう1つの重要な論点として、AIが、人間の価値を低下させる「脅威(threat)」として認識されている、という問題も指摘されている。

ここでまた関連研究を紹介したい。Leung, Paolacci and Puntoni (2018)は、「人は自分のアイデンティティに深く関わる活動について、それらが自動化されることを嫌がる(自分にとってどうでもよいタスクは自動化されることを歓迎する)」ことを明らかにした。この論文では6つの実験を行い、さまざまな日常的タスクの自動化(automation)に対して、人によって反応が大きく分かれることを明らかにした。例えば、

料理が得意な人は自動調理機(投入した材料から自動的に調理を完了する機械)を非常に低く評価するが、料理が好きではない人はその機械の採用意向が高かった。釣りが趣味の人は自動フィッシングマシンを忌避した。これらの結果は、人は、「自己アイデンティティの形成に深く関わっている活動」、すなわち、自分らしさを感じられる活動、自身が誇りを持っている技能、周囲から称賛を得たいと思っている活動が、自動化によって必要なくなっていくことに抵抗する、ということを示唆している。言い換えれば、人は「自動化自体が嫌」なのではなく、「自動化されることで、自分の価値や居場所を脅かされたくない」と思っているのである。

アルゴリズム忌避のメカニズムも同様であると考えられる。つまり、アルゴリズム忌避が生じる原因の1つは、AIが人間にとって大切なもの、すなわち、自己アイデンティティへの脅威として認識されていることに依る。Morewedge (2022)は、数多くのアルゴリズム忌避研究をレビューし、アルゴリズム忌避の正体は、「AIを忌避している」のではなく、「人間を過大評価している」ことに由来すると論じた。人は、AIが間違いを犯す可能性や、公平性への懸念、あるいは、感情的で社会的なタスクが苦手、という特徴によってアルゴリズムを忌避しているのではない。「AIには無理」と考えることで、人間の価値、そして、自己価値を高めたいという欲求を満たしているのだと考えられる。

人事評価は、職務能力や成果を通じて「職業人としての自分」の価値を確認する機会である。そこでAIが主役となってしまうと、評価される従業員にとっては、「自分自身が数値化され、人として扱われていない」「自分の個性を認められていない」と感じられ、自己価値を否定されたような心理的痛みを生みかねない。また、評価する側である人事のプロフェッショナルにとっても、自分のこれまでの仕事が簡単にAIに置き換わってしまうことは、自身が培ってきた能力や成果の価値が否定されたような気持ちになるであろう。いずれの観点からも、人事評価にAIを導入する際には、AIが現場の人間にとって「アイデンティティへの脅威」として認識されないよう、AIを「アシスタント」として位置づけることが肝要と言える。

6 「公平」な調停者としての AI

さて、ここまでアルゴリズム忌避を中心とし、AI 活用への懸念や注意点について論じてきたが、最後に、AI による人事評価の肯定的側面に焦点をあてたい。

繰り返しであるが、AI は論理的で客観的なタスクが得意で、主観的で感情的なタスクが苦手、という素朴信念がある。この素朴信念が、AI による人事評価は成果偏重的で、被評価者の人間性やチームへの配慮などの心理的側面を無視した評価を生むという連想につながり、忌避の反応を生じると言われてきた。

しかし、この AI の特性は、裏を返せば、人間関係の相性や付度を超えて、客観的で公平な評価を可能にするもの、と見ることもできる。Sugitani, Togawa and Motoki (2025) は、日本の職場の文脈に着目し、特に人間関係があまり良くない職場においては、AI による人事評価が肯定的に受け入れられることを明らかにした。以下のような状況を想像して欲しい。あなたは正社員として働いているが、同僚や上司との人間関係が好ましくなく、困っていても助けが得られない。この状況で上司から人事評価を受けるとなると、あなたはどうか感じるだろう。たとえ高い業績を示しても、人間関係に由来する不当な低評価が心配になる。勤務態度や協調性が重視されがちな日本の文脈では、このリスクは高いと感じられよう。ここで AI による人事評価が導入されれば、数量的・論理的な判断を得意とし、感情や関係性を捨象しがちという素朴信念から、職場の人間関係は考慮されにくいという期待が高まる。したがって、人間関係による評価リスクを負っている従業員にとって AI は「公平な評価者」として受け入れられるだろう。

他方、職場で密なコミュニケーションを心がけ、上司・同僚と好ましい関係を築いてきた従業員は、協調性や社会性で高い評価を得られると期待している。そうすると AI 評価では、築いた信頼感が評価されず、営業成績などが点数化されるだけだという「不公平感」が生じ、受け入れがたく感じられるだろう。

Sugitani, Togawa and Motoki (2025) では、3つの実験を行って、職場の人間関係が悪い場合には、人間関係が良好な場合と比べ、AI による人事評価の「公平性」の認知が高まることで、アルゴリズム忌避が減少することを明らかにした。ただし、この結果には文化差が見られた。日本のように、組織メンバー同

士の協調性や他者への配慮が重んじられる職場においては職場の人間関係がアルゴリズム忌避に影響したが、米国のように個人が独立した存在としてとらえられる職場においては、人間関係はアルゴリズム忌避と関連しないことが示唆された。

7 まとめ

本稿では、人事評価における AI 活用の可能性と注意点について、「アルゴリズム忌避」と呼ばれる現象に関する先行研究を紹介しながら論じた。現時点では、人は、これまで人間が担ってきたさまざまな活動（社会的場面・産業場面）について、基本的には AI よりも人間が担当した方が望ましいと感じている。その原因として、AI のパフォーマンスは人間に及ばない、不安定で信頼できない、倫理的懸念やバイアスを含むリスクがある、といった問題が指摘されてきたが、アルゴリズム忌避の本質は、これらの技術的境界ではない。むしろ、AI が人間にとって大切なもの（自己価値、アイデンティティ、相互信頼など）に対する脅威として認識されているという点にあることを指摘した。

したがって、組織活動へ AI 導入を検討するとき、経営者が留意すべきは、「AI をどのように導入するか」ではなく、「AI をどう位置づけるか」である。従業員にとって AI が「あなたの代わり」「ライバル」「監視者」となった時には忌避の態度は強まる。一方で、AI を「あなたのサポーター」「秘書」「面倒な作業を引き受ける補助者」として設計し、最終判断やコミュニケーションによる調整の可能性を人間側に残すのならば、受容されるであろう。また、組織のヒエラルキーや対人関係に由来したコンフリクトが生じた際に、それを調停する第三者として、AI の意思決定を活用するという方法も有効であろう。

参考文献

- 日本経済新聞 (2025) 「新卒採用に AI 面接官、キリン HD や ローソン導入 資料作成負担も軽減」(アクセス日: 2025 年 12 月 4 日)
- ソフトバンク株式会社 (2017) 「新卒採用選考における IBM Watson の活用について」https://www.softbank.jp/corp/group/sbm/news/press/2017/20170529_01/ (アクセス日: 2026 年 1 月 9 日)
- Ali, Muhammad, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove and Aaron Rieke (2019) "Discrimination through Optimization." *Proceedings of the ACM on Human-Computer Interaction*, Vol. 3, No. CSCW, pp. 1-30.

- Bankins, Sarah, Anna Carmella Ocampo, Mauricio Marrone, Simon Lloyd D. Restubog and Sang Eun Woo (2024) "A Multilevel Review of Artificial Intelligence in Organizations: Implications for Organizational Behavior Research and Practice," *Journal of Organizational Behavior*, Vol. 45, No. 2, pp. 159-182.
- Castelo, Noah, Maarten W. Bos and Donald R. Lehmann (2019) "Task-Dependent Algorithm Aversion," *Journal of Marketing Research*, Vol. 56, No. 5, pp. 809-825.
- Dietvorst, Berkeley J., Joseph P. Simmons and Cade Massey (2015) "Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err," *Journal of Experimental Psychology: General*, Vol. 144, No. 1, pp. 114-126.
- Heimstad, Sigurd Birk, Anders Hauge Wien and Tarje Gaustad (2025) "Machine Heuristic in Algorithm Aversion: Perceived Creativity and Effort of Output Created by or with Artificial Intelligence," *Computers in Human Behavior: Artificial Humans*, Vol. 5, 100190.
- Hunkenschroer, Anna Lena and Christoph Luetge (2022) "Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda," *Journal of Business Ethics*, Vol. 178, No. 4, pp. 977-1007.
- Leung, Eugina, Gabriele Paolacci and Stefano Puntoni (2018) "Man Versus Machine: Resisting Automation in Identity-Based Consumer Behavior," *Journal of Marketing Research*, Vol. 55, No. 6, pp. 818-831.
- Longoni, Chiara and Luca Cian (2022) "Artificial Intelligence in Utilitarian vs. Hedonic Contexts: The 'Word-of-Machine' Effect," *Journal of Marketing*, Vol. 86, No. 1, pp. 91-108.
- Morewedge, Carey K. (2022) "Preference for Human, Not Algorithm Aversion," *Trends in Cognitive Sciences*, Vol. 26, No. 10, pp. 824-826.
- Newman, David T., Nathanael J. Fast and Derek J. Harmon (2020) "When Eliminating Bias Isn't Fair: Algorithmic Reductionism and Procedural Justice in Human Resource Decisions," *Organizational Behavior and Human Decision Processes*, Vol. 160, pp. 149-167.
- Sugitani, Yoko, Taku Togawa and Kosuke Motoki (2025) "Socially Excluded Employees Prefer Algorithmic Evaluation to Human Assessment: The Moderating Role of an Interdependent Culture," *Computers in Human Behavior: Artificial Humans*, Vol. 4, 100152.

すぎたに・ようこ 上智大学経済学部経営学科教授。主に『新・消費者理解のための心理学』（共著、福村出版、2025年）など。消費者心理学、マーケティング論専攻。