



コラム連載:「ビジネスと人権」:企業の「なぜ?どうして?」に答える第6回

生成 AI を使い始めたが、人権リスクが心配… 必要な社内ルールは?

株式会社オウルズコンサルティンググループ パートナー
矢守 亜夕美

※2026年2月5日付のサステナブル・ブランド ジャパン(SB-J)の記事を一部変更して掲載しています。

昨今、目覚ましい進化を遂げている生成 AI 技術ですが、業務効率化のために活用を推進する企業も増えて
います。そんな中で、現場からは「生成 AI をどんどん使いたい」との要望がある一方、リスク管理部門からは「情
報漏洩(えい)や人権リスクが懸念されるので、厳しいルールを設けたい」との声がある——企業のサステナビリ
ティ・人権担当者から、こうした「板挟み」の悩みを耳にするようになってきました。「人権の観点からはどう判断
すべきか」と双方から問われ、悩んでいる方も少なくありません。

[以前の連載\(前シリーズの第7回\)](#)でも触れた通り、AI 活用に当たっては人権リスクに注意する必要があり、

中でも特に生成 AI にはさまざまな懸念があります。ですが、ここまで社会に浸透しつつある技術を「危ないから」と禁止し続けるのは現実的でないことも確かです。今回は、「生成 AI を業務で使い始めているが、人権リスクが心配…。どんな社内ルールが必要？」にお答えします。

では、生成 AI の業務活用に際して、人権リスクを踏まえて社内ルールを検討するための視点を整理してみましょう。ポイントは①用途ごとの線引きを検討する、②入力・出力の基本ルールを定める、③運用しながら常に更新していく、の3つです。

「野放し」にしないため、まずはリスクの所在を知る

社内ルールを考える前に、「そもそも何を守らなければならないのか」を押さえておく必要があります。

実務上、特に警戒すべきなのは、AI が特定の属性に対して不公平な判断を下してしまう「差別・バイアス」の問題や、従業員・顧客の機微な情報が AI を通じて漏洩してしまう「プライバシー」の問題などです。AI で自動的に生成された表現が差別やバイアスを含んでいた場合、マイノリティや社会的に弱い立場の人々を深く傷つけてしまう恐れもあります。

こうしたリスクがあるからこそ、何も考えずに「野放し」にするのは危険です。ただし、リスクの大きさは「何に使うか」によって大きく変わります。全てを一律に禁じるのではなく、用途に応じてメリハリをつけることが、現実的なルールづくりの第一歩です。

「何に使うか」でリスクの度合いを線引きする

最初に取り組みたいのが、社内で想定される用途を整理し、「高・中・低」といったリスクレベルに仕分ける作業です。

最も慎重になるべきは、人の人生や権利に直接影響を与える「高リスク領域」です。例えば採用選考や人事評価などで AI がバイアスを含んだ判断を下してしまうと、取り返しのつかない人権侵害につながりかねません。そのためこの領域では原則として利用を控えるか、目的を限定して、人間による厳しい監視下でのみ利用するといった判断が必要になります。

「中リスク領域」としては、例えば広告コピーの作成や自社 Web サイトの記事執筆などが挙げられます。ここでは、誤情報や差別的な表現がそのまま世に出てしまわないよう、必ず人間が内容を確認することを条件とするのが望ましいでしょう。また、必要に応じて「AI で生成した画像・文章である」と明記するなどの対応(透明性の確保)も重要です。一方で、例えば社内向けのアイデア出しや議事録の要約といった領域であれば、人権侵害につながる可能性は相対的には低くなります。基本的なルールさえ守れば、積極的に活用できるゾーンと言えるはずです。

このように「ここから先は注意が必要だが、ここまでは自由に活用してよい」と明確に線を引くことで、現場側も納得してルールを守りやすくなります。

「入力」と「出力」のルールを定め、「どこまでは OK か」を示す

用途の仕分けができたら、次は現場が守るべき具体的な利用ルールの検討です。ここは「入力」と「出力」の 2 つの側面から考えると整理しやすくなります。

入力サイドの鉄則は、個人情報や機微情報を安易に入力しないことです。例えば従業員の氏名・住所や健康情報、人事評価などのデータは、原則として入力しないと決めておくべきでしょう。どうしても業務上必要な場合は、AI の学習に利用されない環境に限定する、上長や情報システム部門の承認を必須にする、といった規定を設けることが考えられます。

出力サイドで何より重要なのは、AI が出した答えをそのまま最終的な意思決定や対外発信に使わない、という点です。特に人の評価や選別につながり得る場面、あるいは社外に情報を出す場面では、必ず人間が内容を読み込み、差別的な判断・表現や事実誤認がないか確認するプロセスを経る必要があります。「AI はあくまで『壁打ち相手』であり、最後に責任を持つのは人間である」という原則を徹底することが、最大のリスクヘッジになります。

ルール作りに際して、併せて意識したいのが、「禁止事項の羅列」だけにしないことです。「これは NG」というルールと同時に、「こうした条件を満たせば利用してよい」という「条件付き OK」のゾーンを示すことで、現場も安心して活用しやすくなります。例えば、「社内向け資料のドラフト作成には使ってよい。ただし機微な情報は入力せず、必ず作成者本人が内容を確認してから配布する」といったイメージです。

最初から完璧を目指さず、常に更新し続ける姿勢こそが重要

心にとどめておきたいのは、社内ルールは一度作れば終わりではないということです。生成 AI の技術進化は目覚ましく、数カ月で状況が一変することも珍しくありません。今日決めたルールが、半年後には時代遅れになっている可能性も十分にあります。だからこそ、最初から完璧なルールを目指す必要はありません。まずは試験的に運用し、現場でヒヤリとした事例などを吸い上げながら、常に柔軟に書き換えていく姿勢が大切です。

人権対応は、決して企業の活動を縛るブレーキではありません。事故を未然に防ぎ、安心してアクセルを踏み込むための安全装置と捉えるべきです。生成 AI の活用でも、「禁止」か「野放し」かではなく、用途ごとの線引きと入力・出力の作法を定め、運用しながら常に更新していく——その積み重ねが現実的な解になります。

まずは自社で「どこからなら安全に始められるか」を言語化すること。その一歩が、技術と人権尊重が両立する未来へのスタートラインになるはずです。

出典：[サステナブル・ブランド ジャパン \(SB-J\)](#)

著者



株式会社オウルズコンサルティンググループ
執行役員／パートナー
矢守 亜夕美／Yamori, Ayumi

A.T. カーニー(戦略コンサルティング)、Google、スタートアップ等を経て現職。東京大学法学部(公法コース)卒。

「ビジネスと人権」チームリードを務めるとともに、サステナビリティ・ルール形成・地政学関連など多岐にわたるプロジェクトを担当。

共著に『すべての企業人のためのビジネスと人権入門』『ビジネスと地政学・経済安全保障』(日経 BP 社)がある他、日本経済新聞、読売新聞、テレビ朝日など、人権・社会課題に関するコメント掲載およびメディア出演実績多数。

経済産業省／中小企業庁「ビジネスと人権」セミナー講師(2021年)、東京都人権プラザ主催「サステナビリティと人権」セミナー講師(2022年)等、登壇実績多数。

労働・人権分野の国際規格「SA8000」基礎監査人コース修了。

ISO30414(人的資本情報開示)リードコンサルタント/アセッサー。

本資料は一般的な情報提供を目的とするものであり、その性質上、特定の個人や事業体に具体的に適用される個別の事情に対応するものではありません。関連する法令等の解釈を行ったものではなく、利用者が本資料を利用したことによる結果について、株式会社オウルズコンサルティンググループは一切の責任を負うものではありません。

また、書面による株式会社オウルズコンサルティンググループの事前承認なしに、第三者への配布・引用・複製を行うことはお断りしております。

株式会社オウルズコンサルティンググループ

〒106-0032 東京都港区六本木 1-4-5

アークヒルズサウスタワー16F

<https://www.owls-cg.com/>