ハマの労務コンサル短信

〈No.32〉新しい相棒 ChatGPT

〈No.25〉低落続く日本の労働生産性
〈No.26〉改正育児・介護休業法「個別の周知」「意向の確認」の義務付けについて
〈No.27〉改めて労働者代表の選出方法を考える
〈No.28〉改めて賃金規定を考える
〈No.29〉改めて産業医を考える
〈No.30〉規制改革推進会議の答申について
〈No.31〉改めて試用期間について考える

ピー・エム・ピー株式会社　代表取締役　鈴木 雅一

✳“それっぽいウソ”をつく

「AIのおかげで労務コンサルの出番が減りましたか？」。先日，こんな質問を顧問先から受けました。かかる実感はありませんが，AIが代替できるのであれば，労務コンサルは退場すべきでしょう。もっとも筆者は仕事で毎日のようにAIを利用していますが，まだ100％の信頼を置くには至りません。誤った情報を平気で返してくるからです。

そこで，AIに「なぜAIは嘘をつくことがあるのですか？」と質問してみました。以下は，筆者が利用しているChatGPT 4 拡張版の返答です。

AI（たとえば私，ChatGPT）が『嘘のような回答をしてしまう理由』は，主に次のような理由によります。

1.『訓練データに誤情報が含まれている』：AIは大量のインターネット上の文章を学習しています。その中には正しい情報だけでなく，誤情報や古い情報，デマなども含まれています。そのため，学習段階で誤った内容を覚えてしまうことがあります。

2.『「もっともらしい文章」を作ることが目的』：AIは「真実を言うこと」ではなく，「入力に対してもっともらしい出力を生成すること」を目的に設計されています。そのため，「情報が曖昧なとき」「データが存在しないとき」「確信がないとき」にも，何かを答えようとして，結果的に“それっぽいウソ”を出してしまうことがあります。

3.『質問の意図を誤解することがある』：AIは自然言語を理解する能力はありますが，人間と同じように「背景」や「文脈」「ニュアンス」を読み間違えることがあります。その結果，質問の主旨とはズレた答えになり，誤解を招く場合があります。（こちらの“質問力”も関係します）

4.『情報が古い』：このAIモデルはある時点までの情報しか知りません。そのため，「最近の出来事」「最新の研究結果」については回答できず，古い情報に基づいた回答が結果的に間違っているということもあります。

補足.『「嘘をつく」のではなく「間違える」』：AIが「嘘をついているように見える」と感じるときでも，実際には人間のように意図的に事実を隠したりごまかしたりしているわけではありません。AIは感情も意図も持たないため，誤答＝ウソと見えるだけなのです。

✳誤答の例を尋ねると…

次に，AIによる嘘＝誤答の例を尋ねてみました。例示された中でいくつか紹介しましょう。

【例1】存在しない論文を紹介した

質問：「ChatGPT，AIによる医療診断の精度に関する論文を教えて」

AIの誤答：「“Deep Learning for Medical Diagnosis” by John Smith et al., published in Nature Medicine, 2021 という論文があります」

問題点：実際にはこの論文も著者も存在しません。もっともらしいタイトルや雑誌名をAIが作り出してしまったケースです。これは「ハルシネーション（幻覚）」と呼ばれます。

●鈴木 雅一（すずき まさかず） www.pmp.co.jp
■本社所在地：〒220-0012 神奈川県横浜市西区みなとみらい2-2-1 ランドマークプラザ5階
ピー・エム・ピー株式会社（PMP）代表取締役・特定社会保険労務士。慶應義塾大学経済学部を卒業（専攻は経済政策，恩師はカトカンで有名な加藤寛教授）。三菱銀行（現三菱UFJ銀行）に入社し，人事企画部門他を経験。その後，米国ケミカル銀行（現JPモルガン・チェース銀行）の日本支店の副社長として銀行と証券人事部門を統括。米国マイクロソフト社の日本法人であるマイクロソフト株式会社の人事部門と総務部門の統括責任者を経て，PMPを創業。また，人事部長のネットワークSMCも運営，人事スペシャリストの研鑽の場を提供している。著書に『アメリカ企業には就業規則がない』（国書刊行会）。

【例２】歴史的事実を誤る

質問：「ナポレオンはいつロシアに侵攻しましたか？」

AIの誤答：「ナポレオンは1821年にロシア遠征を開始しました」

正しくは：1812年です。

原因：年代の記憶や関連する出来事の混同によって誤った数字を生成しています。

【例３】計算ミス

質問：「年利５％で100万円を10年間複利で運用したらいくらになりますか？」

AIの誤答：「約150万円になります」

正しくは：約162.89万円（1,000,000×（１＋0.05）^10）。

原因：複利計算の式を省略・近似して誤った答えを出したりすることがあります。

これらはあくまで一例ですが，『AIは「もっともらしい情報」を出す訓練を受けているため，間違いがあっても自信たっぷりに語ることがあるのが厄介な点です』と自ら言っています。

✳労働法関連の誤答例

続いて，「労働法に絞って誤答例を教えてください」と尋ねました。以下は返答例です。

【例１】変形労働時間制の誤解

質問：「１年単位の変形労働時間制では，１日何時間まで働かせてもいいのですか？」

AIの誤答：「１年単位の変形労働時間制では，労働時間の制限はありません」

正しくは：１年単位の変形労働時間制でも，１日の上限は10時間（労基法施行規則12条の２）。

【例２】有給休暇の買い取りに関する誤答

質問：「使いきれなかった有給休暇は買い取ってもらえますか？」

AIの誤答：「はい，使わなかった有給休暇はすべて会社に買い取ってもらえます」

正しくは：原則として年次有給休暇の買い取りは禁止です（労基法39条）。ただし，退職時の未消化分など一部の例外的な場面では可能です。AIは例外を一般ルールのように答えてしまうことがあります。

【例３】試用期間中の解雇

質問：「試用期間中ならいつでも自由に解雇できますか？」

AIの誤答：「はい，試用期間中は解雇自由なので，理由がなくても解雇できます」

正しくは：試用期間中でも「解雇権濫用法理（労契法16条）」が適用されるため，客観的合理性・社会的相当性が必要です。「自由解雇」と誤認している点が重大な誤答です。

✳公表の注意点を踏まえて

最後に，ChatGPTに「このような誤答についての情報を公表してもかまわないか？」と尋ねてみました。『公表にあたっての注意点』が示されました。

1.『AIの発言が「事実ではない」ことを明記してください』：AI（たとえば私，ChatGPT）が提供する情報は必ずしも正確とは限らず，「誤答」や「架空の情報」を含むことがあります。公表の際は「これはAIが実際に出した誤答の例である」「誤解を招く回答である」など，AIが真実として保証しているわけではないことを明示してください。

2.『誤解を招かないよう，出典や補足説明を付けてください』

3.『OpenAIやChatGPTという名前を出す場合は文脈を適切に』：AIに誤答があることは周知の事実ですが，「OpenAIのAIは使えない」などといった誤解や風評被害につながる表現は避けてください。「あくまでAIは万能ではなく，参考にとどめるべき」などの中立的スタンスが望ましいです。

そして 4.『個別の利用規約にも注意してください』とありましたので，最後に「本文はChatGPT（OpenAI製AIモデル）による出力を，人が確認・編集のうえ公表したものであり，生成内容が正確であることを保証するものではありません」と付記することにします。

AIはもはや，新たな良き相棒という存在ですし，日々AIの間違いを指摘することで次は間違わない返答ができるのではないか，と思って付き合っています。