

2023年12月11日 全24頁

# 生成 AI が日本の労働市場に与える影響②

## 就業者の約80%が生成 AI の影響を受ける可能性

経済調査部 主任研究員 新田堯之

### [要約]

- 前回レポートでは、生成 AI が労働市場に与えるメカニズムの整理と米国の研究や事例からの示唆についてまとめた。本レポートはそれに続く第二弾として、生成 AI が日本の労働市場に与える影響を定量的かつ包括的に分析した。具体的には、職業ごとにタスクのうち生成 AI が自動化可能な割合を示す「自動化対象率」を推計した。その結果、日本の就業者の約80%が何らかの形で生成 AI の影響を受ける可能性があり、同様に約40%の就業者が仕事の半分以上を自動化できると推計された。
- 職業別では、事務従事者や管理的職業従事者、専門的・技術的職業従事者などホワイトカラーの職種が、自動化の影響を強く受けるとの結果が得られた。他方で、運搬・清掃・包装等従事者や建設・採掘従事者などブルーカラーに分類される職業などへの影響は低かった。賃金別に見ると、高賃金の職種ほど自動化対象率が高い傾向が見られた。
- 産業別では、情報通信業や金融業・保険業、不動産業・物品賃貸業など情報やデータを扱うオフィスワーク中心の産業で自動化対象率が比較的高かった。その反面、身体的な動作を伴う場面が多い建設業や漁業、宿泊業・飲食サービス業などは比較的低い水準であった。
- これら一連の結果は、米国を対象にした先行研究と全体的な傾向で一致していた。一方、タスクの自動化対象率が50%を超えた割合は、本レポートの方が先行研究よりも高かった。この背景の一つとして、産業構成の違いなどに加えて、日本は米国に比べてルーティン業務をより多く抱えている可能性が指摘できる。
- 次回の第三弾のレポートでは、今回議論した自動化対象率が高い職業の性質や特徴に関して定量的に分析する。さらに、生成 AI が雇用の代替および補充に及ぼす影響に関してより踏み込んだ議論を展開した上で、リスクリング等に関する政策を提案する予定である。

## 1. はじめに

前回のレポート（生成 AI が日本の雇用に与える影響①）では、ChatGPT を初めとする生成 AI が労働市場に与える影響に関して、そのメカニズムや日本の立ち位置、米国の研究や事例を取り上げつつ整理した。それに続く本レポートは、生成 AI が日本の労働市場に与える影響を定量的かつ包括的に分析した。

具体的には、職業ごとにタスクのうち生成 AI が自動化可能な割合を示す「自動化対象率」を推計した。推計する上では、OpenAI 社が開発した大規模言語モデル（Large Language Model, LLM）GPT-4 を活用した。

分析結果の概要を述べると、総就業者数の約 80% が何らかの形で生成 AI の影響を受ける可能性があり、同様に 40% の就業者が仕事の半分以上を自動化できると推計された。

職業別では、事務従事者や管理的職業従事者、専門的・技術的職業従事者などホワイトカラーの職種を中心に、自動化の影響を強く受けるとの結果が得られた。他方で、運搬・清掃・包装等従事者や建設・採掘従事者などブルーカラーに分類される職業などに対しては、生成 AI が及ぼす影響は低かった。賃金別に見ると、高賃金の職種ほど自動化対象率が高い傾向が見られた。

産業別では、情報通信業や金融業・保険業、不動産業・物品賃貸業などの情報やデータを扱うオフィスワークが中心の産業で自動化対象率が比較的高かった。その反面、身体的な動作を伴う場面が多い建設業や漁業、宿泊業・飲食サービス業などは比較的低い水準であった。また、地域別に見ると、IT 産業や金融業が集中する都市部で生成 AI の影響が高め、地方では低めの傾向が見られた。

これら一連の結果は、米国を対象にした先行研究と全体的な傾向で一致していた。一方、タスクの自動化対象率が 50% を超えた割合は本レポートの方が先行研究よりも高かった。この背景の一つとして、産業構成の違いなどに加え、日本は米国に比べてルーティン業務をより多く抱えている可能性が指摘できる。

筆者の認識では、生成 AI が日本の雇用に与える影響を議論した日本語文献は、特定の職業や産業を取り上げて定性的に議論するケースや、米国など他国の研究を援用して議論するケースがほとんどである。こうした中、本レポートは生成 AI が日本の雇用に与える影響に関し、定量的な分析を実施した数少ない事例として意義があると考えられる。

以降の構成を簡潔に述べる。**2. 職業ごとの自動化対象率の試算方法**では、自動化対象率の試算に用いたデータや手法について解説している。分析結果のみに関心がある場合は、9 ページからの**3. 分析結果**に直接進んでも問題ない。**3. 分析結果**では職業別、年収別、産業別、都道府県別に自動化対象率を集計してその特徴を述べると共に、結果を解釈する際の様々な留意点にも言及している。最後の**4. まとめと今後に向けて**では、全体の結果をまとめつつ、また、次回レポートの内容にも触れている。

## 2. 職業ごとの自動化対象率の試算方法

### 2-1. 職業情報提供サイト（日本版 O-NET）の解説

本レポートで使用した職業情報提供サイト（[日本版 O-NET<愛称：job tag>](#)、以下、日本版 O-NET と表記）とは、500 以上の職業情報を提供するウェブサイトである。米国労働省が運営する職業情報ネットワーク O\*NET および職業情報サイト O\*NET OnLine を参考に、日本における労働市場の透明性を高め、主体的なキャリア形成を支える市場インフラの整備を目指し、2020 年に開設された<sup>1</sup>。

日本版 O-NET は「ジョブ」（職業や仕事）、「タスク」（仕事の内容や作業の細分化）、および「スキル」（仕事に必要な技術や技能）などに関する情報を集約し、可視化することによって、求職者や企業の採用活動を支援する目的で設計された。さらに、職業ごとの能力、必要な知識、向いている興味や価値観を数値化したデータなど、利用者が自身に適した職業選択を行うための情報を提供している。

例えば、保育士のページ（**図表 1**）を見ると、タスクとして、「基本的な日常生活習慣を子どもに教える。」「遊びを通して人との関わり、言語能力、運動能力、知的能力を育てる。」等の情報が記述されている。

図表 1：日本版 O-NET における保育士のタスクの説明例

タスク（職業に含まれるこまかな仕事）	仕事の内容
保育所や託児施設で保護者に代わって小学校に上がる前の子どもたちの世話をする。	
実施率	タスク内容
89.7%	基本的な日常生活習慣を子供に教える。
88.5%	遊びを通して人との関わり、言語能力、運動能力、知的能力を育てる。
87.2%	一人ひとりの発達状況を把握し、健やかな成長のために適切な介助や指導をする。
84.6%	衣服の着脱、玩具や本の整理などの身辺整理ができるように指導する。
82.1%	子供のための食事の準備をし、食事の介助やしつけをする。
79.5%	子どもを観察して保育日誌や連絡ノートをつけ、保護者へその日の活動や様子を伝える。
76.9%	子供たちの発達段階に応じた幼児教育をする。
73.1%	保育している乳幼児の健康状態を見守り、病気や事故の場合は保護者や医療機関と連絡を取る。
62.8%	子供を入浴させたり、プールで水遊びをさせたりする。
52.6%	ゲームなどのレクリエーション活動を計画して運営する。

（出所）職業情報提供サイト（日本版 O-NET）の保育士のページより引用。  
 （URL：<https://shigoto.mhlw.go.jp/User/Occupation/Detail/131>）

<sup>1</sup> 日本政府の「未来投資戦略 2017」（平成 29 年 6 月 9 日閣議決定）、「未来投資戦略 2018」（平成 30 年 6 月 15 日閣議決定）、および「成長戦略フォローアップ」（令和元年 6 月 21 日閣議決定）等に基づいて開設された。

日本版 O-NET のウェブサイトには、「職業情報データダウンロード」というページがある。ここでは、各職業に関する情報がデータ分析に利用しやすいエクセルファイルや CSV ファイルの形式にまとめられている。本レポートでは、当該ページからダウンロードしたファイル<sup>2</sup>を用いて分析を進めた。

## 2-2. 職業ごとの自動化対象率の計算に GPT-4 を使用

### 自動化されやすいタスクを GPT-4 により分類

本レポートでは、ChatGPT を開発した OpenAI 社とペンシルベニア大学の研究 (Eloundou et al. (2023)) の先行研究を参考に、各職業についてタスクベースの自動化対象率を試算した。

初めに、日本版 O-NET に含まれる全 521 職業<sup>3</sup>のうち、タスクが取得できる 456 職業を抽出した (タスクが取得できない 65 職業への対応は**補論 1**を参照)。その上で、各職業のタスクごとに、影響を受ける度合いを 3 カテゴリーに分類した (分類基準は後述)。

この分類時には、従来の研究でよく行われていた人手による分類ではなく、OpenAI 社が開発した大規模言語モデル GPT-4 を活用した<sup>4</sup>。

一つ目の理由は、客観性を担保するためである。人手による分類には、分類した人物の知識や経験、思い込みなどから生じるバイアスが含まれるリスクがある。それゆえ GPT-4 などの AI を用いれば、こうしたリスクを一定程度回避できると見込まれる。

二つ目の理由は、GPT-4 の性能の高さである。第一弾のレポートで述べた通り、GPT-4 は従来よりも多くのタスクで人間と同等以上の結果を収めている。さらに、Fabrizio et al. (2023)によれば、GPT-4 より性能が低い GPT-3.5 ですら、テキスト分類タスクで人手に頼るクラウドワーカー<sup>5</sup>より精度が高かったという。従って、今回の職業に関するタスクを分類する作業でも、GPT-4 は高い精度が期待できる。

三つ目の理由は、効率性の問題である。例えば、各職業のタスクが平均 10 個と仮定すると、タスクの合計は 456 職業×10 タスク=4,560 タスクに達する。これを 1 人で分類すると膨大な手間がかかるため、効率化を図った。

<sup>2</sup> 本レポートでは、独立行政法人労働政策研究・研修機構 (JILPT) 作成「職業情報データベース 解説系ダウンロードデータ ver4.00」および「職業情報データベース 簡易版数値系ダウンロードデータ ver4.00」職業情報提供サイト (日本版 O-NET) より 2023 年 10 月 31 日にダウンロード、を使用した。

<sup>3</sup> 正確には、独立行政法人労働政策研究・研修機構 (JILPT) 作成「職業情報データベース 解説系ダウンロードデータ ver4.00」職業情報提供サイト (日本版 O-NET) に含まれる 521 職業。

<sup>4</sup> Eloundou et al. (2023) では、①GPT-4 による分類と、②GPT-3、GPT-3.5、GPT-4 の出力をレビューした人間のアノテーター (テキストや画像などにラベル付けをする職種) による分類の両方を行った。

<sup>5</sup> 不特定の人 (クラウド=群衆) に業務を外部委託 (アウトソーシング) する「クラウドソーシング」を通じて働く人々を指す。

## タスクの分類基準

分類基準は、Eloundou et al. (2023)の基準 (exposure rubric) と同じものを用いた。すなわち、①E0: エクスポージャーなし (大規模言語モデル (LLM) の使用がタスクの効率を向上させず、むしろ品質を低下させる場合)、②E1: 直接的なエクスポージャー (ChatGPTのような大規模言語モデルを用いてタスクを少なくとも半分の時間で完了できる場合)、③E2: LLM + エクスポージャー (大規模言語モデルに加えて、最新情報を提供するシステムや画像生成 AI などの追加ソフトウェアを組み合わせることで、タスクを高品質かつ効率的に完了できる場合)、である (図表 2)<sup>6</sup>。

GPT-4 に要求を伝えるプロンプト文には、上記の分類基準の文言に若干追記した上で使用した。具体的には、プロンプト文に「あなたは日本で働いていると仮定してください (実際は英語)」という一文を加えた。これは、日本の事情を踏まえた回答を期待したためである。さらに出力形式をデータ分析がしやすい形に指定するなどフォーマットに関する指示を加えた。なお、プロンプト文に用いた言語は、基本的には英語であったが、各職業のタスクのみ日本語の原文をそのまま入力した。

図表 2 : 本レポートにおけるタスクの分類基準の概要

カテゴリ	カテゴリのイメージ	カテゴリに判定される条件
E0: エクスポージャーなし	当該タスクがChatGPTで自動化不可能	<ul style="list-style-type: none"> <li>LLM (大規模言語モデル) の使用により、同等の品質を維持しながら、活動や作業を完了するのに必要な時間がほとんどまたは全く短縮されない場合</li> <li>LLMの使用によって、アウトプットの質が低下する場合</li> </ul>
E1: 直接的なエクスポージャー	当該タスクがChatGPTで自動化可能	<ul style="list-style-type: none"> <li>ChatGPTなど大規模言語モデル (LLM) を活用し、タスクを完了するのに必要な時間を少なくとも半分 (50%) 減少させることができる場合</li> <li>LLMの単独使用では、タスクを完了するのに必要な時間を少なくとも半分減らすことはできない場合</li> </ul>
E2: LLM + エクスポージャー	当該タスクがChatGPTと①最新情報にアクセスできる「Browse with Bing」や②画像生成AI「DALL・E 3」などを組み合わせて自動化可能	<ul style="list-style-type: none"> <li>一方、LLMに追加のソフトウェアを開発することで、特定のタスクを少なくとも半分の時間で高品質に完了させることが可能である場合。</li> <li>インターネットから最新情報を入手するシステムや画像生成システムへのアクセスも含まれる</li> </ul>

(注 1) カテゴリのイメージは筆者の補足。

(注 2) より詳細な基準は Eloundou et al. (2023) の “A.1 Exposure” (pp. 24-26) を参照。

(出所) Eloundou et al. (2023) を参考に大和総研作成

## GPT-4 によるタスク分類結果の一例：専門学校教員の場合

このような方法で、上記の 456 職業のタスクを分類した。分類結果の一例として、ここでは専門学校教員のタスクを取り上げる。図表 3 の「タスクの説明」とは、GPT-4 のプロンプトに記述した各タスクの説明文である。「GPT-4 による判定」とは、GPT-4 から出力されたカテゴリ分類であり、図表 2 で解説した E0、E1、E2 の 3 種類のいずれかに分類される。「GPT-4 による判定理由」は、当該判定を選択した理由の文章<sup>7</sup>である。

<sup>6</sup> より詳細な基準は Eloundou et al. (2023) の “A.1 Exposure” (pp. 24-26) を参照。この論文では、タスクを分類する際に、画像認識や生成に関するタスクは「E3」というラベルで分類されている。しかし、分析を行う際には、「E2」と「E3」のタスクを一緒にして、「E2」というラベルで議論している。本レポートでも、同じ方法を踏襲している。

<sup>7</sup> 判断理由の原文は英語で出力されたが、わかりやすさの観点から大和総研が日本語に翻訳した。

例えば、専門学校教員（**図表 3**）のタスク 2「教育プログラムや研修スケジュールについて枠組みを準備し、講座の目標を策定する。」に対し、GPT-4 は当該タスクが大規模言語モデルで自動化可能だとする E1 に分類した。その判定理由としては、「LLM（大規模言語モデル）は教育プログラムとトレーニングスケジュールの枠組みの準備、およびコース目標の設定を補助し、タスクの完了にかかる時間を少なくとも半分に短縮できる。」と説明している。

他方、タスク 10「学生の知識と能力を向上させるため、視聴覚材を活用して講義を行い、ディスカッションをさせる。」に関しては、GPT-4 は当該タスクが大規模言語モデルで自動化不可能とする E0 と分類した。その判定理由として、「このタスクは視聴覚資料を使った講義の実施、学生の知識とスキルの向上、討論の促進を含むものであり、人間の相互作用を必要とし、LLM（大規模言語モデル）による実行は不可能である。」という文章が出力された。

**図表 3 : タスクの分類結果の例（専門学校教員）**

<p>■タスク 1</p> <ul style="list-style-type: none"> <li>・タスクの説明：講座の内容と指導方法の計画を立てる。</li> <li>・GPT-4による判定：E1</li> <li>・GPT-4による判定理由：LLMはコースの内容や教授方法の計画を補助し、タスクの完了にかかる時間を少なくとも半分に短縮できる。LLMは、既存の知識に基づいた提案やアイデアを提供できる。</li> </ul>	<p>■タスク 2</p> <ul style="list-style-type: none"> <li>・タスクの説明：教育プログラムや研修スケジュールについて枠組みを準備し、講座の目標を策定する。</li> <li>・GPT-4による判定：E1</li> <li>・GPT-4による判定理由：LLMは教育プログラムとトレーニングスケジュールの枠組みの準備、およびコース目標の設定を補助し、タスクの完了にかかる時間を少なくとも半分に短縮できる。</li> </ul>
<p>■タスク 3</p> <ul style="list-style-type: none"> <li>・タスクの説明：実地研修あるいは実習を行い、担当科目の理論、技術、手順、方法を示す。</li> <li>・GPT-4による判定：E0</li> <li>・GPT-4による判定理由：このタスクは実践的な訓練や実習の実施、科目の理論や技術、手順、方法の実演を含むものであるため、人間の相互作用を必要とし、LLMによる実行は不可能である。</li> </ul>	<p>■タスク 4</p> <ul style="list-style-type: none"> <li>・タスクの説明：講義用に書籍、資料、備品や設備を選択し、準備する。</li> <li>・GPT-4による判定：E1</li> <li>・GPT-4による判定理由：LLMは講義用の書籍、資料、機器、施設の選定と準備を補助し、タスクの完了にかかる時間を少なくとも半分に短縮できる。</li> </ul>
<p>■タスク 5</p> <ul style="list-style-type: none"> <li>・タスクの説明：会議、セミナー、研修講座に参加し、研修プログラムに関連する情報を織り込む。</li> <li>・GPT-4による判定：E1</li> <li>・GPT-4による判定理由：LLMは会議、セミナー、研修コースへの参加を補助し、トレーニングプログラム関連の情報を取り入れることで、タスクの完了にかかる時間を少なくとも半分に短縮できる。</li> </ul>	<p>■タスク 6</p> <ul style="list-style-type: none"> <li>・タスクの説明：出席簿や成績表を保管し、学校の規定に従って報告書を作成する。</li> <li>・GPT-4による判定：E1</li> <li>・GPT-4による判定理由：LLMは出席簿と成績表の管理、学校規則に基づく報告書の作成を補助し、タスクの完了にかかる時間を少なくとも半分に短縮できる。</li> </ul>
<p>■タスク 7</p> <ul style="list-style-type: none"> <li>・タスクの説明：課題について添削、評価、コメントをする。</li> <li>・GPT-4による判定：E1</li> <li>・GPT-4による判定理由：LLMは課題の訂正、評価、コメントを補助し、タスクの完了にかかる時間を少なくとも半分に短縮できる。</li> </ul>	<p>■タスク 8</p> <ul style="list-style-type: none"> <li>・タスクの説明：口頭、筆記、実技試験を行い、進捗状況を把握し、学習の効果を評価する。</li> <li>・GPT-4による判定：E0</li> <li>・GPT-4による判定理由：このタスクは口頭、筆記、実技テストの実施、進捗の理解、学習効果の評価を含むものであり、人間の相互作用を必要とし、LLMによる実行は不可能である。</li> </ul>
<p>■タスク 9</p> <ul style="list-style-type: none"> <li>・タスクの説明：学生の学習内容を観察・評価し、進捗状況を判定し、意見を述べ、改善すべき点を提案する。</li> <li>・GPT-4による判定：E1</li> <li>・GPT-4による判定理由：LLMは学生の学習内容の観察と評価、進捗の判定、意見の表明、改善の提案を補助し、タスクの完了にかかる時間を少なくとも半分に短縮できる。</li> </ul>	<p>■タスク 10</p> <ul style="list-style-type: none"> <li>・タスクの説明：学生の知識と能力を向上させるため、視聴覚材を活用して講義を行い、ディスカッションをさせる。</li> <li>・GPT-4による判定：E0</li> <li>・GPT-4による判定理由：このタスクは視聴覚資料を使った講義の実施、学生の知識とスキルの向上、討論の促進を含むものであり、人間の相互作用を必要とし、LLMによる実行は不可能である。</li> </ul>

（注） 判断理由の文章の原文は英語で出力されたが、わかりやすさの観点から日本語に翻訳した（翻訳は大和総研）。

（出所） 独立行政法人労働政策研究・研修機構（JILPT）職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、GPT-4（OpenAI）より大和総研作成（分類はEloundou et al. (2023)を参考にした）

一連のタスク分類作業は各職業に対して3回実施した。その上で、最終結果には最も多く出力されたカテゴリを採用した<sup>8</sup>。これは、GPT-4は同一のプロンプトに対し、同一の回答を常に出力しなかったという事情を踏まえたものである<sup>9, 10</sup>。

## 職業ごとの自動化対象率の計算方法

こうした問題を考慮して、職業ごとに自動化対象率を以下の通りに計算した（**図表4**）。

ステップ1：タスクごとに3回ずつGPT-4によるカテゴリの判定を行う。

ステップ2：タスクごとに最も多く判定されたカテゴリを最終結果として採用する。例えば、

**図表4**のタスク10では、1回目の判定はE1、2回目と3回目の判定はE0のため、最終結果は最も多いカテゴリであるE0とする。

ステップ3：カテゴリごとに分類されたタスクの割合を計算する。例えば、**図表4**のカテゴリE0の割合を計算する際には、最終的にE0と分類されたタスクの数は3個、タスク全体の数は10個のため、E0の割合は $3 \div 10 = 30\%$ と計算できる。日本版O-NETではタスクに対応した実施率が取得できる。日本版O-NETでは、実施率とは「一般的に、その職業に就いている人のうち、そのタスクを実施している人の割合」を指す<sup>11</sup>。なお、職業の実態をできるだけ反映するため、この後の分析では、タスクの実施率で加重平均した値を用いた。

<sup>8</sup> 万が一、3回とも異なるカテゴリが出力された場合は、保守的にE0を最終結果として優先した。

<sup>9</sup> GPT-4をAPI（アプリケーションプログラミングインターフェイス）から呼び出し、temperature（温度）というパラメータをより一貫性のある結果が得られるように対応した。このtemperatureは0から2までの範囲で設定でき、2に近づくにつれランダム性が増えるため、本レポートでは0に設定した。しかし、数学ソフト「Mathematica」の開発者でもあるスティーブン・ウルフラム氏によれば、temperatureを0に設定した場合でも、GPU（Graphics Processing Unit、画像処理装置）の計算が異なる順序で行われることと、ニューラルネットワークの計算時に丸め誤差（有効桁数以降の少数を切り捨てることによって生じる誤差）が原因で、同じプロンプト文およびGPUを用いても出力結果が異なる場合がある（脚注10参照）という。こうした事情を踏まえ、本レポートでは、バイアスがより小さくなるように、複数の結果に基づく多数決方式を採用した。

<sup>10</sup> STEPHAN WOLFRAM Writings “The New World of LLM Functions: Integrating LLM Technology into the Wolfram Language”（2023年5月23日）

（URL：<https://writings.stephenwolfram.com/2023/05/the-new-world-of-llm-functions-integrating-llm-technology-into-the-wolfram-language/>）

<sup>11</sup> 職業情報提供サイト（日本版O-NET）よくあるお問い合わせ

（URL：<https://shigoto.mhlw.go.jp/User/faq>）

図表 4 : 各職業の自動化対象率の計算例 (職業 : 専門学校教員)



タスク名	1回目の判定	2回目の判定	3回目の判定	最終結果 (最も多いカテゴリ)
タスク1	E1	E1	E1	E1
タスク2	E1	E1	E1	E1
タスク3	E0	E0	E0	E0
タスク4	E1	E1	E1	E1
タスク5	E2	E1	E1	E1
タスク6	E1	E1	E1	E1
タスク7	E1	E1	E1	E1
タスク8	E0	E0	E0	E0
タスク9	E1	E1	E1	E1
タスク10	E1	E0	E0	E0

E0の割合 =  $\frac{\text{判定がE0のタスク個数}}{\text{タスク全体の個数}} = \frac{3}{10} = 30\%$

E1の割合 =  $\frac{\text{判定がE1のタスク個数}}{\text{タスク全体の個数}} = \frac{7}{10} = 70\%$

E2の割合 =  $\frac{\text{判定がE2のタスク個数}}{\text{タスク全体の個数}} = \frac{0}{10} = 0\%$

(注) 実際には、タスクの実施率による加重平均を用いた。

(出所) 独立行政法人労働政策研究・研修機構 (JILPT) 職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」より大和総研作成 (イラストはいらすとや (<https://www.irasutoya.com/>) による)

これらのステップに従って、職業の自動化対象率を計算した。その上で、Eloundou et al. (2023)と同様に、各職業がどれほど自動化されやすいのかその範囲を示すため、職業全体の①E1 (下限)、②<sup>12</sup>E1+0.5×E2、③E1+E2 (上限) を計算した (要約統計量は図表 5 を参照)。本レポートのこれ以降の分析では、②E1+0.5×E2のみを自動化対象率として用いた。

図表 5 : 日本版 0-NET ベースの自動化対象率の要約統計量 (単位 : %)

	集計方法	平均値	標準偏差	最小値	下位25%点	中央値	上位25%点	最大値
	①E1	33.9	25.8	0.0	10.0	32.3	53.3	100.0
分析に使用	②E1 + 0.5×E2	37.9	27.1	0.0	13.4	37.1	58.5	100.0
	③E1 + E2	42.0	30.3	0.0	14.8	40.2	64.8	100.0

(注 1) 全て 521 サンプル。

(注 2) タスクの実施率による加重平均を用いた。

(出所) 独立行政法人労働政策研究・研修機構 (JILPT) 職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」より大和総研作成

### 就業者数や年収などを踏まえた議論には他統計とのデータ結合が必要

この方法で日本版 0-NET の職業ごとに生成 AI によるタスクの自動化対象率を計算した。しかし、日本版 0-NET には就業者数や年収などの情報は含まれていない。それゆえ、就業者数や年収などを踏まえた議論を行うためには、各職業の就業者数が取得できる総務省「国勢調査」や各職業の給料が取得できる厚生労働省「賃金構造基本統計調査」など、他の統計とデータを結合する

<sup>12</sup> Eloundou et al. (2023)では、この 0.5 というウェイトの数字の意図は、「(大規模言語モデルの追加ソフトウェア) 技術の展開に際し、補完的なツールやアプリケーションを用いることで追加投資が必要になることを考慮するため」、としている。

必要がある。

このプロセスにおける大きな問題は、各統計の職業カテゴリ名が一致していない点である。そこで、本レポートでは**補論2**の手順のように日本標準職業分類を介して、国勢調査や賃金構造基本統計調査の職業ベースの自動化対象率を計算した。

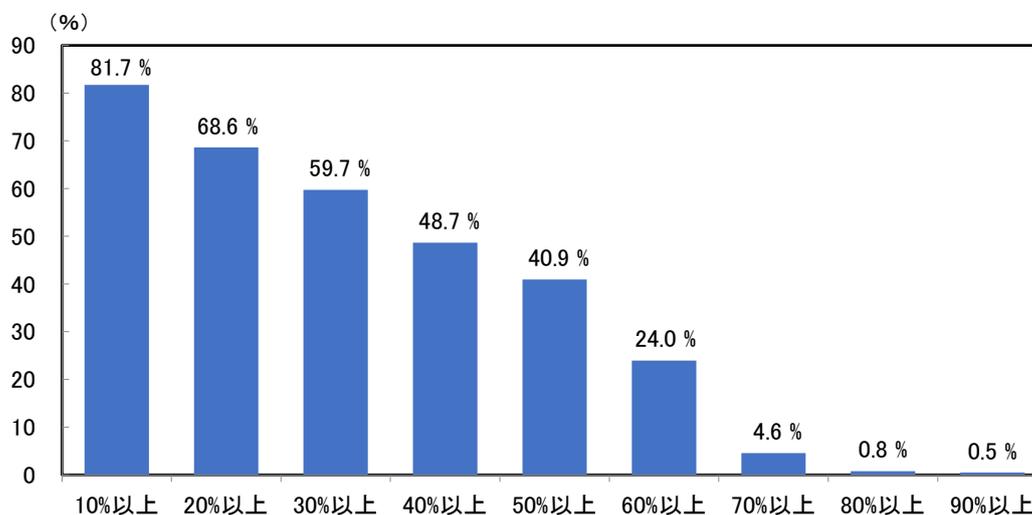
### 3. 分析結果

#### 3-1. 生成AIは多くの就業者に影響を与える結果に

本セクションでは、前章の方法を用いて得られた分析結果を解説する。

まず、令和2年の国勢調査によれば、同年の日本の総就業者数は5,767.4万人である。このうち、タスクの自動化対象率が10%以上の就業者数は4,712.3万人（総就業者数の81.7%）、50%以上の就業者数は2,361.7万人（同40.9%）と推定された（**図表6**）。これは、生成AIが大部分の就業者に対して何らかの形で影響を及ぼし、さらに4割程度の就業者が仕事のタスクの半分程度を生成AIで自動化できる可能性を示唆している。他方、タスクの自動化対象率が70%以上の就業者数は263.0万人（同4.6%）、80%以上は14.7万人（同0.8%）にとどまった。すなわち、タスクのほぼ全てが生成AIによる自動化の影響を受ける労働者はごく少数であった。

**図表6：タスクの自動化対象率が一定ラインを超えた就業者数の割合**



(注) 職業（小分類）ベースで加重平均して集計。宗教家や音楽家など自動化対象率のデータが欠落している職業は自動化対象率をゼロと置いた。

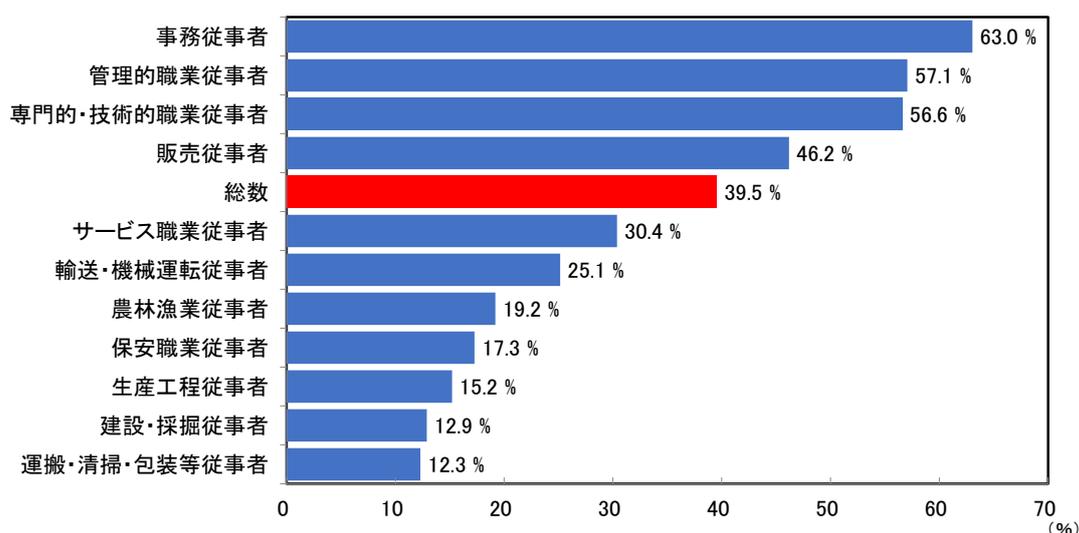
(出所) 総務省「令和2年国勢調査」、独立行政法人労働政策研究・研修機構（JILPT）職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、「解説系ダウンロードデータ ver4.00」より大和総研作成

#### 3-2. 職業別ではホワイトカラーへの影響が特に大きい

次に、職業（大分類）別に見ると、ホワイトカラーに分類される職業の自動化対象率が高い傾向が確認できた（**図表7**、中分類と小分類の詳細な結果は**補論3**を参照）。最も割合が高かった

職業は事務従事者の 63.0%であった。他には、管理的職業従事者（57.1%）や専門的・技術的職業従事者（56.6%）、販売従事者（46.2%）が全体（総数）の 39.5%を上回った。反対に、運搬・清掃・包装等従事者（12.3%）や建設・採掘従事者（12.9%）などブルーカラーに分類される職業などでは、自動化対象率は 10%強と比較的低かった。

図表 7：生成 AI 導入によるタスクの自動化対象率（職業（大分類）ベース）



(注 1) 職業別（小分類、中分類）の就業者数で加重平均した値。

(注 2) 宗教家や音楽家など自動化対象率のデータが欠落している職業は除外した。

(出所) 総務省「令和 2 年国勢調査」、独立行政法人労働政策研究・研修機構（JILPT）職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、「解説系ダウンロードデータ ver4.00」より大和総研作成

より細かい小分類ベースで観察すると、電話応接事務員（90.3%）やデータ・エントリー装置捜査員（75.2%）、総合事務員（66.6%）といったルーティン業務が中心の職業群においてタスクの自動化対象率が高かった（図表 8）。

図表 8：自動化対象率が高いおよび低い職業例（職業（小分類）ベース）

自動化対象率が高い職業例	数値	自動化対象率が低い職業例	数値
電話応接事務員	90.3%	土木従事者	0.0%
税理士	85.1%	航空機操縦士	0.0%
個人教師（学習指導）	73.2%	自動車組立従事者	0.0%
記者，編集者	73.0%	自衛官	2.5%
ソフトウェア作成者	72.7%	ビル・建物清掃員	8.9%

(出所) 総務省「令和 2 年国勢調査」、独立行政法人労働政策研究・研修機構（JILPT）職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、「解説系ダウンロードデータ ver4.00」より大和総研作成

ChatGPT などの大規模言語モデルは、文章やプログラムコードの生成・編集、データ分析、質問に対するリアルタイム応答、などが一般的に得意である。著述家（76.3%）、記者・編集者

(73.0%)、システムコンサルタント・設計者 (79.1%)、ソフトウェア作成者 (72.7%)、個人教師 (学習指導) (73.2%) などの自動化対象率が高かったのは、こうした大規模言語の特徴が反映された結果であろう。

一方で、土木作業員 (0.0%)、自動車組立作業員 (0.0%)、自衛官 (2.5%) のように、身体を動かす場面が多い職種では、自動化の可能性が非常に低かった。この背景としては、大規模言語モデルが物理的なデバイスを備えていないため、物理的に移動する、物を持ち上げる、レバーやハンドルを操作するといった動作が単独では困難であることが挙げられる。航空機の操縦士や鉄道の運転士など、乗り物を操作する職業においても、生成 AI による自動化の割合が低い背景は同様だと考えられる。

次に、賃金構造基本統計調査を用いて年収とタスクの自動化対象率の関係性を探った<sup>13</sup> (図表 9)。総じて、専門職や管理職など高賃金の職種ほど自動化対象率が高い傾向があった。また、平均賃金前後の職種でも、オフィスワーカーを中心にこの割合がかなり高いケースが散見される。一方、ブルーワーカーの数值は低い傾向が確認された。

より詳細な分析をするために、職業をおおまかに三つのグループに分類した。

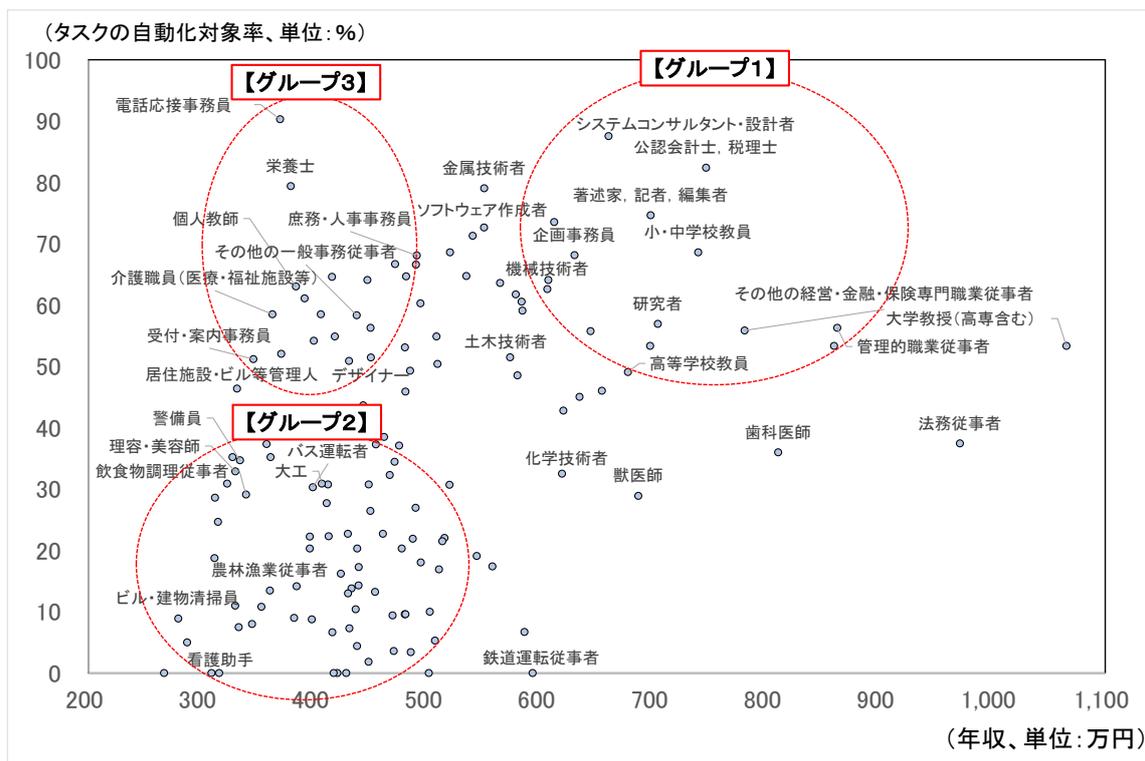
グループ 1 (図表 9 上段右) は、賃金水準と自動化対象率の双方が比較的高い職業グループであり、公認会計士・税理士や管理的職業従事者、(その他の) 経営・金融・保険専門職業従事者などが含まれる。これらは経営スキルや特定の専門技能を持つ職業でありつつも、一部の作業が生成 AI によって自動化可能であることを示唆している。

次に、グループ 2 (図表 9 下段左) は、賃金水準が平均前後かつ自動化対象率が比較的低い職業グループであり、理容・美容師、ビル・建物清掃員、大工などの職業が該当する。これらの職業は、先ほど述べた通り、生成 AI の影響を受けにくい身体的な動作を伴う傾向がある。

最後に、グループ 3 (図表 9 上段左) は、賃金がグループ 2 とそこまで変わらない一方で、自動化対象率が比較的高いグループである。電話・応接事務員や栄養士、受付・案内事務員などがここに分類される。グループ 3 はグループ 1 と同様に自動化対象率が高い一方で、ルーティン業務の割合はグループ 1 より高いと思われる。

<sup>13</sup> なお、先ほど議論した国勢調査と賃金構造基本統計調査の職業分類は異なる場合がある。

図表9：職業ごとの年収と自動化対象率の関係（賃金構造基本統計調査の職業ベース）



(注1) 「年収＝きまって支給する現金給与額×12＋年間賞与その他特別給与額」として計算した。

(注2) 医師（年収：1,428.9万円）、航空機操縦士（年収：1,600.4万円）は年収が極めて高かったため、外れ値として分析対象から除外した。

(注3) 年収の計算には、企業規模計（10人以上）のデータを使用した。

(注4) 宗教家など自動化対象率のデータが欠落している職業は除外した。

(出所) 厚生労働省「令和4年賃金構造基本統計調査」、独立行政法人労働政策研究・研修機構（JILPT）職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、「解説系ダウンロードデータ ver4.00」より大和総研作成

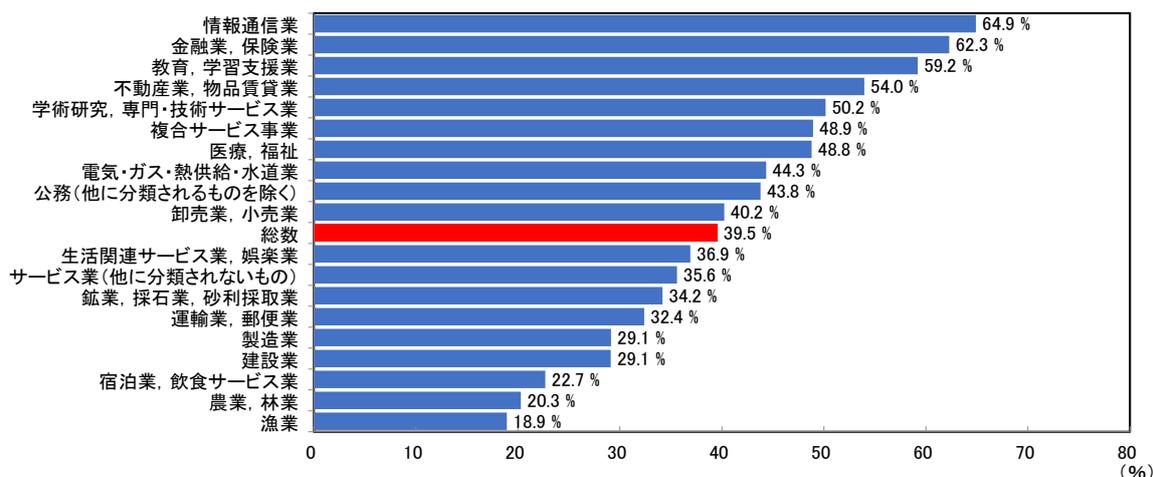
### 3-3. 産業別ではIT産業、金融業、不動産業、教育産業などへの影響が特に大きい

次に、国勢調査の産業（大分類）ごとに自動化対象率を集計した（図表10）。集計に当たっては、当該産業に属する職業別就業者（小分類）の自動化対象率をその人数で加重平均した。

総じて、情報やデータを扱うオフィスワーク中心の産業で数値が比較的高かった。数字が高い順番に、情報通信業（64.9%）、金融業・保険業（62.3%）、教育・学習支援業（59.2%）が特に高かった。これらの産業では、プログラムコードの生成や文章の生成・要約、データ分析、質問への即時応答など生成AIの影響を受けやすい職業の人が多く働いている。従って、こうした産業では、生成AIによって業界の勢力図や従業員の働き方が他の産業よりも大きく変化する可能性がある。

他方、身体的な動作を伴う場面が多い建設業や漁業、宿泊業・飲食サービス業などは生成AIによる影響は比較的低めとの結果が得られた。ただし、これらの業種においても、一部の管理職や事務職など個別の職業に対しては、生成AIの影響をより受ける可能性がある。

図表 10：生成 AI 導入によるタスクの自動化対象率（産業（大分類）ベース）

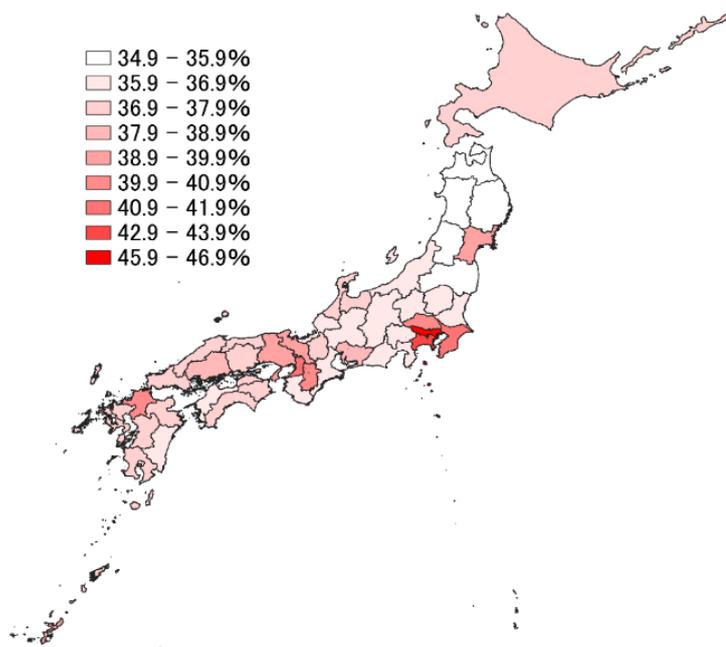


(注) 職業別（小分類）の就業者数で加重平均した値。宗教家や音楽家など自動化対象率のデータが欠落している職業は除外した。

(出所) 総務省「令和 2 年国勢調査」、独立行政法人労働政策研究・研修機構（JILPT）職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、「解説系ダウンロードデータ ver4.00」より大和総研作成

最後に、都道府県別に生成 AI の自動化対象率を見ると、多くの地域が 40%前後に収まった（図表 11）。それゆえ、職業別や産業別の時に比べるとデータのばらつきは比較的小さい。

図表 11：タスクの自動化対象率（都道府県ベース）



(注) 職業別（小分類）の就業者数で加重平均した値。宗教家や音楽家など自動化対象率のデータが欠落している職業は除外した。

(出所) 国土交通省「国土数値情報 行政区域データ（令和 5 年）」、総務省「令和 2 年国勢調査」、独立行政法人労働政策研究・研修機構（JILPT）職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、「解説系ダウンロードデータ ver4.00」より大和総研作成

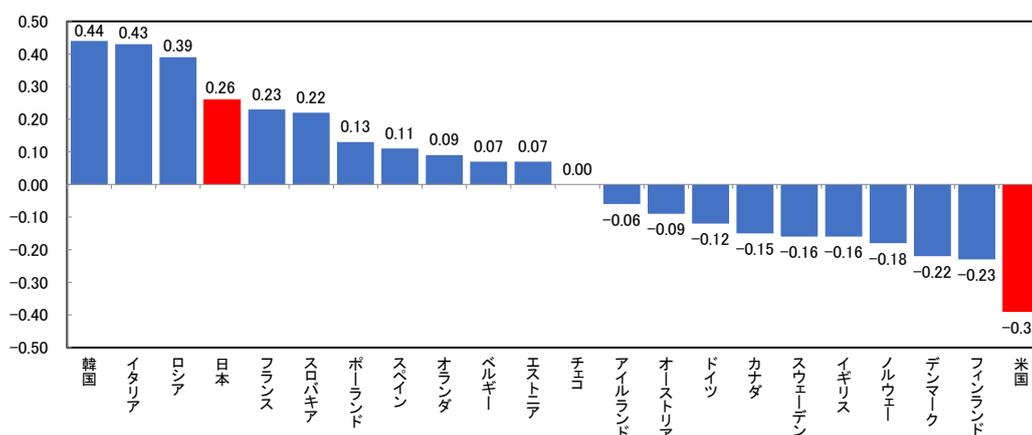
自動化対象率が高い地域は、東京都を中心とする1都3県、大阪府とその周辺、福岡県、宮城県など都市部を中心とする地域であった。とりわけ東京都の数値は突出していた。他方、宮城県を除く東北地方をはじめ、地方の数値は低めであった。両者の差は、IT産業や金融業など自動化対象率が高い産業の集積度合いの違いから生じたものと思われる。

### 3-4. 米国と全体的な傾向は一致、ただし、一部では違いも

本レポートの結果と米国を対象にした先行研究の結果を比較したところ、全体的な傾向は一致していた。

まず、米国を対象にした先行研究では、2021年の就業者のうち、自動化対象率が10%および50%を超えた割合はそれぞれ約83%および約23%<sup>14</sup>であった。本レポートの試算では、10%を超えた割合は81.7%と米国の場合とほぼ一致した結果であった。一方、50%を超えた割合は40.9%と本レポートの方が先行研究よりも高かった。この背景の一つとして、産業構成の違いに加え、日本は米国に比べてルーティン業務をより多く抱えている可能性が指摘できる。例えば、De La Rica and Gortazar(2015)がOECDの国際成人力調査(Programme for the International Assessment of Adult Competencies, PIAAC)<sup>15</sup>から構築したルーティン業務の集中指数(Routine task-intensity, RTI)を見ると、日本は22か国中4番目の高さだったのに対し、米国は全体で最も低かった(図表12)。

図表12：ルーティン業務の集中指数(Routine task-intensity, RTI)の国際比較



(注) 具体的な算出方法は De La Rica and Gortazar (2015) を参照。

(出所) De La Rica and Gortazar (2015) より大和総研作成

また、米国のケースでは、本レポートと同じく、会計士、数学者、税理士、ライター・作家、翻訳者、ウェブデザイナー、ジャーナリストなどいわゆるホワイトカラーの職業の自動化対象率が高かった。賃金との関係でも、賃金が高いほど自動化対象率が高いという関係が観察され

<sup>14</sup> 本レポートの自動化代替率と対応する Eloundou et al. (2023) の Figure 3 の右図の  $\beta$  (model) と比較した。

<sup>15</sup> 調査は 2011 年から 2012 年にかけて実施された。

た。身体的な動作や機械の操作を伴う職業、具体的にはレンガ職人、バイク整備士、大工、オイル・ガス掘削機操作者などの自動化対象率はゼロであった。これら職業別の影響は、本レポートで分析した日本の結果とほぼ一致している。

### 3-5. 分析の留意点

本レポートの分析にはいくつか留意すべき点がある。

まず、タスクの分類に用いた GPT-4 の出力結果に関しては、現状では完全な再現性がない。さらに、今回の分析にはデータや手法面に制約や限界があった。従って、分析結果の数字はある程度の幅を持って見る必要がある。

制約の1つはデータセットであった。日本版 0-NET の開発およびデータベースの整備は、職業情報の提供などに加え、日本の労働市場を分析する上でも非常に重要な取り組みだと評価できる。他方、2023年11月現在、日本版 0-NET に掲載されている職業数は521職業と米国の 0\*NET の1,016職業の半分程度にとどまる。また、記述されているタスクの数も比較的少なめである。例えば、公認会計士のタスク数は日本版<sup>16</sup>では9個であり、30個も示している米国版<sup>17</sup>の1/3以下である。加えて、既述の通り、日本版 0-NET の一部の職業はタスクが一つも記述されていない。

職業分類に関しては、日本版 0-NET、日本標準職業分類、国勢調査などデータセット間で分類が1対1で対応していない。このため、本レポートでは一部の職業に関しては、一定の仮定を置き、単純平均や加重平均を人手で計算するケースもあった。

本レポートのタスクに基づくアプローチに関しても、いくつかの限界が存在する。第一に、各職業のタスクを分解し、個別に分析するアプローチ自体の限界が挙げられる。職業におけるタスクはしばしば相互に関連しており、単独で実施するという仮定が成立しない場合がある。例えば、あるタスク A が別のタスク B の前提条件となっているケースや、複数のタスクが同時に行われる場合などが該当する。しかし、データの制約上、今回はこの点を議論することはできなかった。

さらに重要なのは、本レポートで議論した生成 AI による自動化は「人間の仕事を代替すること」と必ずしもイコールではない点である。生成 AI などの新たなテクノロジーは定型的な作業から非定型的な作業へのシフトを促し、「人間の仕事を補完」する場合がある。このため、本レポートで議論した自動化対象率は、よく話題に上がる AI による失業リスクと直接結びつけるべきではない。

<sup>16</sup> 職業情報提供サイト（日本版 0-NET）公認会計士  
(URL : <https://shigoto.mhlw.go.jp/User/Occupation/Detail/90>)

<sup>17</sup> 0\*NET OnLine Accountants and Auditors  
(URL : <https://www.onetonline.org/link/summary/13-2011.00>)

## 4. まとめと今後に向けて

本レポートは、生成 AI が日本の労働市場に与える影響を定量的かつ包括的に分析した数少ない試みである。定量的な影響を分析するために、本レポートでは、職業ごとにタスクのうち生成 AI の影響を受ける割合を示す「自動化対象率」を推計した。この推計する上では、OpenAI 社が開発した GPT-4 と呼ばれる大規模言語モデルを活用した。

分析結果を見ると、日本の就業者の約 80% が何らかの形で生成 AI の影響を受ける可能性があり、同様に約 40% の就業者が仕事の半分以上を自動化できると推計された。これは、生成 AI が幅広い職種のタスクに変化をもたらす可能性を示唆している。

職業別では、事務従事者や管理的職業従事者、専門的・技術的職業従事者などホワイトカラーの職種を中心に、自動化の影響を強く受けるとの結果が得られた。これらの職種は、文章やプログラムコードの生成、データ分析など大規模言語モデルが比較的得意なタスクに関わることが多いためであろう。他方で、運搬・清掃・包装等従事者や建設・採掘従事者などブルーカラーに分類される職業などに対しては、生成 AI が及ぼす影響は比較的低かった。賃金別に見ると、高賃金の職種ほど自動化対象率が高い傾向が見られた。

産業別の分析では、情報通信業や金融業・保険業、不動産業・物品賃貸業など情報やデータを扱うオフィスワークが中心の産業で自動化対象率が比較的高かった。その反面、身体的な動作を伴う場面が多い建設業や漁業、宿泊業・飲食サービス業などの数字は比較低い水準であった。また、地域別に見ると、IT 産業や金融業が集中する都市部で生成 AI の影響が高め、地方では低めの傾向が見られた。

これら一連の結果は米国を対象にした先行研究と全体的な傾向は一致した。一方、タスクの自動化対象率が 50% を超えた割合は本レポートの方が先行研究よりも高かった。この背景の一つとして、産業構成の違いなどに加え、日本は米国に比べてルーティン業務をより多く抱えている可能性が指摘できる。

次回のレポートでは、今回議論した自動化対象率が高い職業の性質や特徴に関して定量的に分析する。さらに、生成 AI が雇用の代替および補完に及ぼす影響に関してより踏み込んだ議論を展開する方針である。その上で、生成 AI の活用を踏まえたリスクリング等に関する政策を提案する予定である。

以上

---

**【参考文献】**

Fabrizio, G., M. Alizadeh, and M. Kubli (2023), “ChatGPT Outperforms Crowd-Workers for Text-Annotation Tasks,” Papers 2303.15056, arXiv.org, revised Jul 2023.  
<https://doi.org/10.48550/arXiv.2303.15056>

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2023), “GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models,” Papers 2303.10130, arXiv.org, revised Mar 2023. <https://arxiv.org/abs/2303.10130v4>

Sara De La Rica and L. Gortazar (2015), “Differences in Job De-Routinization in OECD countries: Evidence from PIAAC,” *Working Papers 2015-11*, FEDEA. <https://ideas.repec.org/p/fda/fdaddt/2015-11.html>

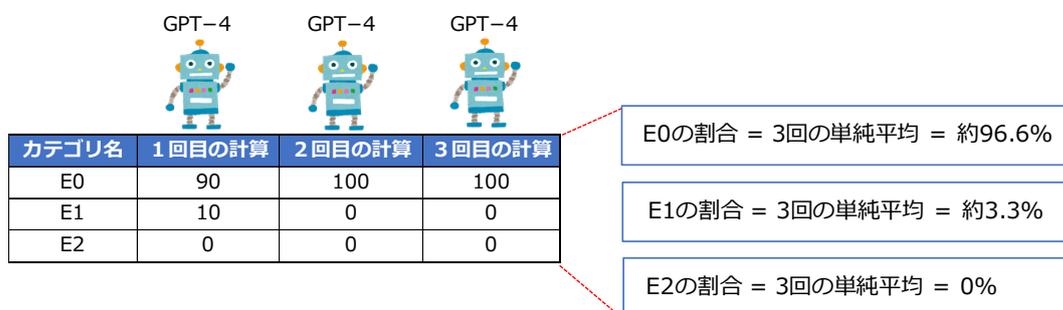
## 補論 1：タスクが取得できない職業に対する自動化対象率の計算方法

既述の通り、日本版 O-NET に含まれる全 521 職業のうち 65 職業はタスクが取得できない。一方、「仕事の内容<sup>18</sup>」に関するテキストは全ての職業で取得できる。例えば、保育士の仕事の内容は「専門的知識及び技術をもって、子どもの保育を行うとともに、保護者に対し子育てに関する支援や助言・指導を行う。認可保育所をはじめとする保育施設において、保護者が仕事のため日中に家庭で保育ができない場合など、主に就学前の保育を必要とする子どもの保育に従事する保育士が多くを占めている。(以下略)」という約 800 文字の説明文がある。

また、一部の職業には、よく使う道具、機材、情報技術等が示されている。例えば、経理事務の職業では、当該項目には「経理・会計管理のソフト、クラウド、表計算ソフト (Excel、スプレッドシート等)、パソコン」と記述されている。

そこで、本レポートでは、タスクが取得できる場合と同じ分類基準を用いつつ、GPT-4 に対して、「仕事の内容」のテキストを基に、「この仕事内容に含まれるタスクのうち何パーセントが E0、E1、E2 に分類されるのか推定して下さい。」と指示した。これにより、タスクが取得できない職業に対しても、「E0 の割合が 50%、E1 の割合が 40%、E2 の割合が 10%」、のように各カテゴリに分類されるタスクの割合が計算できる。計算はタスクごとに 3 回ずつ実施し、この平均値を最終結果として採用した<sup>19、20</sup>。

補論図表 1：タスクが取得できない職業の自動化対象率の計算例（職業：合板製造）



(出所) 独立行政法人労働政策研究・研修機構 (JILPT) 職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」より大和総研作成 (イラストはいらすとや (<https://www.irasutoya.com/>) による)

<sup>18</sup> 日本版 O-NET のウェブサイトでは「どんな仕事？」と表現されている。

<sup>19</sup> 万が一、各カテゴリのシェアの合計値が 100%に一致しない場合は、シェアの合計値が 100%に一致するように加重平均を用いて調整した。

<sup>20</sup> このケースでは、元データからタスクの実施率を取得できない。しかし、GPT-4 のプロンプト文で「この仕事内容に含まれるタスクのうち何パーセントが E0、E1、E2 に分類されるのか推定して下さい。」と指示しているため、本レポートでは、出力結果をタスクの実施率で加重平均した自動化対象率と解釈した。

## 補論 2 : 国勢調査や賃金構造基本統計調査とのマッチング

以下、日本版 0-NET と国勢調査や賃金構造基本統計調査とマッチングするための方法を説明する（補論図表 2）。

はじめに、日本版 0-NET と日本標準職業分類（小分類）の職業をマッチングした。マッチングの際には、独立行政法人労働政策研究・研修機構（JILPT）「職業情報データベース 解説系ダウンロードデータ」に記述されている対応表を使用した。ここで、日本標準職業分類（小分類）の職業一つに対して、複数の日本版 0-NET の職業が対応する場合がある。この場合は、対応する日本版 0-NET の職業の自動化対象率の単純平均を、該当する日本標準職業分類の職業の自動化対象率とした。その上で、日本標準職業分類（中分類）の自動化対象率を当該分類に含まれる小分類の単純平均とした。同様に、日本標準職業分類（大分類）の自動化対象率も、当該分類に含まれる中分類の単純平均を計算した。

次に、日本標準職業分類（小分類）と国勢調査の職業（小分類）をマッチングした。マッチングの際には、総務省統計局「令和 2 年国勢調査に用いる職業分類」の「日本標準職業分類（平成 21 年 12 月設定）との分類項目比較対照表」の対応表を用いた。なお、国勢調査の職業一つに対して、複数の日本標準職業分類の職業が対応する場合は、対応する日本標準職業分類の職業の自動化対象率の単純平均を該当する国勢調査の職業の自動化対象率とした。国勢調査の中分類、大分類、総数の自動化対象率を求める際には、当該分類に属する 1 階層低い職業分類（例えば、中分類より 1 階層低い職業分類は小分類）の就業者数で加重平均した。

最後に、日本標準職業分類と賃金構造基本統計調査の職業をマッチングさせた。賃金構造基本統計調査の職業の多くは日本標準職業分類をベースにして作成されたため、1 対 1 で対応するケースが多い。他方、前二つのケースとは異なり、両者の職業対応リストは存在しない。こうした中、日本標準職業分類では「システムコンサルタント」と表記される職業名が賃金構造基本統計調査では「システムコンサルタント・設計者」と表記されるケースがあるなど、全て人手でマッチングすると煩雑な作業になりかねない。そこで、本レポートでは、Levenshtein 距離および Jaro-winkler 距離という文字列の類似度を計算するアルゴリズム<sup>21</sup>を利用して、より効率よくマッチングを行った<sup>22</sup>。

<sup>21</sup> 例えば、Levenshtein 距離とは、文字列 A を 1 文字ずつ操作（削除・挿入・置換）して文字列 B に一致するまでの最小回数を指す。

<sup>22</sup> なお、本文の手法で直接マッチングできない場合は、人手により以下の対応を行った。

・マッチング可能と考えられる職業が見つかった場合：類似した職業の情報を活用して計算した。例えば、国勢調査の「卸売店主・店長」は直接マッチングできなかったため、類似していると考えられる「小売店主・店長」（国勢調査）の数値を用いた。

・マッチング可能と考えられる職業が見つからない場合：当該職業の数値は欠落とした。例えば、日本標準職業分類や国勢調査における「宗教家」や「音楽家」などは対応する 0-NET の職業が見つからなかったため、数値を欠落とした。

## 補論図表 2 : マッチングの手順

### 1. 日本版O-NETの職業ベースで集計

職業名	E1の割合
運用・管理 (IT)	40%
ヘルプデスク (IT)	60%
セキュリティエキスパート (オペレーション)	50%

職業名	E1の割合
電気通信技術者	20%
システムエンジニア (基盤システム)	60%

職業名	E1の割合
デジタルビジネスインベーター	70%
セキュリティエキスパート (脆弱性診断)	50%
セキュリティエキスパート (デジタルフォレンジック)	60%

A. 日本版O-NETと日本標準職業分類の職業対応リストを利用してマッチング

B. 日本標準職業分類と国勢調査の職業対応リストを利用してマッチング

C. 文字列の類似度を計算するアルゴリズムを利用してマッチング

### 2. 日本標準職業分類の職業ベースで集計

職業名	E1の割合
システム運用管理者	$(40+60+50) \div 3 = 50\%$
通信ネットワーク技術者	$(20+60) \div 2 = 40\%$
その他の情報処理・通信技術者	$(70+50+60) \div 3 = 60\%$

これらのデータセットを分析に使用

### 3. 国勢調査の職業ベースで集計

職業名	E1の割合
その他の情報処理・通信技術者	$(50+40+60) \div 3 = 50\%$

### 3. 賃金構造基本統計調査の職業ベースで集計

職業名	E1の割合
その他の情報処理・通信技術者	60% (日本標準職業分類の「その他の情報処理・通信技術者」と同じ)

(注1) 「日本版 O-NET と日本標準職業分類の職業対応リスト」には、独立行政法人労働政策研究・研修機構 (JILPT) 職業情報データベース「解説系ダウンロードデータ ver4.00」を用いた。

(注2) 日本標準職業分類と国勢調査の職業対応リストには、総務省統計局「令和2年国勢調査に用いる職業分類」の「日本標準職業分類 (平成21年12月設定) との分類項目比較対照表」を利用した。

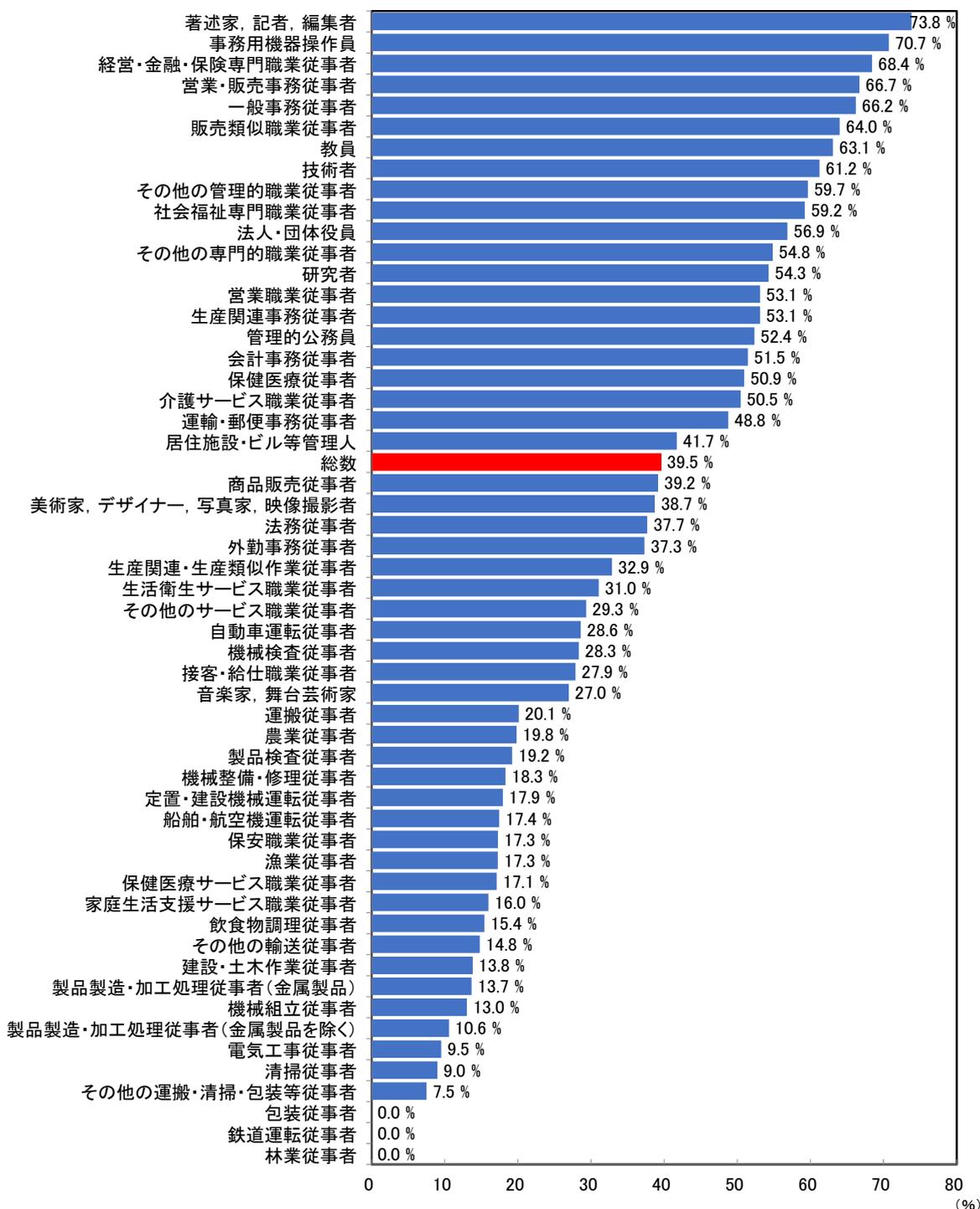
(注3) 文字列の類似度を計算するアルゴリズムには、Levenshtein 距離および Jaro-winkler 距離を利用した。

(注4) 職業分類は全て小分類。

(出所) 大和総研作成

## 補論3：国勢調査の職業（中分類・小分類）ごとの自動化対象率

補論図表3：生成AI導入による自動化対象率（職業（中分類）ベース）

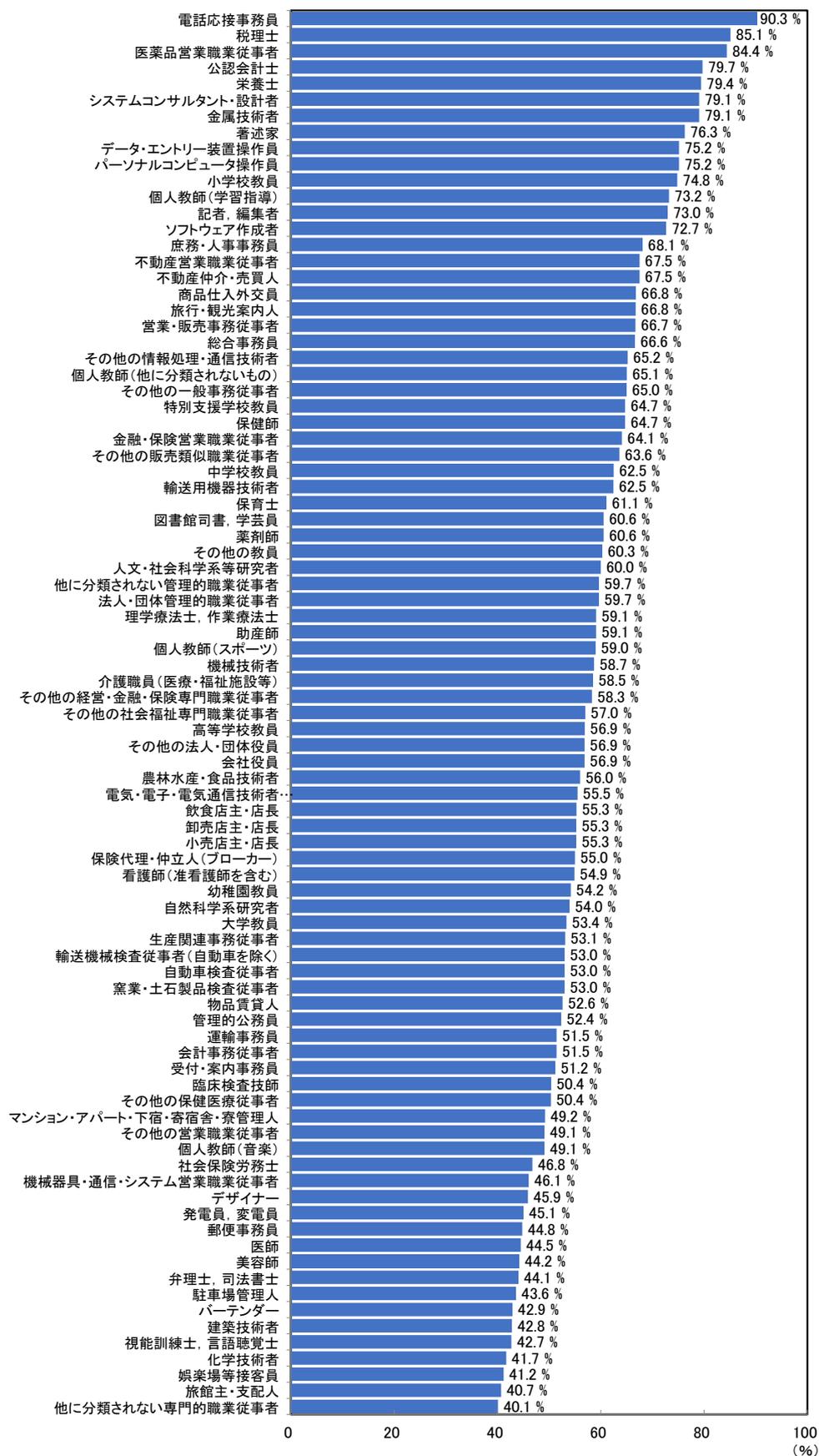


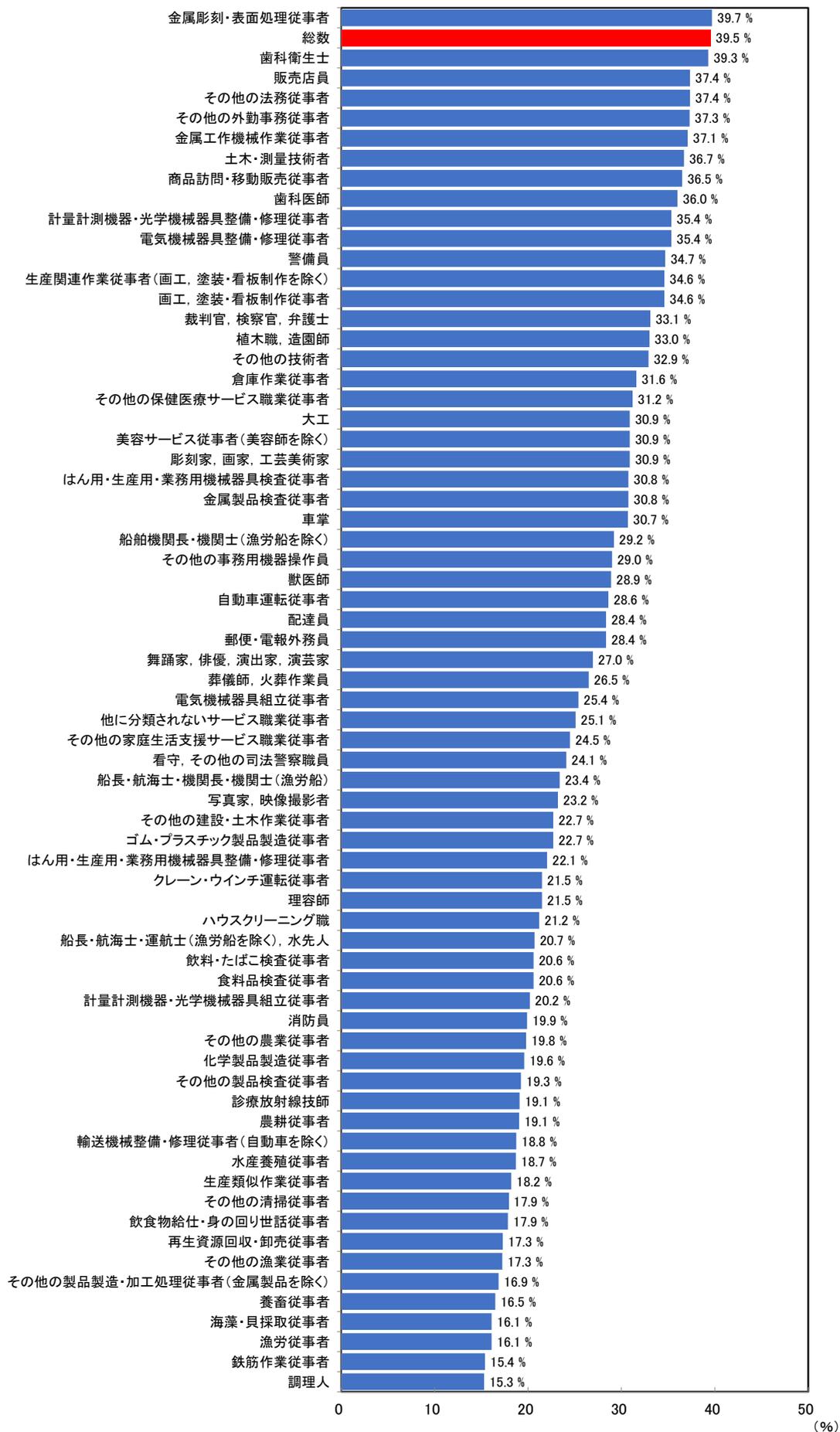
(注1) 職業別（小分類）の就業者数で加重平均した値

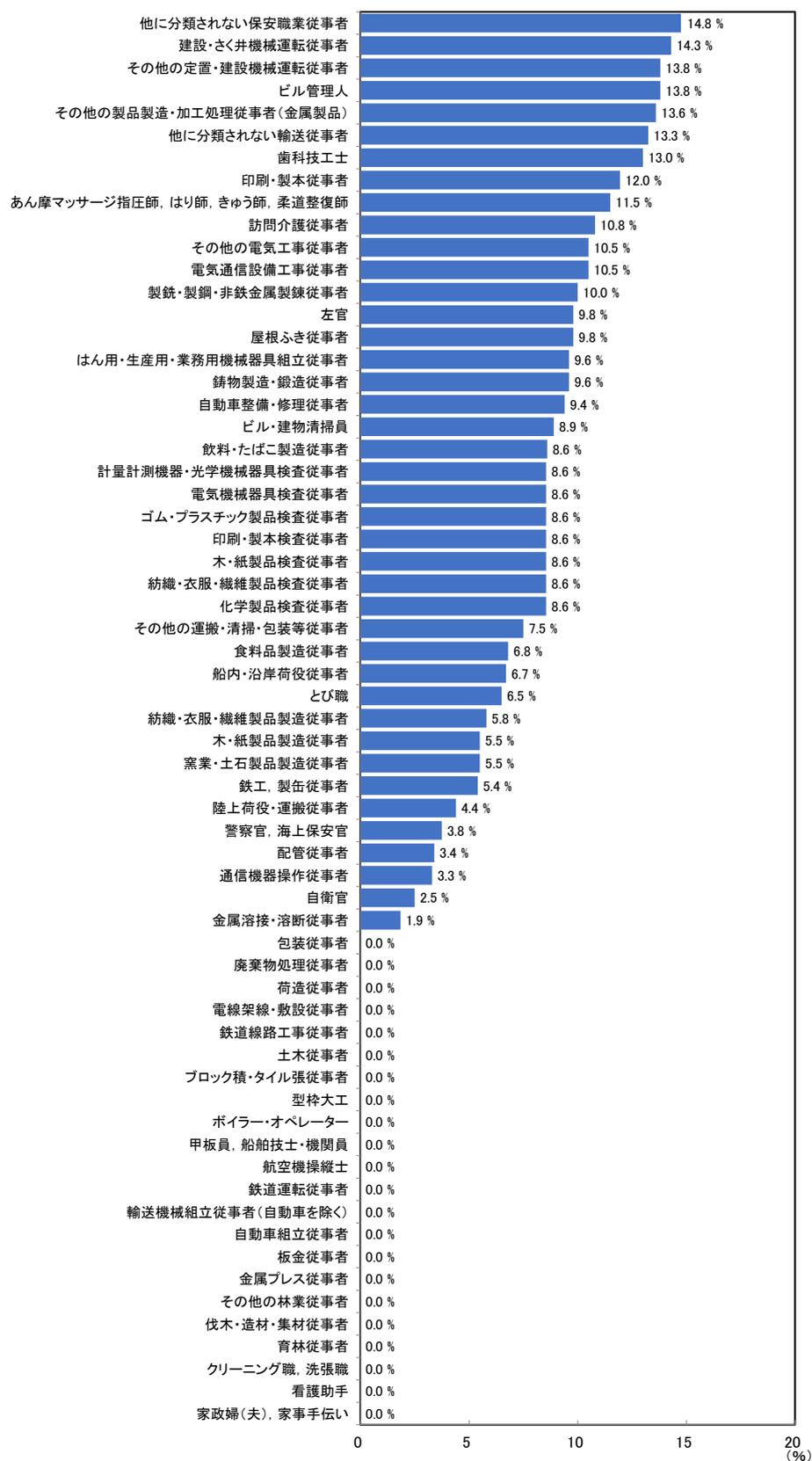
(注2) 宗教家や音楽家など自動化対象率のデータが欠落している職業は除外した。

(出所) 総務省「令和2年国勢調査」、独立行政法人労働政策研究・研修機構（JILPT）職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、「解説系ダウンロードデータ ver4.00」より大和総研作成

補論図表 4 : 生成 AI 導入による自動化対象率 (職業 (小分類) ベース)







(注1) 電気・電子・電気通信技術者…は「電気・電子・電気通信技術者(通信ネットワーク技術者を除く)」を指す。

(注2) 宗教家や音楽家など自動化対象率のデータが欠落している職業は除外した。

(出所) 総務省「令和2年国勢調査」、独立行政法人労働政策研究・研修機構(JILPT)職業情報データベース「簡易版数値系ダウンロードデータ ver4.00」、「解説系ダウンロードデータ ver4.00」より大和総研作成