

# 保険会社における 生成AI活用実践（採用編）

生成AIは、単なる業務効率化に留まらず、保険会社社員と同等レベルでの業務実行可能性を秘めている。保険会社用AIのキャリア開発プログラムを策定し、5つの生成AIについて保険会社用AIとして活用できるかどうか評価試験を行った。

## 入社1年後の新人と 代替できるかを検証

近い将来、複数のAIが人と同じように個々の専門性を発揮しながら保険会社で活躍するようになるだろう。保険会社におけるこれまでのAI活用はAI-OCRのような事務の高度化や効率化が主だった。自然な対話が可能になった生成AIの登場により、創造的な業務までもがAIに代替される可能性がある。

NRIでは、このような将来を見据え、保険会社の人材のキャリア開発プログラム（CDP）を参考にして「AI用の開発プログラム」を策定した（図表1）。段階的に専門性が高度化する三層構造とし、一層目は保険会社入社1年後の社員を想定した「保険新人AI」である。保険業務を1年相当経験した知識や実務スキルを有する。二層目は入社3年目相当の「保険アソシエイトAI」。より高度な専門性に加え、保険会社ごとの商品知識や企業理念などを備えている。そして、個々の専門性を磨き、三層目となる「保険専門

AI」に分化していくというステップである。保険専門AIについては継続的なフィードバックと追加学習（育成）が必要である。専門家による学習に加え、AI同士による育成も可能となると考えている。

こうしたCDPを策定した後、数多く存在する生成AIから性能や得意領域を考慮し、生成AIを「採用」（選択）することになるが、ここでは「保険新人AI」における生成AIの採用とその評価について解説する。なお、採用では追加学習を行っていない、あくまで素のモデル（Pre-trained）を用いていることに留意して頂きたい。また、「育成」については続編で解説することとする。

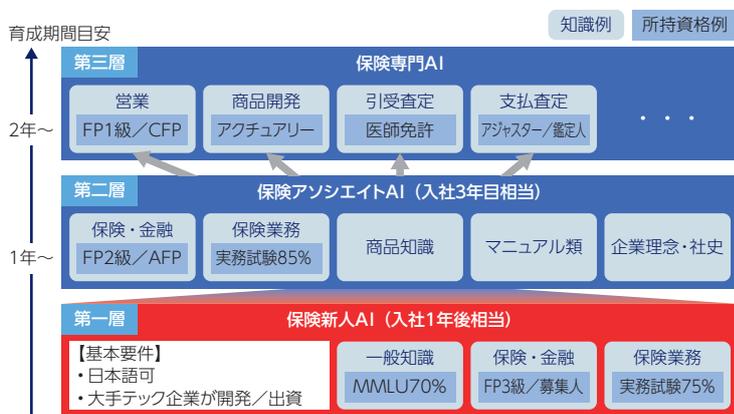
## 保険新人AIの採用における評価基準

まず基本要件等から生成AIの採用候補を選定した。日本語に対応し、技術革新の将来性から大手テック企業が開発または出資していることに加え、一般知識の担保指標であるMMLUスコアが70%以上<sup>1)</sup>という基準を設定した。その結果、GPT-4、GPT-3.5、Palm2、

Claude2、Gemini Proが対象候補<sup>2)</sup>となった。この選定は人材採用であれば、書類選考とSPI試験の能力測定にあたる。今後の課題ではあるが、人材採用と異なり生成AIのポリティカルコンパス<sup>3)</sup>といった要因を考慮する必要もあるかもしれない。

次に、選考試験として保険業務特化の学科試験と実技試験の2つの試験を設けた。学科試験では、保険商品のみならず、他の金融商品も取り扱う可能性を踏まえ、ファイナンシャルプランニング技能検定試験（以下、

図表1 保険会社用AI開発プログラムイメージ



(出所) 野村総合研究所

**NOTE**

- 1) MMLU (Massive Multitask Language Understanding)。自然言語モデルの知識量と問題解決能力を評価するために設計された代表的なベンチマーク。初等数学、歴史、コンピュータサイエンス、法律などの57のタスクをカバーしている。
- 2) 使用モデルはGPT-3.5 : gpt-3.5-turbo-0613、Palm2 : text-bison@001、Gemini Pro : gemini-pro である。GPT-4、Claude2は2023年11月時点で公開されているブラウザ版を使用した。Temperatureの設定はGPT3.5、Palm2、Gemini Proは0、GPT4、Claude2はデフォルト値で実施した。
- 3) 生成AIのポリティカルコンパス：ワシントン大学、カーネギーメロン大学、西安交通大学の研究者により2023年7月に発表された。From Pretraining Data to Language Models to Downstream Tasks: Tracking the Trails of Political Biases Leading to Unfair NLP Models

**図表2 設問例：20代女性への保険提案結果**

		GPT-4	GPT-3.5	Claude2	Palm2	Gemini Pro
スコア (4点満点)		4	4	3	3	3
スコア根拠		リスク対策・資産形成に対する商品を提案	リスク対策・資産形成に対する商品を提案	リスク対策に対する商品のみ提案	リスク対策に対する商品のみ提案	パートナー有無で商品に変化なし
パートナー有り	リスク対策	生命保険 医療保険 学資保険	生命保険 医療保険 学資保険	生命保険 医療保険 団体信用生命保険	生命保険 学資保険	生命保険 医療保険 がん保険
	資産形成	貯蓄型保険	個人年金保険	—	—	貯蓄型保険
パートナー無し	リスク対策	生命保険 医療保険 就業不能保険	医療保険	養老保険	医療保険 がん保険	生命保険 医療保険 がん保険
	資産形成	貯蓄型保険	個人年金保険	—	—	貯蓄型保険

(出所) 野村総合研究所

FP 3級の学科試験一回分60問（2023年5月試験）を解かせた。入社1年後相当の保険・金融知識として、FP合格水準である60%の得点率が必要と定義した。

実技試験では、保険会社社員が保険募集人資格を原則取得することから保険提案に関する設問をNRIが作成し、出力された回答が保険実務に適用できるかをNRI社内の保険会社出身者によって4段階評価を行った。

図表2はその設問の一例である。“年収800万、資産形成への意識が高い都内在住20代女性に対し、パートナー有無別に保険外交員として保険提案する”という設問では、ケースごとに妥当な保険商品をリスク対策と資産形成の両面から提案できれば4点とした。ケースが変わっても提案に変化がない場合や資産形成に触れていない場合は3点、提案商品の妥当性に疑問がある場合は2点、誤った回答は1点とした。保険新人AI水準として、得点率75%（平均3点）が必要と定義した。

**素のモデル (Pre-trained) の試験結果**

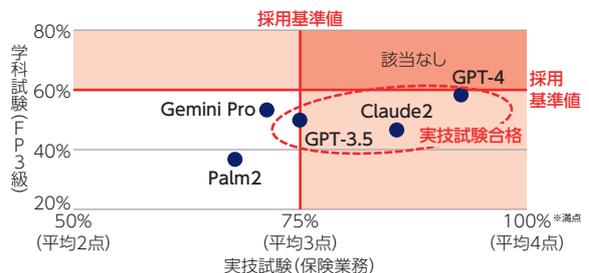
選考の結果（図表3）、学科試験（FP3級）では保険

新人AIの基準に達する生成AIはなかった一方、実技試験（保険業務）ではGPT-4、Claude2、GPT-3.5が保険新人AIの基準となる得点率75%（平均3点）を上回った。最高スコアであるGPT-4は学科試験：58.3%、実技試験：93%の得点率であった。学科試験は追加学習で点数の向上が見込めることから、実技試験で合格水準だった生成AIを採

用し、次号の「育成編」では追加学習を行い得点率がどのように向上したかを示したい。

今回はFP3級と保険業務の設問を評価基準としたが、試験内容や得点率の考え方は保険会社各社で様々だろう。次々と新たな生成AIが登場する昨今では、自社の評価基準を定め、常に代替可能性を検討する必要がある。

**図表3 素のモデル (Pre-trained) の試験結果**



(出所) 野村総合研究所

**Writer's Profile**

**鷺山 寛子** Hiroko Washiyama  
 保険デジタル企画部 未来保険研究室  
 エキスパートコンサルタント  
 専門は保険業界の新規事業・サービス創出支援  
 focus@nri.co.jp