

企業の労務業務における AI RAG (Retrieval-Augmented Generation) の 効果的な活用方法に関する研究

— 労務 DX とガバナンス強化の観点から

2025/12/17

株式会社コマースロボティクス

代表取締役 伊藤彰弘

取締役 CFO 庄子敬弥

要約（Abstract）

日本企業における労務業務は、少子高齢化に伴う人材不足、ホワイトカラーの低生産性、法令改正への継続的対応といった複合的課題を抱えている。近年、生成AIの活用が進む中で、企業内業務への適用においては、誤回答リスクや説明責任、実務定着といった観点から慎重な設計が求められている。

本研究は、労務業務におけるAI RAG（Retrieval-Augmented Generation）の効果的な活用方法について、実務検証と設計思想の整理を通じて検討した。分析の結果、労務RAGの価値は、単なる自動応答や問い合わせ削減にとどまらず、RAGを中心に入・専門家・ナレッジが連携する業務基盤の構築にあることを明らかにした。

具体的には、労務RAGを「答えを出すAI」ではなく「根拠を提示するAI」として設計し、自動応答の対象をレベル1～3に限定することで、誤回答やブラックボックス化のリスクを抑制できることを示した。また、Q&A形式データを中心としたデータ設計、更新管理を重視したメタデータ設計、労務業務に特化したプロンプト設計が、回答精度と利用定着に寄与することを確認した。

さらに、RAGと労務担当者とのシームレスな連携、ならびに社会保険労務士などの専門家との連携が、労務人材不足の補完や判断品質の向上に有効であることを示した。社員向けの労務RAGにおいては、回答精度が利用定着の前提条件であり、制度説明と労務システムの利用方法を統合した回答が、自己解決率の向上に寄与することが明らかとなった。

本研究は、労務業務におけるAI活用を技術論にとどめず、設計・運用・人材の視点から体系的に整理した点に新規性がある。本稿で示した労務RAGの活用モデルは、実務における安全性と有効性を両立する指針として、今後の企業内AI活用に資するものと考える。

第1章 序論

- 1.1 研究背景
- 1.2 日本企業における労務業務の構造的課題
- 1.3 生成AI・RAG技術の登場と企業実務への影響
- 1.4 本研究の目的
- 1.5 研究の意義と新規性
- 1.6 本論文の構成

第2章 労務業務の現状分析と課題整理

- 2.1 労務業務の定義と業務範囲
- 2.2 企業規模別に見る労務業務の特徴
- 2.3 労務担当者への問い合わせ業務の実態
- 2.4 ナレッジの分散・属人化問題
- 2.5 法改正・制度変更への追随負荷
- 2.6 従来のIT化・FAQ化の限界

第3章 AI RAG技術の概要

- 3.1 生成AIの基本構造
- 3.2 RAG (Retrieval-Augmented Generation) の仕組み
- 3.3 従来型生成AIとRAGの違い
- 3.4 RAGにおける「検索」「根拠提示」「回答生成」のプロセス
- 3.5 企業内業務におけるRAGの優位性
- 3.6 労務領域における適用可能性

第4章 労務業務におけるAI RAG活用モデル

- 4.1 労務業務×RAGの全体アーキテクチャ
- 4.2 社員向け問い合わせ対応への活用
- 4.3 労務担当者・管理部門向け支援への活用
- 4.4 規程・社内ルール・法令データの扱い方
- 4.5 「答えを出すAI」ではなく「根拠を提示するAI」としての設計
- 4.6 チケット管理・業務フローとの連携モデル

第5章 RAG導入におけるナレッジ設計とデータ整備

- 5.1 労務ナレッジの体系化の重要性
- 5.2 学習対象とすべきドキュメントの分類
- 5.3 Q&A形式とドキュメント形式の使い分け
- 5.4 メタデータ設計と検索精度への影響
- 5.5 ナレッジ正規化と表現ゆれ対策
- 5.6 法令・社内規程の更新管理

第6章 労務RAG導入効果の定量・定性分析

- 6.1 労務業務における KPI 設計
- 6.2 問い合わせ対応工数削減効果
- 6.3 回答品質・再質問率の変化
- 6.4 労務担当者の負荷軽減と役割変化
- 6.5 社員満足度・自己解決率への影響
- 6.6 組織全体のガバナンス強化効果

第 7 章 導入時のリスクと課題

- 7.1 誤回答・誤解釈リスク
- 7.2 法的責任と AI 回答の位置づけ
- 7.3 ブラックボックス化への懸念
- 7.4 セキュリティ・個人情報保護への対応
- 7.5 労務担当者の AI リテラシー問題
- 7.6 運用フェーズでの課題

第 8 章 ケーススタディ（または実証分析）

- 8.1 中小企業における RAG 導入事例
- 8.2 中堅企業・大企業での活用事例
- 8.3 問い合わせ対応プロセスの Before / After
- 8.4 労務担当者・社員の評価
- 8.5 成功要因と失敗要因の整理

第 9 章 労務業務における AI RAG の今後の展望

- 9.1 法改正対応の自動化可能性
- 9.2 労務×他管理業務（人事・経理）との連携
- 9.3 マルチボット・役割別 AI の可能性
- 9.4 労務業務の高度化と人の役割の再定義
- 9.5 AI RAG がもたらす労務業務の未来像

第 10 章 結論

- 10.1 本研究の総括
- 10.2 研究結果の実務的示唆
- 10.3 本研究の限界
- 10.4 今後の研究課題

第1章 序論

1.1 研究背景

近年、日本社会は少子高齢化の進行により、あらゆる産業分野において慢性的な人手不足に直面している。特に企業の管理部門に属する労務・人事業務においても、その例外ではなく、限られた人員で高度かつ多岐にわたる業務を遂行することが求められている。

加えて、日本企業におけるホワイトカラーの労働生産性は、国際的に見ても低い水準にあると指摘されている。その要因の一つとして、付加価値を生まない定型業務や確認作業、社内調整業務に多くの時間が割かれている点が挙げられる。労務業務においても、制度説明や規程確認といった反復的な問い合わせ対応が日常的に発生しており、本来注力すべき制度設計や労務リスク管理といった高付加価値業務への時間配分を阻害している。

このような状況下で、企業には単なる人員補充ではなく、業務構造そのものを見直し、生産性を抜本的に向上させる取り組みが求められている。その解決策の一つとして、近年急速に発展している生成AI技術の業務活用が注目を集めている。

特に、社内規程やマニュアル、法令解釈文書といった企業固有の知識を参照しながら回答を生成するRAG (Retrieval-Augmented Generation)技術は、労務業務における知識集約型作業を効率化する手段として高い可能性を有している。RAGを活用することで、反復的な問い合わせ対応を自動化・高度化すると同時に、労務担当者がより戦略的・専門的な業務へと役割転換することが可能となる。

一方で、労務業務は法的リスクやコンプライアンス要件が極めて高い領域であり、生成AIの導入に際しては誤回答や判断のブラックボックス化といった課題への配慮が不可欠である。そのため、労務業務におけるAI活用には、単なる自動化ではなく、「根拠に基づいた情報提供」を前提とした慎重かつ体系的な設計が求められる。

以上の背景から、日本企業が直面する人手不足とホワイトカラーの低生産性という構造的課題に対し、AIRAGをいかに効果的かつ安全に活用できるかを明らかにすることは、学術的にも実務的にも重要な意義を持つと考えられる。

1.2 日本企業における労務業務の構造的課題

日本企業の労務業務は、就業規則、賃金規程、勤怠ルール、各種運用細則など、多層的かつ文書中心のナレッジによって支えられている。しかし、これらの情報は必ずしも体系的に整理・統合されているとは限らず、担当者の経験や暗黙知に依存して運用されているケースが多い。

また、社員からの労務関連の問い合わせは、「有給休暇の取得条件」「残業時間の扱い」「育児・介護休業制度」など、日常的かつ反復的な内容が多いにもかかわらず、問い合わせのたびに個別対応が行われることが一般的である。このような対応は業務効率を低下させるだけでなく、回答のばらつきや属人的判断によるリスクを生む要因ともなっている。

従来、これらの課題に対してはFAQの整備や社内ポータルの構築といったIT施策が講じ

られてきたが、制度改正への追随の難しさや検索性の低さといった問題から、十分な効果を上げているとは言い難い。

1.3 生成 AI・RAG 技術の登場と企業実務への影響

近年の大規模言語モデル（Large Language Models : LLM）の進化により、自然言語による高度な情報処理が可能となった。しかし、単体の生成 AI は学習データに依存する性質上、最新情報や企業固有のルールを正確に反映できないという課題を有している。

この課題を補完する技術として注目されているのが、RAG（Retrieval-Augmented Generation）である。RAG は、外部または内部の情報源から関連文書を検索し、その内容を根拠として回答を生成する仕組みであり、回答の正確性や説明可能性を高めることができる。

企業実務において RAG を活用することで、社内規程やマニュアル、法令解釈文書などを参照しながら回答を提示することが可能となり、特に正確性が求められる労務業務との親和性が高いと考えられる。

1.4 本研究の目的

本研究の目的は、企業の労務業務において AI RAG（Retrieval-Augmented Generation）を効果的かつ安全に活用するための方法論を明らかにすることである。労務業務は知識集約型であり、生成 AI との親和性が高いと考えられる一方で、実務レベルにおいては RAG の本格的な活用が十分に進んでいるとは言い難い。

その背景には、いくつかの構造的要因が存在する。第一に、労務業務は法令遵守やコンプライアンスが強く求められる領域であり、誤った情報提供が法的リスクや労使トラブルにつながる可能性が高い。そのため、生成 AI による自動回答に対する心理的・組織的な抵抗が大きい。第二に、労務ナレッジの多くが就業規則や各種規程、運用細則、過去の個別対応といった多様な文書や暗黙知として分散しており、RAG が前提とする「検索可能で構造化された知識基盤」が十分に整備されていない点が挙げられる。ナレッジの未整理は、RAG の精度や信頼性を低下させる要因となる。第三に、労務業務においては「一意の正解」が存在しないケースが多く、個別事情や運用判断が必要となる場面が少なくない。このような特性に対し、生成 AI をどの範囲まで適用すべきか、また人と AI の役割分担をどのように設計すべきかについて、明確な指針が確立されていないことも導入の障壁となっている。

さらに、従来の労務 DX は勤怠管理や給与計算といった基幹システムの導入が中心であり、ナレッジ活用や問い合わせ対応といった非定型業務は後回しにされてきた。この結果、RAG のような知識活用型 AI を組み込むための業務プロセス設計や運用モデルが十分に検討されてこなかった。

以上のように、労務分野における RAG 活用が進んでいない要因は、技術的制約というよりも、業務特性、ナレッジ構造、リスク認識、運用設計の不在に起因していると考えられる。

本研究では、これらの阻害要因を整理した上で、労務業務の特性に適合した RAG の活用モデルを提示し、単なる自動化ではなく、判断の質とガバナンスを高める支援技術としての AI 活用のあり方を明らかにすることを目的とする。

1.5 研究の意義と新規性

本研究の意義は、労務業務という高い専門性とリスクを伴う領域に対して、AI RAG を単なる省力化ツールとしてではなく、ガバナンスを強化する知識基盤として位置づける点にある。

従来の研究や実務事例では、生成 AI の活用効果が定性的に語られることが多かったが、本研究では労務業務特有の制約条件を踏まえたうえで、導入設計・ナレッジ構造・運用モデルを体系的に整理することを試みる。この点において、労務業務と RAG 技術を結びつけた実務的・理論的貢献が期待される。

1.6 本論文の構成

本論文は全 10 章から構成される。第 1 章では研究背景および目的を示した。第 2 章では、日本企業における労務業務の現状と課題を整理する。第 3 章では AI RAG 技術の概要を説明し、第 4 章では労務業務への適用モデルを提示する。第 5 章以降では、ナレッジ設計・導入効果・リスク・ケーススタディを通じて実務的検討を行い、第 10 章において結論と今後の課題を述べる。

第 2 章 労務業務の現状分析と課題整理

2.1 労務業務の定義と業務範囲

労務業務とは、企業が従業員を雇用し、適切に労働条件を管理・運用するために行う一連の業務を指す。具体的には、労働基準法をはじめとする各種労働関連法令への対応、就業規則および諸規程の整備・運用、勤怠管理、給与・社会保険手続き、労働時間管理、休暇制度の運用、労使トラブル対応などが含まれる。

加えて、近年では働き方改革や多様な雇用形態の普及により、テレワーク規程、副業・兼業対応、育児・介護休業制度の運用など、労務業務の対象範囲は拡大傾向にある。このように労務業務は、単なる事務処理にとどまらず、法令遵守、制度設計、従業員対応といった複数の側面を併せ持つ知識集約型業務である点に特徴がある。

2.2 企業規模別に見る労務業務の特徴

労務業務の運用形態は、企業規模によって大きく異なる。中小企業では、総務担当者や管理部門の一部として労務業務が兼任されるケースが多く、専門人材の不足や知識の偏在が課題となりやすい。一方、中堅・大企業では専任の労務担当者が配置されるものの、組織の複雑化に伴い、規程や運用ルールが多層化し、現場との情報伝達に齟齬が生じやすい。

また、企業規模が拡大するにつれて、従業員からの問い合わせ件数も増加し、労務部門が問い合わせ対応業務に多くの工数を割かざるを得ない状況が生まれる。この結果、規模の大小を問わず、労務業務は人手不足と業務負荷増大という共通の課題を抱えることとなっている。

2.3 労務担当者への問い合わせ業務の実態

企業における労務担当者への問い合わせ業務は、社員からの制度確認や手続きに関する質問を起点として発生する。これらの問い合わせは、メール、社内チャット、口頭など複数のチャネルを通じて寄せられ、その内容は有給休暇、労働時間、社会保険、育児・介護休業制度など、日常的かつ反復的なものが大半を占めている。

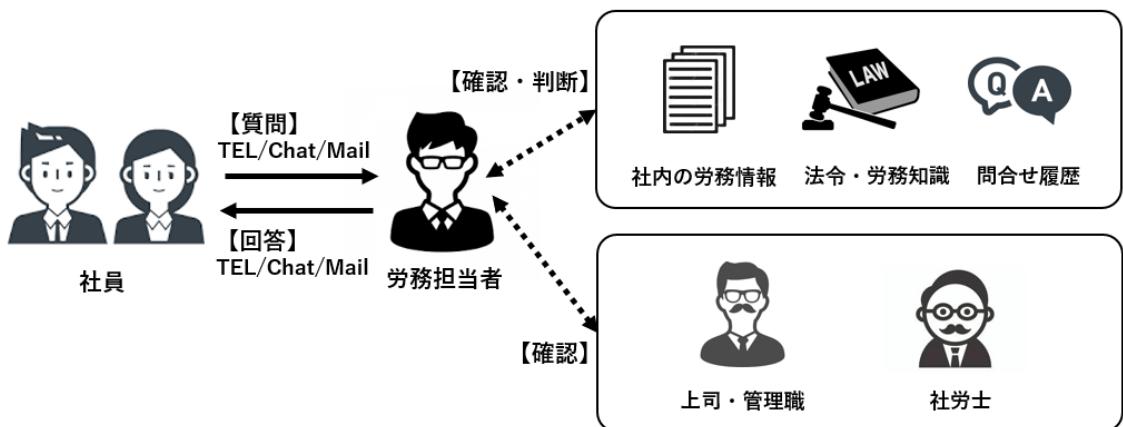


図1 労務業務フロー

図1は、日本企業における一般的な労務担当者の問い合わせ対応業務フローを示したものである。労務担当者は、問い合わせを受領した後、まず内容を整理し、関連する就業規則、各種規程、運用マニュアル、法令情報などを確認する。この段階で判断が可能な場合は、そのまま社員に回答が行われる。

一方で、労務業務は法令解釈や個別事情への配慮が求められるケースが多く、担当者自身の知識や経験のみでは判断が困難な場面も少なくない。このような場合、労務担当者は上司や管理職に確認を行ったり、社外の社会保険労務士などの専門家に助言を求めるプロセスを経ることとなる。

この確認プロセスは、回答の正確性を担保する上で重要である一方、以下のような課題を内包している。第一に、確認先が人に依存するため、即時性が確保しにくく、回答までに時間を要する点である。第二に、同様の問い合わせであっても、都度同じ確認プロセスが繰り返されることから、知識が蓄積・再利用されにくい点が挙げられる。第三に、確認結果がメールや口頭で共有されることが多く、組織的なナレッジとして残りにくいという問題がある。

その結果、労務担当者は問い合わせ対応のたびに「調査・確認・判断」を個別に行う必要があり、業務が非効率化するとともに、対応内容の属人化が進行する。この構造は、問い合わせ件数の増加に伴って労務担当者の業務負荷を指数関数的に高め、戦略的な労務施策や制度設計といった高付加価値業務への時間配分を困難にしている。

以上のように、労務担当者への問い合わせ業務は、単なる質疑応答ではなく、複数の知識源と人的判断を経由する複雑な業務フローとして成立している点に特徴がある。この構造を前提とせずに単純なFAQ化や自動応答を導入しても、根本的な課題解決には至らないことが示唆される。

2.4 ナレッジの分散・属人化問題

労務業務に関するナレッジは、就業規則や各種規程、社内マニュアル、過去のメール対応、担当者の経験知など、複数の媒体や形式で分散して管理されていることが多い。このような状態では、必要な情報に迅速にアクセスすることが困難であり、担当者個人の記憶や経験に依存した運用が常態化しやすい。

特に、法改正への対応や例外的な運用判断が必要なケースでは、過去の対応履歴や解釈が十分に共有されていないことが問題となる。結果として、ナレッジの属人化は業務効率の低下だけでなく、担当者交代時の引き継ぎ負荷や組織的なリスクを高める要因となっている。

2.5 法改正・制度変更への追随負荷

労務業務は、労働関連法令の改正や行政通達の変更に強く影響を受ける。法改正が行われるたびに、企業は就業規則や運用ルールの見直し、従業員への周知、システム設定の変更など、多岐にわたる対応を迫られる。

しかし、これらの対応は必ずしも一元的に管理されているわけではなく、規程の改訂と実際の運用が乖離するケースも少なくない。また、過去の制度と新制度が混在する移行期間においては、問い合わせ対応の難易度がさらに高まる。こうした法改正への追随負荷は、労務担当者の業務負担を増大させる大きな要因となっている。

2.6 従来のIT化・FAQ化の限界

これまで多くの企業では、労務業務の効率化を目的として、マニュアルサイトやFAQ検索サイト、社内ポータルといった情報提供型のITシステムを導入してきた。これらのシステムは、就業規則や各種制度説明を文書として集約し、社員が自ら参照することを前提として設計されており、一定の業務補助効果を果たしてきた。

しかし、これらの従来型システムは、利用者が適切な検索語を入力し、該当する情報を自ら見つけ出すことを前提とする「検索中心型」の設計思想に基づいている。労務業務においては、制度や用語自体が複雑であることから、社員が自らの疑問を適切なキーワードに変換できない場合が多く、結果としてマニュアルやFAQが存在していても利用されず、労務担当

者への直接的な問い合わせが継続的に発生する構造が温存されている。

また、FAQ 検索サイトは質問と回答を一对一で対応づける形式が一般的であるが、労務業務では雇用形態や勤続年数、個別事情、社内運用などの条件によって回答が変化するケースが少なくない。このため、あらゆるケースを想定した FAQ を事前に整備することは現実的ではなく、FAQ の増殖による検索性・保守性の低下という新たな問題を引き起こしている。さらに、法改正や社内規程の変更が発生するたびに、マニュアルや FAQ の内容を人手で更新する必要があり、その作業負荷は労務担当者に集中する。更新漏れや古い情報の残存は、誤解や労使トラブルの原因となるだけでなく、情報システム自体への信頼性低下を招く要因ともなる。

このような課題を整理すると、従来の IT 化は「情報を置いておく」ことには適している一方で、「利用者の文脈を理解し、必要な情報を取捨選択して提示する」ことは想定されていないことが分かる。すなわち、労務業務における問題は IT 導入の有無ではなく、検索前提の情報提供モデルそのものに内在する限界に起因している。

この点を明確にするため、表 1 において、従来型 FAQ・マニュアルサイトと AI RAG を労務業務への適合性という観点から比較する。

表 1 従来型 FAQ システムと AI RAG の比較（労務業務を対象として）

観点	従来型 FAQ／マニュアルサイト	AI RAG (Retrieval-Augmented Generation)
基本思想	情報を整理して「置いておく」	必要な情報を「探し出し、整理して提示する」
利用前提	利用者が適切な検索語を入力できる	利用者は自然文で質問できる
質問対応方式	Q&A の事前定義が必要	未定義の質問にも対応可能
文脈理解	文脈理解なし（キーワード一致）	文脈・意図を考慮した検索・生成
個別事情への対応	困難（FAQ の増殖が必要）	条件に応じた情報の組み合わせが可能
法令・規程の参照	該当ページを利用者が自力で確認	根拠文書を自動で抽出・提示
回答の一貫性	担当者・FAQ 作成者に依存	参照データに基づき一貫性を確保
更新対応	人手で全 FAQ を修正	元データ更新で反映可能
ナレッジ再利用	限定的（検索ヒット次第）	過去知識を横断的に再利用
労務担当者の役割	FAQ 整備・問い合わせ対応	判断・監督・高度案件対応
主な課題	検索されない、古くなる、増えすぎる	ナレッジ設計と運用ルールが重要

第 3 章 AI RAG 技術の概要

3.1 生成 AI の基本構造

生成 AI とは、大規模なデータを学習した言語モデルを用いて、人間の自然言語に近い文章を生成する技術である。近年では大規模言語モデル（Large Language Models : LLM）の発

展により、質問応答、要約、翻訳など、多様なタスクにおいて高い性能を示している。生成 AI は、入力されたテキスト（プロンプト）に基づき、学習時に獲得した統計的な言語パターンを用いて、次に続く語や文を予測・生成する仕組みを持つ。このため、自然な文章生成能力に優れる一方で、学習データに含まれない情報や、企業固有のルール、最新の法令改正といった情報を正確に反映することが困難であるという特性を有している。特に企業内業務においては、生成 AI が「もっともらしいが誤った回答」を生成する可能性がある点が課題として指摘されており、正確性や説明責任が求められる領域では慎重な利用が必要とされている。

3.2 RAG (Retrieval-Augmented Generation) の仕組み

RAG (Retrieval-Augmented Generation) は、生成 AI に検索機能を組み合わせることで、外部または内部の知識を参照しながら回答を生成するアーキテクチャである。RAG は、従来の生成 AI が抱えていた「学習データに依存する」「根拠を示せない」といった課題を補完する技術として注目されている。

RAG の基本的な構造は、質問理解 → 情報検索 → 回答生成という三層構造で整理できる。この全体像を示したものが 図 3-1 である。

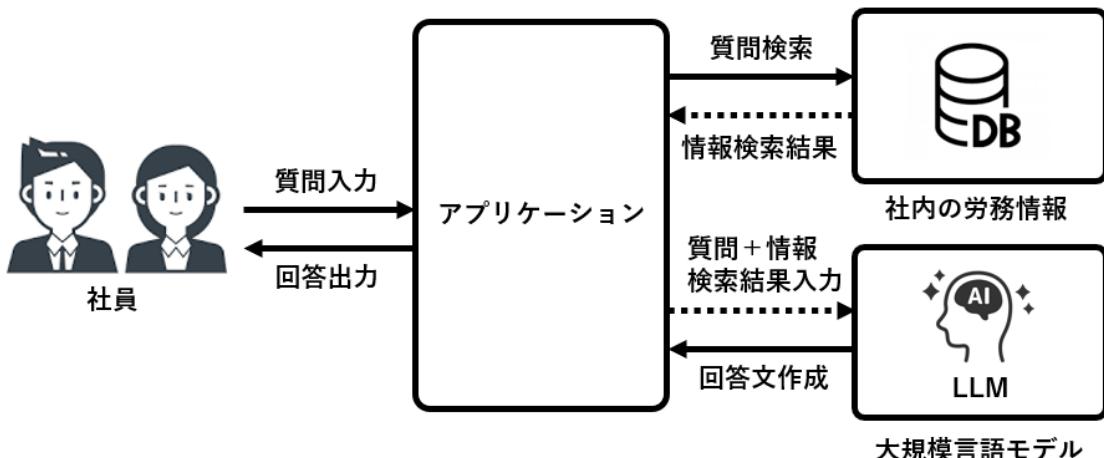


図 2 RAG (Retrieval-Augmented Generation) の概念図

図 2 に示すように、RAG は主に以下の要素から構成される。

第一に、ユーザー入力（質問）である。利用者は自然文で質問を行い、専門的な検索語や制度名称を正確に把握している必要はない。

第二に、検索（Retrieval）コンポーネントである。ここでは、社内規程、マニュアル、FAQ、法令資料、過去の問い合わせ履歴など、あらかじめ登録されたナレッジベースから、質問内容と意味的に関連性の高い情報が抽出される。この検索は、単純なキーワード一致ではなく、文脈や意味的類似性を考慮した方法で行われる点に特徴がある。

第三に、生成 AI (LLM) である。検索によって取得された情報は、生成 AI の入力として与えられ、回答の根拠として利用される。生成 AI は、新たな知識を創出するのではなく、取得した情報を整理・要約し、利用者にとって理解しやすい形で回答を生成する。

第四に、回答および根拠提示である。RAG では、生成された回答とともに、参照した文書や該当箇所を明示することが可能であり、回答の妥当性や説明可能性を確保できる。

RAG の特徴的な点

このような構造により、RAG は以下の特徴を有する。

- 企業固有のナレッジを動的に参照できる
- 最新の法令・規程情報を反映できる
- 回答の根拠を提示できるため、説明責任を果たしやすい

特に、正確性と再現性が求められる企業内業務においては、RAG は生成 AI 単体よりも実務適用性が高いと評価できる。

労務業務との接続

労務業務においては、回答そのものよりも「なぜその回答になるのか」という根拠が重要視される。RAG の概念図が示すように、検索と生成を明確に分離し、参照情報を可視化できる点は、労務領域における AI 活用の前提条件を満たすものである。

次節では、従来型の生成 AI と RAG の違いを整理し、なぜ RAG が企業内業務、とりわけ労務業務に適しているのかを比較の観点から検討する。

3.3 従来型生成 AI と RAG の違い

従来型の生成 AI は、あらかじめ学習された知識に基づいて回答を生成するため、企業固有の文脈や最新の情報を正確に反映することが難しい。一方、RAG は検索機能を通じて外部知識を動的に参照する点に特徴がある。

この違いは、企業内業務において極めて重要である。特に労務業務では、法令や社内規程といった「正解の根拠」が明確に存在するため、回答の妥当性を説明できない生成 AI は実務適用が困難である。RAG は、回答に用いた文書や条文を明示できる点で、従来型生成 AI に比べて高い説明可能性を有している。

表 2 従来型生成 AI と RAG (Retrieval-Augmented Generation) の比較

観点	従来型生成 AI (LLM 単体)	RAG (Retrieval-Augmented Generation)
基本構造	学習済みモデルのみで回答生成	検索機能 + 生成 AI を組み合わせ
参照情報	学習時点のデータに依存	外部・内部データを動的に参照
最新情報への対応	困難 (再学習が必要)	可能 (データ更新で対応)
企業固有ルールの反映	原則不可	可能 (社内ナレッジ参照)

回答の根拠提示	不可または限定的	可能（参照文書・箇所を提示）
説明可能性	低い（ブラックボックス化しやすい）	高い（根拠に基づく説明が可能）
誤回答リスク	高い（もっともらしい誤答）	相対的に低い（参照情報に依存）
文脈理解	モデル内部の推論に依存	検索+文脈理解の組み合わせ
利用者の信頼性	判断が難しい	判断・確認がしやすい
運用負荷	モデル選定・再学習が必要	ナレッジ管理・更新が中心
適した用途	創作・一般知識・アイデア生成	業務支援・規程参照・判断補助
労務業務への適合性	低い	高い

3.4 RAG における「検索」「根拠提示」「回答生成」のプロセス

RAG の処理プロセスは、大きく「検索」「根拠提示」「回答生成」の三段階に整理できる。第一に、ユーザーの質問を解析し、関連性の高い文書やデータを検索する。この際、単純なキーワード検索ではなく、文脈や意味的類似性を考慮した検索が行われる。第二に、検索によって取得された情報の中から、回答の根拠となる部分が抽出・整理される。この段階では、どの文書のどの部分が参照されたのかを明示することが可能であり、回答の信頼性を担保する役割を果たす。第三に、抽出された根拠情報を基に、生成 AI が自然言語による回答を生成する。ここで重要なのは、RAG が「新たな判断を創出する」のではなく、「既存の情報を整理し、理解しやすい形で提示する」点にある。

3.5 企業内業務における RAG の優位性

企業内業務において RAG が有する最大の優位性は、組織固有の知識を活用しながら、回答の正確性と説明可能性を両立できる点にある。

従来の検索型システムでは、利用者が情報を解釈し、判断する必要があった。一方、RAG は検索結果を単に提示するのではなく、複数の情報を統合し、文脈に即した形で提示することができる。この特性は、知識が分散し、属人化しやすい管理部門業務との親和性が高い。また、RAG は元となるナレッジを更新することで回答内容が自動的に変化するため、法改正やルール変更への追隨が容易である。この点は、更新負荷が高い従来型 FAQ との大きな違いである。

3.6 労務領域における適用可能性

労務業務は、正確性、再現性、説明責任が強く求められる領域であり、生成 AI の安易な適用はリスクを伴う。しかし、RAG を用いることで、社内規程や法令を根拠として提示しながら情報提供を行うことが可能となり、労務業務との適合性が高まる。

特に、社員からの問い合わせ対応においては、RAG を活用することで反復的な質問への対応を効率化しつつ、労務担当者は判断が必要な高度案件に集中することが可能となる。また、

上司や社労士に確認した知見をナレッジとして蓄積・再利用することで、組織としての対応力を高めることが期待される。

以上のことから、RAG は労務業務を単に自動化する技術ではなく、判断を支援し、知識を組織化する基盤技術として位置づけることができる。次章では、この特性を踏まえ、労務業務に特化した RAG 活用モデルについて検討する。

第4章 労務業務における AI RAG 活用モデル

4.1 労務業務×RAG の全体アーキテクチャ

労務業務に AI RAG を適用するにあたっては、単に生成 AI を導入するのではなく、既存の業務フローおよびナレッジ構造とどのように接続するかを明確にした全体アーキテクチャ設計が不可欠である。

労務 RAG の基本構成は、

- ①ユーザーインターフェース
- ②検索対象となるナレッジ基盤
- ③RAG 処理層
- ④人による判断・監督層

の四層構造として整理できる。社員や労務担当者は自然文で質問を行い、RAG は社内規程や法令、過去の対応履歴などから関連情報を検索・整理し、根拠を提示した回答を生成する。最終的な判断や例外対応は人が担う設計とすることで、正確性とガバナンスを両立させる。このようなアーキテクチャにより、RAG は「業務を代替する存在」ではなく、「業務判断を支援する基盤」として機能する。

4.2 社員向け問い合わせ対応への活用

企業の管理部門に寄せられる社員からの問い合わせのうち、労務・勤怠に関するものは大きな割合を占めており、管理部門業務における主要な負荷要因となっている。特に、有給休暇、各種休暇制度、勤怠処理、残業に関する質問は定型性が高く、内容の重複も多いことから、AI による支援との親和性が高い領域である。

社員向け問い合わせ対応に AI RAG を適用するにあたっては、すべての問い合わせを一律に自動化するのではなく、対応難易度に応じて適用範囲を明確に区分することが重要である。本研究では、労務関連の問い合わせを自動回答の難易度に基づき、レベル 1 からレベル 5 までの 5 段階に分類した（図 3）。

労務関連の自動回答難易度レベル



図3 労務関連の自動回答難易度

社員向け問い合わせ対応において、AI RAG の主な適用対象となるのは、レベル 1 からレベル 3 までの領域である。これらの領域では、制度の有無、基本ルール、申請方法、システム操作手順といった情報を、就業規則や社内規程、労務関連システムのマニュアルを横断的に参照しながら提示することが可能であり、RAG による自動化効果が高い。

一方で、レベル 4 およびレベル 5 に該当する問い合わせについては、AI が最終的な判断や回答を行うことは適切ではない。これらの領域では、RAG はあくまで一次的な情報整理や相談窓口の案内にとどめ、実際の対応は労務担当者や専門家が行う設計が求められる。このような役割分担を明確にすることで、誤回答やブラックボックス化といったリスクを抑制することができる。

また、社員向け RAG においては、制度説明に加え、労務システム上の確認方法や操作手順を併せて案内することが有効である。多くの社員は労務システムのマニュアルを参照しない傾向があり、「有給休暇の残日数」といった本来は自己確認可能な内容についても問い合わせが発生する。RAG がシステムマニュアルを参照しながら具体的な確認手順を示すことで、社員の自己解決率向上と問い合わせ削減の両立が可能となる。

以上のことから、社員向け問い合わせ対応における AI RAG の活用は、図 4.1 に示した自動回答難易度レベルを前提に、適用範囲を限定して設計することが重要である。この段階的な導入により、管理部門の負荷軽減と社員満足度の向上を同時に実現することが可能となり、労務 RAG 活用モデルの中核的ユースケースとして位置づけられる。

4.3 労務担当者・管理部門向け支援への活用

AIRAG の活用は、社員向け問い合わせ対応にとどまらず、労務担当者および管理部門自身の業務支援においても高い有効性を有する。労務担当者の業務は、制度説明に加え、運用・判断・教育という複数のプロセスから構成されており、それぞれ異なる性質のナレッジと支援が求められる点に特徴がある。

まず運用プロセスにおいては、休職・復職対応や長時間労働への対応など、継続的かつ手順性の高い業務が中心となる。この領域では、「休職・復職ガイド」や各種運用マニュアルといった時系列で整理されたドキュメントが重要な学習データセットとなる。RAG を用いることで、個別ケースに応じて該当手順や留意点を迅速に参照・整理することが可能となり、対応の属人化を抑制できる。

次に判断プロセスでは、安全衛生運用や労務上の判断が求められる場面が多く、単一の規程参照では対応が困難なケースが少なくない。このため、「労務判断基準集」や「労務対応事例集」といった、過去の判断や社内解釈を整理したナレッジが重要となる。RAG はこれらの資料を横断的に検索し、判断に至った根拠や類似事例を提示することで、労務担当者の意思決定を支援する役割を果たす。

一方で、法的解釈やリスク判断を伴う判断が難しい内容については、AI や社内ナレッジのみで完結させることは適切ではない。このようなケースでは、顧問社労士などの専門家とシームレスに連携することが有効となる。RAG システム内に、社労士へ問い合わせをエスカレーションできる仕組みを組み込むことで、労務担当者は追加の調査や情報整理を行うことなく、専門家の見解を迅速に得ることが可能となる。

さらに重要なのは、この専門家への問い合わせ内容と回答結果をデータとして記録・蓄積する点である。社労士から得られた回答のうち、再利用性が高く一般化可能なものについては、内容を整理した上で RAG のナレッジとして登録することで、次回以降の類似問い合わせに活用できる。この循環により、RAG は単なる検索・回答システムにとどまらず、専門家知見を内製化・資産化する仕組みとして機能する。

さらに教育プロセスにおいては、経験の浅い労務担当者や新任担当者に対する知識移転が課題となる。この領域では、「新人労務担当ガイド」や基礎的な制度解説資料を学習データセットとして整備することが有効である。RAG を活用することで、日常業務の中で生じる疑問に対して即時に参照可能な学習支援環境を提供でき、OJT に過度に依存しない教育体制を補完することが可能となる。

加えて、これら運用・判断・教育の各プロセスを横断する形で、労務担当者向け FAQ の整備も重要である。労務担当者自身が判断に迷いやすいポイントや、過去に専門家確認を要した事項を FAQ として整理し、RAG の検索対象に含めることで、調査・確認工数の削減と対応品質の均質化が期待される。

このように、労務担当者・管理部門向けの AIRAG 活用は、運用の標準化、判断支援、教育支援に加え、専門家との連携と知見の蓄積を包括的に実現する基盤として機能する。結果と

して、労務担当者は反復的な確認作業から解放され、より高度な制度設計やリスクマネジメントに注力できる体制の構築につながる。

4.4 規程・社内ルール・法令データの扱い方

労務 RAG の精度と信頼性は、参照するナレッジデータの設計に大きく依存する。労務領域においては、就業規則、賃金規程、各種運用細則、法令条文、行政通達など、性質の異なる情報が混在している。

これらのデータは、「法令」「社内規程」「運用ルール」「補足説明」といった区分で整理し、どの情報がどのレベルの根拠を持つのかを明確にして RAG に取り込む必要がある。特に、法令と社内ルールの優先関係を明示することは、誤った解釈を防ぐ上で重要である。

また、法改正時には元データを更新することで、RAG の回答も自動的に変化する設計とすることで、FAQ の個別修正に比べて運用負荷を大幅に軽減できる。

4.5 「答えを出す AI」ではなく「根拠を提示する AI」としての設計

労務業務における AI 活用において最も重要な設計思想の一つは、AI を「即断的に答えを出す存在」として位置づけるのではなく、「判断に必要な根拠を整理し提示する支援主体」として位置づける点にある。労務領域では、制度解釈や運用方法が企業ごとに異なり、かつ例外的な取り扱いも多いため、単純な正誤判定型の回答は誤解やリスクを生じやすい。

この課題に対応するため、本研究では、RAG による回答を運用ガイドの標準フォーマット（回答テンプレート）に基づいて生成する設計を提案する。このテンプレートは、RAG の回答内容を一定の型に収めることで、回答のブレを抑制し、労務担当者および社員にとって理解しやすい形で情報を提供することを目的としている。

本研究で提案する回答テンプレートは、以下の 6 要素から構成される。

第一に、結論として、質問に対する要点を一行で簡潔に示す。これにより、利用者は最初に全体像を把握でき、読み進めるか否かの判断が容易となる。

第二に、根拠として、該当する就業規則や規程、法令条文を明示する。例えば「就業規則第〇条」や「労働基準法第〇条」といった形で出典を示すことで、回答の正当性と説明責任を担保する。

第三に、運用手順として、実務上必要となる具体的な行動を整理する。申請先、提出期限、必要な証憑、システム上の操作方法などを明示することで、「分かったが次に何をすべきか分からぬ」という状態を防ぐ。

第四に、例外・注意点として、グレーゾーンや例外的な取り扱いが想定されるケースを補足する。これにより、制度の形式的理解にとどまらず、実務で問題が生じやすいポイントを事前に把握できる。

第五に、よくある勘違いを一つ提示する。これは、過去の問い合わせやトラブル事例から抽出された誤解を明示するものであり、誤運用や不要な再問い合わせを減少させる効果があ

る。

第六に、必要に応じたエスカレーション条件を明示する。社労士への確認が必要となる判断ラインや、人による最終判断が求められる条件をあらかじめ示すことで、AIの適用範囲を明確にし、過度な自動化を防止する。

このような回答テンプレートに基づく設計により、RAGは単なる質問応答システムではなく、企業内における労務判断の共通言語を形成する仕組みとして機能する。回答の構造が標準化されることで、労務担当者間の判断差や説明のばらつきが抑制され、組織としての一貫性とガバナンスが強化される。

以上のことから、労務RAGにおいては、生成される文章の自然さや網羅性以上に、「どのような構造で根拠を提示するか」という設計思想が重要であり、本研究で提案する回答テンプレートは、その中核的要素として位置づけられる。

4.6 チケット管理・業務フローとの連携モデル

労務RAGの効果を最大化するためには、単体のチャットツールとして導入するのではなく、既存のチケット管理や業務フローと連携させることが重要である。

RAGによる一次対応の結果、判断が必要な案件は自動的にチケットとして起票し、担当者や期限を設定することで、対応漏れや属人化を防止できる。また、対応完了後の判断内容をナレッジとして蓄積することで、次回以降の問い合わせ対応に再利用する循環型モデルを構築できる。

このように、RAGを業務フローの中に組み込むことで、問い合わせ対応からナレッジ蓄積までを一体的に管理することが可能となり、労務業務全体の生産性と品質向上につながる。

第5章 RAG導入におけるナレッジ設計とデータ整備

5.1 労務ナレッジの体系化の重要性

AI RAGを労務業務に効果的に適用するためには、労務ナレッジを単に網羅的に収集するのではなく、明確な目的と用途を前提として体系化することが不可欠である。目的や利用者を考慮せずにナレッジを整備した場合、検索精度や回答品質が安定せず、実務での活用が進まない要因となる。

本研究では、労務ナレッジの体系化にあたり、まず用途（誰が利用するか）という観点から整理することが重要であると考える。労務RAGの主な利用者は、①社員、②管理職、③労務担当者の三者に大別され、それぞれが求める情報の粒度や内容は大きく異なる。社員からの問い合わせでは、制度の概要や申請方法、システム操作といった分かりやすさが重視される。一方、管理職からの問い合わせでは、部下対応や判断に関わるルールの解釈や留意点が求められる。さらに、労務担当者からの問い合わせでは、法令や規程に基づく詳細な判断材料や過去事例への参照が重要となる。

次に、目的（何のために使うか）という観点からの整理も欠かせない。労務RAGにおける

ナレッジ活用の目的は、大きく二つのパターンに分けられる。一つは、参照すべき書類や根拠を明示した上で、回答案を作成することであり、主に社員向けや管理職向けの問い合わせ対応に該当する。もう一つは、判断を支援することであり、労務担当者や管理職が複雑なケースに対応する際の論点整理や選択肢提示を目的とする。

これらの用途と目的を組み合わせることで、労務ナレッジは「誰に」「何を」「どの深さで」提供すべきかが明確になる。例えば、社員向けの回答案作成を目的とするナレッジでは、Q&A形式や平易な表現が適している。一方、労務担当者向けの判断支援を目的とするナレッジでは、法令条文、通達、過去事例、専門家見解といった情報を構造化して整理する必要がある。

以上のことから、労務ナレッジの体系化は、単なる分類作業ではなく、用途（利用者）と目的（役割）を起点とした設計行為であるといえる。この視点を欠いたナレッジ整備は、RAGの性能を十分に引き出すことができず、運用の形骸化を招く可能性が高い。労務RAGの導入においては、まずナレッジ体系の設計思想を明確にした上で、段階的にデータ整備を進めることが、持続的な活用に向けた前提条件となる。

5.2 学習対象とすべきドキュメントの分類

労務RAGに取り込むドキュメントは、無差別に収集するのではなく、役割と根拠レベルに応じて分類・整理する必要がある。労務業務は制度理解とシステム運用が密接に結びついており、法令や社内規程だけでなく、実際に利用している労務関連システムの仕様や操作手順も労務業務においてAI RAGを有効に活用するためには、単に既存の規程やマニュアルを学習させるだけでは不十分である。生成AIは、労務に関する一般的な制度知識や法令の概要といった「一般教養」に相当する知識については広く学習している一方で、企業固有の運用ルールや最新の制度運用、実務上のつまずきポイントについては十分な情報を持たない。

このため、RAGで参照すべき学習対象ドキュメントを適切に定義することが重要となる。第一に、就業規則や各種規程のみでは判断できない細則が挙げられる。多くの企業では、就業規則に制度の原則が記載されているものの、実際の運用における判断基準や例外対応は、別途細則や運用メモとして管理されているケースが多い。これらの細則は、社員や新任労務担当者が最も迷いやすい領域であり、RAGにおいて優先的に学習させるべき重要なデータセットである。

第二に、生成AIが学習していない、あるいは学習が不十分な労務知識が存在する点である。例えば、残業規制など、労働法改正後の具体的運用方法や、毎年多くの社員がつまずく年末調整の申告要領・記載方法の詳細、労働災害に該当するかの判断などは、一般的な教養知識としては扱われにくい。これらの情報は、企業の実務現場に即した形で整理・定義する必要があり、RAG活用において特に価値の高いデータセットとなる。

第三に、労務担当者によるチェックにおいて差戻しが頻発する申請書類も、重要な学習対象である。差戻しが発生する背景には、社員側の理解不足や記載ミス、添付書類の漏れ等があ

り、これは制度趣旨や申請ルールが十分に伝わっていないことを示している。差戻し理由や修正ポイントをドキュメント化し、RAG の参照対象とすることで、申請品質の向上と労務担当者の確認工数削減が期待される。

第四に、会社横断的に頻出する社員からの質問が挙げられる。特定部署に限らず繰り返し寄せられる質問は、制度や運用に関する共通の理解ギャップを示しており、RAG における FAQ データセットとして整備する意義が大きい。これにより、社員の自己解決率を高めると同時に、問い合わせ対応の標準化を図ることができる。

第五に、勤怠管理システムや労務管理システムの利用方法などで難解な部分である。これらの管理システムは高機能である一方、操作手順や画面遷移が直感的でない部分も少なくない。システムの公式マニュアルだけでは理解が難しい操作ポイントや、社内特有の設定に基づく運用方法を整理したドキュメントは、RAG において高い実務価値を持つ学習対象となる。

以上のように、労務 RAG において重要なのは、生成 AI が既に保有している一般的な労務知識を再学習させることではなく、一般教養には含まれない企業固有かつ実務密着型の労務データセットを定義・整備することである。これらのデータセットを体系的に整理し RAG に取り込むことで、労務業務の属人化を抑制し、実務に即した高精度な支援を実現することが可能となる。

本研究では、労務 RAG における学習対象ドキュメントを、以下の五つに分類する。

第一に、法令・行政資料である。労働基準法、労働安全衛生法、育児・介護休業法などの関連法令、ならびに行政通達・ガイドラインは、最上位の根拠として扱われるべき情報である。

第二に、社内規程・社内ルール文書である。就業規則、賃金規程、各種運用細則、社内ガイドラインなどは、企業固有の判断基準として重要な位置を占める。

第三に、労務関連システムのマニュアル・操作仕様書である。勤怠管理システム、給与計算システム、人事管理システムなどの利用方法や設定仕様は、実務における問い合わせ内容と密接に関連している。社員や管理職からの質問の多くは、「制度として可能かどうか」だけでなく、「システム上どのように操作すればよいか」という点に集中するため、これらのマニュアルを RAG の検索対象に含めることは不可欠である。

第四に、運用マニュアル・補足資料である。法令や社内規程をどのように実務へ落とし込んでいるかを示す内部資料や、例外対応の考え方をまとめた文書は、判断を補助するナレッジとして価値を持つ。

第五に、過去の問い合わせ・対応履歴である。労務担当者が上司や社労士に確認した結果を含むこれらの情報は、実務判断の蓄積として有用であるが、法令や規程と同等の根拠として扱わないよう、位置づけを明確にした上で管理する必要がある。

このように、労務 RAG では「制度ナレッジ」と「システム運用ナレッジ」を分離しつつ統合的に扱う設計が重要となる。制度として正しい内容であっても、システム上の制約により実行できないケースも存在するため、両者を併せて参照できることが、実務に即した回答生

成につながる。

加えて、労務関連システムの利用に関しては、多くの社員が公式マニュアルを参照しないという実態も考慮する必要がある。例えば、「自分の有給休暇は何日残っているか」といった、制度やシステム上は容易に確認可能な内容についても、社員から労務担当者へ直接問い合わせが寄せられるケースは少なくない。

このような問い合わせは、社員の理解不足というよりも、「どこを、どのように確認すればよいか分からず」ことに起因している場合が多い。労務関連システムでは、有給残日数や申請状況を画面上で確認できるにもかかわらず、操作手順が直感的でない場合や、必要な画面への導線が分かりにくいことが、問い合わせ発生の要因となっている。

この点において、RAG を活用し、「システム上での確認方法そのもの」を回答として提示できることは大きな意義を持つ。RAG が、該当する画面の確認手順や操作ポイントを案内することで、社員は労務担当者に問い合わせることなく、自身で情報を確認・解決できるようになる。これは、単なる制度説明ではなく、システム利用を前提とした実務的支援であり、労務担当者の対応工数削減と社員の自己解決率向上の双方に寄与する。

以上のことから、労務 RAG における学習対象としては、制度や規程そのものに加え、社員がつまずきやすい労務システムの利用方法や確認手順を整理したデータセットを含めることが重要である。

表3 労務 RAG における学習対象ドキュメントの分類と根拠レベル

ドキュメント分類	主な内容例	根拠レベル	RAG での役割	注意点
法令・行政資料	労働基準法、育児・介護休業法、通達、ガイドライン	最上位	判断の最終根拠	改正履歴・施行日管理が必須
社内規程・社内ルール	就業規則、賃金規程、運用細則	高	企業固有の判断基準	法令との優先関係を明示
労務関連システムマニュアル	勤怠・給与・人事システムの操作・設定仕様	中	実務実行方法の補足	制度と操作の混同に注意
運用マニュアル・補足資料	社内解釈、例外対応方針、Q&A 資料	中～低	判断補助・説明補完	更新漏れ・属人化リスク
過去の問い合わせ・対応履歴	上司・社労士回答、対応メモ	参考情報	類似事例の提示	根拠として断定しない

5.3 Q&A 形式とドキュメント形式の使い分け

AI RAG の検索精度および回答品質を高めるためには、学習対象とするドキュメントの形式を適切に使い分けることが重要である。労務業務においては、Q&A 形式のデータと、規程・ナレッジ文書・マニュアル等のドキュメント形式のデータが混在しており、それぞれが果た

す役割は異なる。

本研究において当社で実施した検証結果によれば、RAG のベクトルデータベースにおける検索精度は、第 1 優先として Q&A 形式のデータ、第 2 優先としてナレッジ文書やマニュアルを参照する構成が最も有効であることが確認された。すなわち、検索時に Q&A 形式のデータを優先的にヒットさせ、その補完情報としてナレッジ文書やマニュアルを参照する設計が、回答の一貫性および実務適合性の向上に寄与する。

Q&A 形式のデータは、社員や労務担当者が実際に使用する自然文に近い形で整理されており、質問意図との一致度が高い。そのため、検索クエリとの意味的距離が短く、RAG による初期検索段階において高い精度を発揮する。一方で、Q&A 単体では背景や制度全体の文脈が不足する場合があり、誤解や例外対応への配慮が不十分となる可能性がある。

これに対して、ナレッジ文書やマニュアルは、制度の全体像や運用ルール、システム操作手順を体系的に整理した情報源であり、Q&A で示された結論や回答を補強する役割を果たす。RAG がこれらのドキュメントを第 2 優先で参照することで、回答の根拠提示や具体的な運用手順の説明が可能となり、実務上の再質問や誤運用の抑制につながる。

さらに当社の検証では、RAG の検索精度および回答品質は、汎用的な AI プロンプトを用いるだけでは十分に向上せず、労務業務に特化したプロンプト設計を行うことが不可欠であることも確認された。労務領域では、「法令」「社内規程」「運用」「例外」といった観点を意識的に区別する必要があり、これらを前提とした指示文（プロンプト）を設計することで、検索結果の選別精度および回答構造の安定性が大きく向上する。

以上を踏まえると、労務 RAG における効果的な設計としては、Q&A 形式のデータを検索の起点とし、その背後にナレッジ文書およびマニュアルを配置する多層構造に加え、労務業務に最適化された AI プロンプトを組み合わせることが重要である。この構成により、検索精度と説明責任の両立が可能となり、「迅速に理解できる回答」と「根拠に基づく説明」を同時に実現することができる。

このような Q&A 形式、ドキュメント形式、プロンプト設計の役割分担は、労務 RAG の精度向上のみならず、ナレッジ整備および運用設計の優先順位を明確にする指針としても有効であり、RAG 導入初期における実務的な設計指針となる。

5.4 メタデータ設計と検索精度への影響

AI RAG の検索精度および回答の信頼性を担保する上で、メタデータ設計は極めて重要な要素である。特に労務業務においては、参照する情報の正確性や最新性が求められるため、単に文書を学習させるだけでは不十分であり、更新を前提としたデータ管理設計が不可欠となる。

本研究における当社の労務データセットに関する調査・検証の結果、RAG で扱うデータの多くが、社内規程、運用ガイド、申請書式、法令・通達など、不定期で更新が必要となるデータセットで構成されていることが確認された。これらのデータは、制度改定や法令変更、

社内運用の見直しに伴い内容が変更されるため、古い情報を参照した場合には誤回答や誤運用につながるリスクが高い。

このような特性を踏まえると、労務 RAG におけるメタデータ設計では、詳細で複雑な属性情報を付与することよりも、更新管理を容易にし、常に最新データを参照できる仕組みを構築することが重要となる。本研究では、その実践的手法として、各データセットに書類 No (ドキュメント ID) およびバージョン番号を付与し、元データを一元管理する方法が有効であることを確認した。

具体的には、就業規則や細則、運用ガイド等の各ドキュメントに一意の書類 No を設定し、改訂のたびにバージョン番号を更新する。RAG 上では、この書類 No をキーとしてデータを管理し、最新版のみを検索・参照対象とする更新機能を実装することで、古い情報が検索結果に混在することを防止できる。

なお、検証の結果、過去バージョンのデータを RAG 内に残すことは、検索精度および回答の安定性の観点から有効ではないことも明らかとなった。古いバージョンが残存している場合、検索時に意味的に近いが内容が異なる文書がヒットし、誤った根拠提示や回答の搖らぎを生じさせる要因となる。このため、RAG においては履歴管理を主目的とせず、常に最新かつ正とされたデータのみを参照対象とする設計が望ましい。

以上のことから、労務 RAG におけるメタデータ設計の要点は、検索精度向上を目的とした属性の付与ではなく、更新管理と参照統制を中心としたシンプルな設計にあるといえる。書類 No とバージョン管理を基軸としたメタデータ設計により、RAG は常に最新の制度・運用情報に基づいた回答を生成でき、労務業務に求められる正確性と説明責任を高い水準で両立することが可能となる。

5.5 ナレッジ正規化と表現ゆれ対策

RAG を用いた業務支援においては、一般にナレッジの正規化や表現ゆれ対策が検索精度および回答品質に大きな影響を与えるとされている。特に専門用語や略語、俗称が混在する業務領域では、同義語辞書の整備や表現統一が重要な設計要素として論じられることが多い。一方で、本研究において当社が実施した労務分野の検証では、労務用語に関する表現ゆれが RAG の回答精度に与える影響は限定的であることが確認された。具体的には、労務分野で頻出する約 300 語の用語を対象に、正規表現と表現ゆれを含むデータセットを用いて検証を行った結果、現行の生成 AI は、これらの用語を一般常識レベルの知識として既に広く学習しており、検索および回答の精度に顕著な差は生じなかった。

この結果は、労務分野における多くの基本用語や表現が、生成 AI にとって新規性の高い専門知識ではなく、社会制度や一般的業務知識として学習済みであることを示唆している。したがって、「有給」「年休」「有休」や「残業」「時間外労働」といった代表的な表現ゆれについては、RAG の検索段階や回答生成段階で自然に吸収され、精度低下の要因とはなりにくい。

以上の検証結果を踏まえると、労務 RAG においては、用語正規化や表現ゆれ対策に過度な工数をかける必要はないと判断できる。複雑な同義語辞書の構築や、細かな表記統一ルールの策定は、コストに見合う効果を生まない可能性が高い。

むしろ、労務 RAG の精度向上において優先すべきは、第 5.2 節および第 5.3 節で論じたような、学習対象ドキュメントの選定、更新管理、Q&A 形式の活用、プロンプト設計といった上流設計である。これらの要素に比べると、表現ゆれ対策は相対的に影響度が低い位置づけとなる。

ただし、本節の結論は、すべての労務ナレッジに対して正規化が不要であることを意味するものではない。社内独自の略称や、特定企業内でのみ通用する用語、システム固有の項目名などについては、必要に応じて補足的な正規化を行うことが望ましい。重要なのは、表現ゆれ対策を目的化せず、影響度に応じて最小限に設計する姿勢である。

以上のことから、労務分野におけるナレッジ正規化は、RAG 設計全体の中で優先度を適切に見極め、必要最小限にとどめることが、実務的かつ合理的なアプローチであるといえる。

5.6 法令・社内規程の更新管理

労務業務においては、法改正や社内規程の変更が頻繁に発生するため、ナレッジの更新管理が不可欠である。RAG 導入後も、元となるナレッジが最新でなければ、誤った回答を生成するリスクは避けられない。

更新管理においては、文書の改訂履歴や適用開始日を明確に管理し、古い情報が参照されない仕組みを整える必要がある。また、法改正に伴う影響範囲を整理し、関連するナレッジを一括で更新できる体制を構築することが望ましい。

このように、ナレッジを「一度作って終わり」にするのではなく、継続的にメンテナンスする前提で設計することが、労務 RAG を実務で定着させる鍵となる。

第 6 章 労務 RAG 導入効果の定量・定性分析

6.1 労務業務における KPI 設計

労務 RAG 導入効果を検証するためには、従来の業務指標とは異なる観点から KPI を設計する必要がある。労務業務は直接的な売上を生まない管理業務であるため、工数削減や品質向上、リスク低減といった間接的効果を可視化する指標が重要となる。

本研究では、労務 RAG の効果測定指標を以下の三つの観点で整理する。

第一に、**効率性指標**（問い合わせ対応時間、対応件数、一次解決率など）である。

第二に、**品質指標**（再質問率、誤回答率、根拠提示率など）である。

第三に、**組織的効果指標**（労務担当者の役割変化、ガバナンス強化、社員満足度など）である。

これらの KPI を導入前後で比較することで、労務 RAG の導入効果を多面的に評価することが可能となる。

6.2 問い合わせ対応工数削減効果

労務業務における AI RAG 導入効果の一つとして、問い合わせ対応に要する工数の削減が挙げられる。ただし、RAG は導入直後から最大限の効果を発揮する仕組みではなく、スマートスタートを前提とし、運用を通じて段階的に精度を高めていくことが重要である。

初期導入段階では、RAG が対応可能な問い合わせ範囲を、有給休暇、勤怠、残業といった定型性の高い領域に限定することで、比較的短期間で一定の工数削減効果が得られる。一方で、この段階では、すべての問い合わせを完全に自動化するのではなく、AI による一次対応や回答案提示にとどめ、人による確認や補足を併用する運用が現実的である。

RAG の運用が進むにつれて、社員からのフィードバックや労務担当者による回答修正履歴が蓄積される。これらの情報をもとに、Q&A データやナレッジ文書の追加・修正を行うことで、RAG が対応可能な問い合わせ範囲が徐々に拡大し、人手を介さずに完結する問い合わせの割合が増加する。結果として、問い合わせ 1 件あたりの対応時間が短縮され、管理部門全体としての工数削減効果が段階的に顕在化する。

このように、問い合わせ対応工数削減効果は、RAG 導入の瞬間に一律に発生するものではなく、継続的な改善プロセスを通じて累積的に創出される効果として捉える必要がある。

6.3 回答品質・再質問率の変化

RAG 導入による効果は、工数削減だけでなく、回答品質の向上および再質問率の低下という質的側面にも表れる。特に労務業務では、曖昧な回答や根拠の不明確な説明が、追加質問や誤解を招く要因となりやすい。

RAG は、Q&A データ、ナレッジ文書、マニュアル等を参照しながら回答を生成するが、初期段階では回答表現や参照根拠にばらつきが生じる場合もある。このため、運用初期においては、回答内容に対する人による確認や修正を積極的に行い、その結果を学習データとして反映させることが不可欠である。

当社の検証では、回答に対する修正履歴や「分かりにくかった」「再度質問が発生した」といったフィードバックを分析し、データセットの更新や回答テンプレートの調整を行うことで、回答の一貫性と説明の明確性が徐々に向上する傾向が確認された。これに伴い、同一内容に関する再質問の発生率も低下し、社員の自己解決率が向上する。

また、回答精度に問題が生じた場合には、AI モデルそのものを変更するのではなく、参照データの見直しや更新、Q&A の追加、プロンプト設計の修正といった運用レベルでの改善が有効である点も重要である。これにより、過度なシステム改修を行うことなく、継続的に回答品質を高めることが可能となる。

以上のことから、RAG による回答品質向上および再質問率低下は、一時的な成果ではなく、フィードバックを学習させながら「育てていく」運用モデルの中で実現される効果であるといえる。労務 RAG の評価においては、導入直後の数値のみで判断するのではなく、改善サ

イクルを前提とした中長期的視点が求められる。

6.4 労務担当者の負荷軽減と役割変化

労務 RAG の導入効果は、問い合わせ対応工数の削減といった定量的側面にとどまらず、労務担当者の心理的負荷の軽減や役割の変化といった定性的側面にも表れる。労務業務は、法令遵守やリスク判断を伴う性質上、担当者が個人で判断を抱え込みやすく、特に少人数体制の企業においては、労務担当者が孤立しやすい業務領域であるといえる。

この点において、RAG は単なる業務効率化ツールではなく、労務担当者にとっての「相談役」として機能する可能性を持つ。RAG を通じて、就業規則や過去の対応事例、判断基準を即座に参照できる環境が整うことで、担当者は一人で判断を下すのではなく、「根拠を確認しながら検討する」というプロセスを自然に取ることができるようになる。この意味で、RAG はもう一人の社員、あるいは補助的な同僚として位置づけることができる。

特に当社の運用事例では、新人労務担当者にとって RAG が心強い支援ツールとなったことが確認されている。新人担当者は、制度知識や過去事例の蓄積が十分でない中で判断を求められる場面が多く、不安や心理的負担を感じやすい。RAG を活用することで、過去の類似事例や社内ルールを即座に参照できるため、上司や先輩に逐一確認する前に、自身で論点整理を行うことが可能となった。

この変化は、単に確認作業の時間短縮にとどまらず、労務担当者の役割そのものにも影響を与える。RAG が定型的な確認や一次的な情報整理を担うことで、労務担当者は例外対応や制度設計、専門家との連携といった、より高度な業務に集中できるようになる。結果として、労務担当者は「問い合わせ対応者」から「組織の労務ガバナンスを担う専門職」へと役割を進化させることができるとなる。

以上のことから、労務 RAG の導入は、業務負荷の軽減だけでなく、労務担当者の孤立を防止し、安心して判断・成長できる環境を整備するという点においても重要な意義を持つ。RAG は、労務業務の効率化ツールであると同時に、人を支える仕組みとしての価値を有しているといえる。

6.5 社員満足度・自己解決率への影響

労務 RAG の導入効果は、管理部門側の工数削減や業務効率化だけでなく、社員の満足度および自己解決率の向上という観点からも評価されるべきである。ただし、労務 RAG が社員に継続的に利用されるか否かは、回答精度が一定水準以上で担保されているかどうかに大きく依存する。

労務に関する問い合わせは、社員にとって身近かつ切実なテーマであり、回答内容の正確性や分かりやすさに対する期待値が高い。そのため、RAG による回答精度が低い場合や、曖昧な表現が多い場合には、社員は早期に利用を断念し、従来どおり労務担当者へ直接問い合わせる行動に戻る傾向がある。すなわち、労務 RAG は「使われ続けるか否か」が精度に直

結するシステムであるといえる。

この点において、RAG 導入初期からすべての問い合わせを対象とするのではなく、回答精度を担保しやすい領域に限定して提供するスマートスタートが有効である。例えば、有給休暇残日数の確認方法や勤怠申請手順といった定型性の高い質問に対して、正確かつ一貫した回答を提供することで、社員は「RAG を使えば解決できる」という信頼感を形成する。さらに、社員からの利用履歴や再質問の発生状況を分析し、回答精度に課題が見られる領域については、Q&A データやナレッジ文書の見直し、プロンプトの調整を行うことで、段階的に精度を改善することが可能である。この改善サイクルを継続することで、RAG への信頼性が高まり、社員の自己解決率も徐々に向上する。

当社の運用においても、回答精度が安定している領域では、社員が労務担当者へ問い合わせる前に RAG を利用する行動が定着し、結果として自己解決率の向上が確認された。一方で、精度が不十分な領域については利用が進まず、社員の利用定着には「まず正確に答えること」が前提条件であることが明らかとなった。

以上のことから、労務 RAG の評価においては、単純な利用回数や対応件数だけでなく、回答精度を基盤とした信頼形成と、それに伴う社員満足度および自己解決率の変化を重視する必要がある。労務 RAG は、精度を犠牲にして広範囲に展開するものではなく、信頼を積み重ねながら利用領域を拡張していく仕組みとして設計・運用されるべきである。

6.6 組織全体のガバナンス強化効果

労務 RAG 導入の最も重要な効果の一つは、組織全体のガバナンス強化である。RAG は、参照したナレッジや判断根拠を可視化するため、対応内容のトレーサビリティを確保できる。

これにより、労務対応の一貫性が向上し、属人的判断や説明不足によるリスクが低減される。また、過去の判断や対応履歴をナレッジとして蓄積することで、組織としての学習効果が生まれ、制度運用の成熟度が高まる。

このようなガバナンス強化効果は、直接的な数値化が難しいものの、内部監査対応の容易化やトラブル件数の減少といった形で間接的に評価することが可能である。

第 7 章 導入時のリスクと課題

7.1 誤回答・誤解釈リスク

労務業務に AI RAG を導入する際の主要なリスクの一つとして、誤回答や誤解釈が挙げられる。労務領域は、法令遵守や個人の権利義務に直結するため、誤った情報提供が企業リスクに直結する点に特徴がある。このリスクを低減するためには、RAG に過剰な期待を置かず、適用範囲を明確に限定した設計が不可欠である。

本研究では、誤回答・誤解釈リスクを防止するための基本方針として、RAG による自動応答の対象をレベル 1 からレベル 3 までに集中させることが重要であると考える。レベル 1

(AI による完全自動化領域) およびレベル 2 (AI 前段対応・人間確認領域) では、定型的かつ再現性の高い問い合わせが中心であり、適切なナレッジ設計と回答テンプレートを用いることで、誤回答リスクを相対的に低く抑えることが可能である。

また、レベル 3 (AI による半自動化・人手併用領域) においては、RAG は最終判断を行うのではなく、関連情報や判断材料を整理・提示する役割にとどめることが望ましい。これにより、労務担当者は RAG を判断補助ツールとして活用しつつ、人による最終確認を前提とした対応が可能となる。

一方で、レベル 4 (人間対応が必須の領域) およびレベル 5 (メンタル不調やハラスメント等、リスクが極めて高い領域) については、RAG が主体的に回答すべきではない。これらの領域では、状況の個別性や感情的要素、法的リスクが大きく、AI による自動判断は適切でない。そのため、RAG はあくまで一次的な情報整理や窓口案内にとどめ、実際の対応は人間が行う設計が正しい。

以上のように、誤回答・誤解釈リスクを抑制するためには、「どこまでを AI に任せ、どこからを人が担うか」という役割分担を明確に定義することが重要である。RAG を万能な回答者として扱うのではなく、自動応答が有効な領域に集中させることで、安全性と実務効果を両立する設計が可能となる。

7.2 法的責任と AI 回答の位置づけ

労務業務において AI RAG を活用する際には、回答内容に対する法的責任の所在を明確にし、AI の役割を適切に位置づけることが不可欠である。特に社員向けに情報提供を行う場合、AI による回答があたかも最終判断であるかのように受け取られるリスクが存在するため、設計段階から慎重な対応が求められる。

本研究では、社員向けの AI 回答については、自動応答レベル 3 までを適用範囲とすることを基本方針とする。レベル 1 からレベル 3 に該当する問い合わせは、制度説明や手続き案内、判断材料の整理といった性質を持ち、RAG が自ら判断を下すのではなく、あくまでデータセットから根拠を参照・提示する形で回答することが可能である。この設計により、AI 回答は「助言」や「参考情報」として位置づけられ、最終的な判断責任が人間側にあることを明確にできる。

一方で、RAG が根拠を示さずに断定的な結論のみを提示する場合、社員が AI 回答を公式見解として誤認するおそれがある。そのため、社員向けの回答においては、就業規則や社内ルール、法令条文といった参照元を明示する根拠提示型の回答設計を必須とすることが重要である。これにより、AI は判断主体ではなく、情報整理・提示主体としての役割に限定される。

また、AI 回答の妥当性を継続的に担保するためには、Good/Bad の評価機能を RAG に組み込むことが不可欠である。社員や労務担当者が回答内容に対して評価を付与できる仕組みを設けることで、不適切な回答や誤解を招く表現を早期に検知することが可能となる。こ

の評価結果は、Q&A データやナレッジ文書の修正、プロンプト設計の見直しといった改善活動に活用され、RAG の信頼性向上につながる。

さらに、Good/Bad 評価の蓄積は、AI 回答の品質管理だけでなく、法的リスク管理の観点からも有効である。問題のある回答がどのような文脈で発生したかを追跡できるため、事後的な説明責任や内部監査への対応も容易となる。

以上のことから、法的責任の観点においては、AI RAG を最終判断者として扱うのではなく、根拠を参照・提示する支援ツールとして位置づけることが重要である。自動応答レベルの限定、根拠提示型回答、Good/Bad 評価機能を組み合わせることで、AI 活用と法的安全性を両立した運用設計が可能となる。

7.3 ブラックボックス化への懸念

AI 活用においてしばしば指摘される課題の一つに、判断過程が不透明となる、いわゆるブラックボックス化への懸念がある。労務業務においては、判断根拠の説明責任が特に重要であるため、AI による回答がどのような情報に基づいて生成されたのかが把握できない状態は、実務上およびガバナンス上のリスクとなり得る。

この懸念に対する有効な対応策として、本研究では、RAG の適用範囲を自動応答レベル 3 までに限定する設計が重要であると考える。レベル 1 からレベル 3 に該当する問い合わせは、制度説明や手続き案内、判断材料の整理といった性質を持ち、就業規則や社内ルール、ナレッジ文書などの明示的な根拠を参照しながら回答することが可能な領域である。この範囲に限定することで、AI の回答内容と参照根拠の対応関係を可視化しやすくなり、ブラックボックス化のリスクを低減できる。

一方で、レベル 4 およびレベル 5 に該当する問い合わせは、個別性や感情的要素、法的・倫理的リスクが高く、AI による判断過程の説明が困難な領域である。これらの領域については、RAG が最終的な回答や判断を行うのではなく、一次的な情報整理や窓口案内にとどめる設計が適切である。すなわち、RAG は相談内容を整理し、対応窓口や次のアクションを提示する役割に限定し、実際の判断や対応は人間が担う。

このような役割分担により、AI が判断主体として振る舞う場面を意図的に制限することができ、結果として「AI が何を根拠に結論を出したのか分からぬ」というブラックボックス化への懸念を抑制できる。また、労務担当者や社員にとっても、AI の役割と限界が明確となり、過度な依存や誤解を防ぐ効果が期待される。

以上のことから、労務 RAG におけるブラックボックス化への懸念は、技術的な説明性の向上だけでなく、適用範囲の限定と役割設計によって管理可能な課題であるといえる。自動応答レベル 3 までに対象を絞り、レベル 4・5 では一次対応にとどめる設計は、説明責任と実務的安全性を両立する現実的なアプローチとして位置づけられる。

7.4 セキュリティ・個人情報保護への対応

労務業務では、個人情報や機微情報を扱う場面が多く、RAG 導入にあたっては情報セキュリティと個人情報保護への配慮が不可欠である。特に、問い合わせ内容や対応履歴には、特定の個人を識別できる情報が含まれる可能性がある。

RAG に取り込むナレッジやログについては、アクセス制御や権限管理を適切に設計し、必要以上の情報が参照されないようにする必要がある。また、外部の生成 AI サービスを利用する場合には、データの保存場所や学習利用の有無についても十分な確認が求められる。

これらの対策を怠ると、情報漏えいや法令違反といった重大なリスクにつながる可能性がある。

7.5 労務担当者の AI リテラシー問題

労務業務において AI RAG を有効に活用し、実際の業務成果につなげるためには、労務担当者自身が RAG システムの基本的な原理を理解していることが不可欠である。RAG は単なる自動回答ツールではなく、適切なデータ設計と継続的な運用を前提とした仕組みであるため、利用者側のリテラシー不足は、効果を十分に引き出せない要因となり得る。

具体的には、労務担当者には、どのようなデータセットが RAG の回答精度向上に寄与するのかを理解することが求められる。就業規則や規程だけでなく、Q&A 形式のナレッジ、運用ガイド、判断基準、システム操作に関する補足資料など、労務業務に即したデータを適切に選定・整理する能力が重要である。

また、Q&A データの作成方法に関する理解も欠かせない。質問文と回答文をどの粒度で整理するか、どのような表現を用いると検索精度が向上するかといった点は、RAG の性能に直接影響する。これらは AI モデル側の問題ではなく、運用設計上の知識として労務担当者が把握しておく必要がある。

さらに、RAG は一度構築して終わりではなく、データセットの継続的なメンテナンスが前提となる。法令改正や社内ルール変更への対応、誤回答が発生した際のデータ修正、古い情報の更新といった運用作業を適切に行うためには、RAG の仕組みを理解した上で判断が求められる。

加えて、社員や労務担当者自身から得られる利用後のフィードバックをどのように反映させるかという点も重要である。Good/Bad 評価や再質問の発生状況を分析し、Q&A やナレッジ、プロンプトの改善につなげる運用知識がなければ、RAG の精度向上は期待できない。最後に、労務領域に特化したプロンプトエンジニアリングの理解も、AI リテラシーの一部として位置づけられる。労務業務では、「根拠提示」「例外条件」「エスカレーション判断」といった観点を明示的に指示するプロンプト設計が不可欠であり、汎用的な指示文では十分な性能を発揮しないことが多い。

以上のことから、労務担当者に求められる AI リテラシーとは、AI を操作する技術ではなく、RAG を業務システムとして運用・改善していくための知識と判断力であるといえる。このリテラシーを組織的に育成しない限り、RAG 導入は一時的な施策にとどまり、継続的

な成果には結びつかない可能性が高い。

7.6 運用フェーズでの課題

労務 RAG の導入における課題は、初期構築段階にとどまらず、むしろ運用フェーズにおいて顕在化する。RAG は一度導入すれば自律的に性能が向上するシステムではなく、継続的な改善活動を前提とした運用設計が不可欠である。

運用フェーズにおける重要な課題の一つが、利用者からの Good/Bad フィードバックをどのように反映させるかという点である。社員や労務担当者による評価は、回答の分かりやすさや正確性、誤解を招く表現の有無を把握するための重要な指標となる。しかし、これらのフィードバックを収集する仕組みがあっても、それをデータセットの修正や Q&A の追加、プロンプトの改善に結びつける運用プロセスが整備されていなければ、RAG の品質は向上しない。

また、時間の経過に伴い、法令改正や社内ルールの変更、運用実態の変化が生じることも、労務 RAG に特有の課題である。これらの変化に対応するためには、データセットを定期的に見直し、更新を継続することが不可欠である。古い情報が残存したままでは、RAG の回答精度が低下し、社員の信頼を損なう要因となる。

さらに、RAG の回答品質は、参照データだけでなく、プロンプト設計にも大きく依存する。運用フェーズでは、実際の利用状況や誤回答の傾向を分析し、回答テンプレートや指示文の調整といったプロンプトチューニングを継続的に行う必要がある。これにより、同じデータセットを用いていても、回答の一貫性や根拠提示の明確さを向上させることが可能となる。以上のように、労務 RAG の運用フェーズにおける課題は、「利用者の声を反映し続けること」「時間の経過に応じて RAG を賢く育てていくこと」に集約される。RAG を静的なシステムとして扱うのではなく、データ更新とプロンプト調整を繰り返す動的な業務基盤として位置づけることが、長期的な成果を得るための前提条件である。

第8章 ケーススタディ（または実証分析）

8.1 中小企業における RAG 導入事例

中小企業における労務 RAG の導入効果は、企業ごとの労務管理の整備状況によって大きく左右される。本研究では、社会保険労務士へのヒアリング結果を踏まえ、RAG 導入と労務管理成熟度の関係性について分析を行った。

社会保険労務士の実務経験によれば、RAG のデータセットとなる就業規則、各種規程、細則、Q&A、ならびにそれらの運用ルールが一定水準以上で整備されている企業においては、RAG による自動回答比率が 70%以上に達する可能性があることが確認されている。これは、RAG の性能そのものというよりも、企業側における労務ルールの明確化・文書化が進んでいることの裏返しであると解釈できる。

一方で、規程や運用ルールが曖昧であったり、文書化されていない事項が多い企業では、

RAG が参照すべき根拠情報が不足するため、自動回答の割合は低くなる傾向にある。このことから、RAG の回答精度や自動回答比率は、企業の労務管理の整備状況を示す一つの指標として機能し得ると考えられる。

この視点に立つと、RAG は単なる業務効率化ツールではなく、労務管理のアセスメント手段としての側面を持つ。すなわち、RAG 導入時の回答精度や対応可能範囲を分析することで、規程整備の不足、運用ルールの曖昧さ、Q&A 不足といった課題を可視化することが可能となる。

この点において、社会保険労務士にとって RAG 導入は、新たなビジネス機会を生み出す可能性がある。RAG の導入・運用支援を通じて、企業の労務管理状況を診断し、規程整備や運用改善、Q&A 整備といったコンサルティングにつなげることができるためである。RAG は、社労士の専門知見を置き換えるものではなく、専門家の価値を可視化・拡張するツールとして機能するといえる。

以上のことから、中小企業における労務 RAG の導入は、企業側にとっては労務管理の成熟度を高める契機となり、社会保険労務士にとっては支援領域を拡張する手段となる。RAG を介した労務管理の高度化は、企業と専門家の双方にとって相互に価値を創出する取り組みであると位置づけられる。

8.2 中堅企業・大手企業での活用事例

中堅企業および大手企業における労務 RAG の活用は、組織規模の拡大に伴う労務環境の複雑化と、社員数の増加による問い合わせ対応負荷の増大を背景として、特に高い実務的価値を持つ。本研究では、大手社会保険労務士事務所へのヒアリング結果および実務事例をもとに、中堅企業から大手企業における労務 RAG の活用形態を整理した。

中堅企業では、従業員数の増加とともに、雇用形態の多様化や勤怠制度の複線化、事業部ごとの運用差異が生じやすくなる。この段階では、定型的な制度説明や申請手順に関する問い合わせが中心となるため、社員向けの労務問い合わせ対応 RAG を導入し、レベル 1~2 の領域を中心に自動化することで、労務担当者の負荷軽減が実現される。

一方、大手企業においては、労務環境がさらに高度かつ複雑化し、社員から寄せられる問い合わせの内容も、専門性や法的リスクを伴うものが多くなる傾向がある。このような企業では、単純な FAQ 対応にとどまらず、判断を要するレベル 3 程度までの領域を RAG で支援することに、大きな価値がある。

具体的には、大企業において頻出する高度な労務相談として、以下のような法的リスクを伴う判断が挙げられる。

- 解雇・懲戒の妥当性に関する論点整理（普通解雇・懲戒解雇等）
- 労働問題に関する初期判断（ハラスメント事案、労災認定の線引き）
- 長期休職者の復職可否に関する判断補助（医師意見との整合性確認）

これらのテーマは、最終判断を AI が行うべきものではないが、関連法令、過去事例、社内

ルールを整理し、論点や判断材料を構造化して提示することには高い実務価値がある。RAG がレベル 3 としてこれらの情報整理を担うことで、労務担当者や専門家は、ゼロから調査を行うことなく、迅速かつ質の高い判断に集中できる。

また、大手企業では、以下のような複雑な制度設計や運用に関する相談も多く発生する。

- 労働時間制度（フレックスタイム制、変形労働時間制、裁量労働制）の導入・変更
- 副業・兼業規定の設計および運用判断
- 就業規則改定における法令適合性の点検
- 労働基準監督署への対応が必要となる案件の整理

これらの領域においても、RAG は制度の概要説明や判断軸の整理、必要書類や手続きの提示といった役割を担うことで、労務担当者の作業工数を大幅に削減することが可能となる。以上のように、大手企業における労務 RAG の価値は、単純な問い合わせ対応の自動化にとどまらず、専門性の高い判断業務を支援する点にある。社員向けの問い合わせ対応 RAG による一次対応と、レベル 3 までの判断補助を組み合わせることで、労務部門全体の生産性向上と対応品質の平準化が実現される。

さらに、判断が難しいケースについては、RAG と社会保険労務士などの専門家がシームレスに連携することで、リスクを抑えつつ高度な対応が可能となる。以上のことから、中堅企業から大手企業にかけては、社員向け問い合わせ対応 RAG と、専門性の高い判断支援 RAG を組み合わせた多層的な活用モデルが、現実的かつ有効なアプローチであるといえる。

8.3 問い合わせ対応プロセスの Before / After

RAG 導入前の問い合わせ対応プロセスでは、社員からの質問を受けた労務担当者が、規程確認、システム操作確認、必要に応じた上司・社労士への確認を経て回答を行うため、対応に一定の時間を要していた。また、同様の質問であっても都度同じプロセスが繰り返されていた。

RAG 導入後は、社員からの質問に対して RAG が一次対応を行い、原則的な扱いや関連規程を提示する。判断が必要な案件のみが労務担当者へエスカレーションされるため、対応プロセスが分岐・整理された。これにより、対応時間の短縮と、問い合わせ対応の標準化が実現された。

8.4 労務担当者・社員の評価

労務 RAG の有効性は、システム機能そのものだけでなく、実際に利用する労務担当者および社員の評価を通じて検証されるべきである。本研究では、RAG 導入後の利用状況やフィードバックをもとに、両者の評価を整理した。

まず社員の評価においては、回答精度が利用促進の前提条件であることが明確となった。労務に関する質問は、個人の権利や待遇に直結するため、社員は回答内容の正確性や一貫性に対して高い期待を持っている。RAG による回答が不正確であったり、曖昧な表現が多かつ

たりする場合、社員は次第に RAG を利用しなくなり、従来どおり労務担当者への直接問い合わせに戻る傾向が見られた。すなわち、回答精度が一定水準を下回る RAG は、自然と利用されなくなるという評価が得られた。

また、社員の利用行動に関する重要な知見として、多くの社員は労務システムのマニュアルを自発的に参照しない傾向があることが確認された。例えば、「自分の有給休暇の残日数は何日あるか」といった、システム上で容易に確認可能な内容であっても、確認方法が分からぬ、あるいは調べる手間を避けたいという理由から、労務担当者へ問い合わせが寄せられるケースが多い。

この点において、RAG が労務システムのマニュアルや操作手順をデータセットとして参照し、「どこを見れば、どのように確認できるか」を具体的に案内できることは、社員から高く評価された。制度説明だけでなく、システム上での確認方法を含めて回答することで、社員は労務担当者に問い合わせることなく、自身で問題を解決できるようになる。この結果、自己解決率の向上と問い合わせ件数の抑制が同時に実現された。

一方、労務担当者の評価においては、RAG と人による対応がシームレスに連携しているかどうかが、システム全体の有効性を左右する重要な要素として挙げられた。RAG が一次回答を行い、その内容や参照根拠、システム案内を含めた回答履歴を労務担当者が引き継げる仕組みがあることで、対応の重複や説明の齟齬が減少し、業務効率と対応品質の双方が向上する。

以上の評価結果から、労務 RAG の利用定着には、回答精度の担保に加え、社員の行動特性（マニュアルを読まない）を前提とした設計が不可欠であることが示唆される。労務システムマニュアルを含む実務データを RAG のデータセットに組み込み、AI と人が連携する対応プロセスを構築することが、社員・労務担当者双方から評価される労務 RAG の条件であるといえる。

8.5 成功要因と失敗要因の整理

本章では、これまでのケーススタディおよび実務検証を踏まえ、労務 RAG 導入における成功要因と失敗要因を整理する。労務 RAG は技術的に高度な仕組みである一方で、その成否はシステム単体の性能ではなく、人および専門家との連携設計に大きく左右される点に特徴がある。

1. 成功要因

第一の成功要因は、RAG と労務担当者とのシームレスな連携である。RAG が一次回答を行い、その内容や参照根拠、回答履歴を労務担当者が容易に確認・引き継げる仕組みが整備されている場合、対応の重複や説明の齟齬が減少する。これにより、労務担当者は RAG を補助的なツールとして信頼し、実務の中で積極的に活用するようになる。

第二の成功要因は、RAG と社会保険労務士などの専門家とのシームレスな連携である。労務判断が難しいケースにおいて、RAG が相談内容を整理し、論点や関連情報をまとめた上

で専門家へエスカレーションすることで、相談の質と効率が向上する。さらに、専門家からの回答や助言をナレッジとして蓄積し、RAG に再学習させることで、同様の問い合わせへの対応力が継続的に強化される。

第三に、適用範囲を明確に限定した設計も成功要因として挙げられる。自動応答レベル 1～3 に対象を絞り、レベル 4・5 は人が対応するという役割分担を明確にすることで、誤回答や過度な AI 依存を防止し、安心して運用できる環境が構築される。

第四に、回答精度を最優先とした運用方針である。社員向け RAG においては、精度が一定水準を下回ると利用が定着しないため、スマールスタートと継続的な改善を前提とした運用が有効である。

2. 失敗要因

一方で、労務 RAG 導入における主な失敗要因としては、RAG を単独で完結するシステムとして設計してしまうことが挙げられる。労務担当者や専門家との連携が不十分な場合、AI の回答に対する不信感が生じ、現場で利用されなくなる可能性が高い。

また、精度が不十分な状態で利用範囲を過度に拡大することも失敗につながる。回答品質が安定しないまま全社員向けに展開すると、誤回答や再質問が増加し、RAG そのものへの信頼を損なう結果となる。

さらに、運用体制や改善プロセスを整備せずに導入することも課題である。Good/Bad 評価やフィードバックを活用せず、データセットやプロンプトの更新を行わない場合、RAG は時間の経過とともに実務との乖離が生じ、形骸化してしまう。

3. 小括

以上の整理から、労務 RAG 導入の成功は、技術の高度さよりも、RAG を中心とした「人・専門家・ナレッジ」の連携構造をいかに設計できるかにかかっているといえる。RAG を孤立した自動化ツールとしてではなく、労務業務全体を支える協働基盤として位置づけることが、持続的な成果を生む鍵となる。

第 9 章 労務業務における AI RAG の今後の展望

9.1 法改正対応の自動化可能性

労務業務における大きな負荷の一つが、法改正や行政解釈の変更への対応である。従来、法改正が行われるたびに、労務担当者は関連条文の確認、社内規程との整合性確認、影響範囲の整理、社員への周知といった多段階の対応を手作業で行ってきた。

AI RAG を活用することで、法改正対応の一部は自動化・半自動化が可能となる。具体的には、改正法令や通達をナレッジとして取り込み、既存の社内規程や運用ルールとの関連箇所を自動的に抽出・可視化することが考えられる。これにより、労務担当者は「何が変わったか」を一から調査するのではなく、「どこに影響が及ぶか」を確認する作業に集中できる。ただし、法改正の最終的な解釈や社内適用方針の決定は、人による判断が不可欠である。RAG はその判断を支援する補助的役割を担うものとして位置づけられる。

9.2 労務×他管理業務（人事・経理）との連携

労務業務は、人事評価、採用、経理処理など、他の管理業務と密接に関連している。例えば、労働時間管理は給与計算や原価管理に影響を及ぼし、休業制度は人事配置や評価制度とも関係する。

今後、AI RAG は労務領域にとどまらず、人事・経理といった他の管理業務と連携することで、横断的な業務支援基盤へと発展する可能性がある。各部門固有のナレッジを RAG で整理・接続することで、「制度としてどうか」「実務上どう処理するか」「会計上どう扱うか」といった複合的な問い合わせに対して、一貫した情報提供が可能となる。

このような連携は、管理部門全体の生産性向上だけでなく、部門間の認識齟齬や判断のばらつきを抑制する効果も期待される。

9.3 マルチボット・役割別 AI の可能性

労務業務における AI 活用は、単一の汎用ボットによる対応から、役割別に設計されたマルチボット構成へと進化する可能性が高い。例えば、社員向けの問い合わせ対応ボット、労務担当者向けの調査支援ボット、管理職向けの判断補助ボットといった形で、利用者の役割に応じた RAG を用意することが考えられる。

役割別にボットを分離することで、参照するナレッジの範囲や回答の粒度、表現を最適化でき、誤解や情報過多を防ぐことができる。また、アクセス権限の制御もしやすくなり、セキュリティ面での利点も大きい。

このようなマルチボット構成は、労務業務の複雑性を吸収しつつ、利用者ごとに適切な支援を提供する現実的なアプローチといえる。

9.4 労務業務の高度化と人の役割の再定義

AI RAG の普及は、労務業務の自動化を進めるだけでなく、労務担当者の役割そのものを再定義する契機となる。反復的な問い合わせ対応や情報検索といった業務は RAG に委ねられ、労務担当者はより高度な判断や制度設計、リスクマネジメントに注力することが可能となる。

この変化は、労務担当者に求められるスキルの変化も意味する。単なる制度知識や手続き能力に加え、AI が提示する情報を評価・判断し、組織に適用する力が重要となる。すなわち、労務担当者は「情報の提供者」から「判断と設計の専門家」へと役割を移行していくと考えられる。

9.5 AI RAG がもたらす労務業務の未来像

AI RAG の導入は、労務業務を単に効率化する手段にとどまらず、労務業務そのものの在り方を再定義する可能性を持つ。本研究で示してきた事例および分析を踏まえると、将来的な

労務業務は、AIを中心に入れ専門家、ナレッジが有機的に連携する構造へと進化していくと考えられる。

まず、RAGと労務担当者とのシームレスな連携が、労務業務の基盤となる。RAGは定型的な問い合わせへの一次対応や、判断に必要な情報整理を担い、労務担当者はその内容を踏まえて最終的な判断や例外対応を行う。この役割分担により、労務担当者は単なる問い合わせ対応者から、制度運用や組織ガバナンスを担う専門職へと役割を高度化させることが可能となる。

次に、RAGと社会保険労務士などの専門家とのシームレスな連携が、労務業務の信頼性を高める重要な要素となる。RAGが相談内容を整理し、論点を明確化した上で専門家へエスカレーションすることで、相談の質と効率が向上する。さらに、専門家の助言や判断結果をナレッジとして蓄積し、RAGに反映することで、組織全体の対応力が継続的に強化される。これらを統合した形として、将来の労務業務は、RAGを中心とした「人・専門家・ナレッジ」の連携構造に基づいて運営されるようになると考えられる。RAGは単独で完結する自動化ツールではなく、組織内外の知見を結び付けるハブとして機能し、労務判断の透明性と再現性を高める役割を果たす。

このような未来像においては、労務業務の属人化は抑制され、判断根拠や対応履歴が組織的な資産として蓄積される。結果として、企業は法令遵守と業務効率を両立しながら、労務リスクに対してより強靭な体制を構築できるようになる。

以上のことから、AI RAGがもたらす労務業務の未来とは、AIが人を置き換える世界ではなく、AIを中心に入れ専門家、ナレッジが協働することで、労務業務の質と持続性を高める世界であると結論づけられる。

第10章 結論

10.1 本研究の総括

本研究は、企業の労務業務におけるAI RAG (Retrieval-Augmented Generation) の効果的な活用方法について、実務的検証と設計思想の整理を通じて明らかにすることを目的とした。日本企業における人材不足、ホワイトカラーの低生産性、労務業務の属人化といった課題を背景に、RAGを単なる自動応答技術としてではなく、業務基盤としてどのように設計・運用すべきかを中心的な論点として論じてきた。

本研究の分析から、労務 RAG の価値は、FAQ自動化や問い合わせ削減といった表層的な効率化にとどまらず、人・専門家・ナレッジをつなぐ中核的な役割にあることが明らかとなった。特に、RAGを「答えを出すAI」ではなく「根拠を提示するAI」として設計し、労務担当者や社会保険労務士とシームレスに連携させることで、実務的な安全性と有効性を両立できることを示した点に、本研究の意義がある。

10.2 研究結果の実務的示唆

本研究から得られた実務的示唆は、以下の三点に集約される。

第一に、労務 RAG はスマールスタートと継続的改善を前提に導入すべきであるという点である。RAG は導入直後から大きな成果を求めるものではなく、定型的な問い合わせ領域（自動応答レベル 1～3）に対象を限定し、利用者のフィードバックを反映しながら精度を高めていく運用が不可欠である。

第二に、RAG 単体ではなく、人との連携を前提とした設計が成功の鍵となる点である。RAG と労務担当者のシームレスな連携により、担当者の負荷軽減や孤立防止が実現され、RAG と社会保険労務士などの専門家との連携により、高リスク案件への対応品質が向上する。この「連携構造」を設計段階から組み込むことが、実務定着の条件となる。

第三に、社員向け RAG の利用定着には回答精度が最重要であるという点である。社員は労務システムのマニュアルを積極的に参照しない傾向があり、自己確認可能な内容であっても問い合わせが発生する。RAG が制度説明とシステム操作案内を統合して提供することで、自己解決率が向上するが、回答精度が低い場合には自然と利用されなくなる。この点を踏まえた慎重な展開が求められる。

10.3 本研究の限界

本研究にはいくつかの限界が存在する。第一に、本研究の実証は、当社の検証データおよび一部の社労士事務所へのヒアリングに基づいており、すべての業種・企業規模に一般化できるとは限らない。

第二に、RAG の性能評価については、定量的指標（工数削減率、再質問率など）のモデル化を中心としており、長期的な労務リスク低減効果や組織文化への影響については、十分に検証できていない。

第三に、生成 AI 技術そのものは急速に進化しており、本研究で前提とした技術水準や設計方針が、将来的に変化する可能性がある点も留意が必要である。

10.4 今後の研究課題

今後の研究課題として、以下の点が挙げられる。

第一に、労務 RAG と人的資本経営との関係性の深化である。RAG による労務業務の高度化が、従業員エンゲージメントや組織ガバナンスにどのような影響を与えるかについて、より長期的な分析が求められる。

第二に、専門家連携型 RAG の標準モデル化である。社会保険労務士や弁護士との連携を前提とした RAG 運用モデルを体系化し、業界横断的に適用可能なフレームワークとして整理することが、今後の実務展開において重要となる。

第三に、RAG 運用人材の育成と組織設計に関する研究である。労務担当者に求められる AI リテラシーや運用スキルをどのように育成し、組織として定着させるかは、技術以上に重要なテーマである。

以上を踏まえると、AIRAG は労務業務を置き換える技術ではなく、人と専門家、ナレッジを結び付ける協働基盤として発展していくべきものである。本研究が、その実務的かつ理論的検討の一助となることを期待したい。