

# AIによる自動車部品製造会社の業務のデジタル化

何 若文\* 郑 春慧\* 黄 剑波\*

## 抄 録

自動車産業のデジタル化・スマートファクトリー化による業務効率化の進展に対応して競争力を維持するため、検索拡張生成と大規模言語モデルを連携させたAIエージェントシステムを開発した。開発したシステムは、知識検索や翻訳など8つの業務シナリオに適用した。

## 1. まえがき

自動車産業ではデジタル化・スマートファクトリー化を進めることで業務効率化が急速に進展しており、当社でもこれらの技術を実務に適用していくことが必須である。本稿ではJATCO (Guangzhou) Automatic Transmission Ltd.における事務分野の取り組みを紹介する。この分野では知識検索、翻訳、資料整理、契約審査など多岐にわたる業務があり、特定の業務だけでなく、それらを幅広く支援する必要がある。本研究では、検索拡張生成（RAG：Retrieval-Augmented Generation）と大規模言語モデル（LLM：Large Language Model）を組み合わせた方式を中核とし、前処理・後処理やプロンプトを最適化することで実業務に適用可能なシステムを構築した。

### 1.1 用語の解説

#### 1) LLM<sup>[1]</sup>

大量のデータで事前学習された大規模AIモデルで、多様な言語処理（翻訳、解析等）を実行できる。

#### 2) RAG<sup>[2]</sup>

LLMの回答精度を向上させるため、外部データベース（用語辞書、社内文書等）から関連知識をリアルタイムで検索し補強する技術である。

#### 3) 混合類似度（Hybrid Similarity）

RAGシステムの検索評価指標で、キーワード類似度とベクトル類似度の重み付き融合を求め、クエリと文書の字面・意味の総合的な適合性を示す。値は0%から100%までで、数値が高いほど総合的一致度が高い。

#### 4) キーワード類似度（Keyword Similarity）

RAGシステムの検索評価指標で、クエリと文書のキーワード重複度に基づき字面一致度を示す。値は0%から100%までで、数値が高いほどキーワードの一致度が高い。

#### 5) ベクトル類似度（Vector Similarity）

RAGシステムの検索評価指標で、クエリと文書の高次元ベクトル変換後の意味関連性を示す。値は0%から100%までで、数値が高いほど意味の一致度が高い。

## 2. 手法及び考え方

### 2.1 業務分析とコア技術選択

本研究では、社内アンケートを通じて特定された高頻度・高工数な業務シナリオに焦点を当てた。これには、知識検索、翻訳、プログラムコーディング、契約審査など8つのコア業務のシナリオ（Table 1）が含まれており、汎用的に活用される場面から部門固有の専門的応用まで、多様な実践ケースを網羅している。

技術的な実現手法として、RAGをコア検索エンジンとして採用し、文書処理とベクトル<sup>[3]</sup>検索を担当させ、高品質な知識検索を提供する。同時に、LLMを言語モデルサービスの基盤として活用し、システムのコアとなる言語理解及び生成機能を提供する。

\*JGZ情報管理科

Table 1 Matrix of departments and works

Department	Knowledge retrieval	Information organization	Program coding	Translation and Interpretation	Request and submission for approval	Contract review	Application proxy service	PC kitting
Production Engineering	R	R	R					-
Information System	A	A, R	A, R	A, R	R	R	R	A, R
Management	R	R	R		C	C	A	-
Finance	R	R	R					-
Legal Affairs	R	R	R		A	A		-
Purchasing	R	R	R					-
Manufacturing	R	R	R					-

RACI: I (Informed)  
 C (Consulted)  
 R (Responsible)  
 A (Accountable)

本システムはRAGとLLMの前処理・後処理に新しい処理手法を検討・導入し、各シナリオの最終的な出力による効果をも高める。Fig. 1に示すように、全体構成はRAGとLLMの連携による知識検索・言語処理の統合フローとなる。

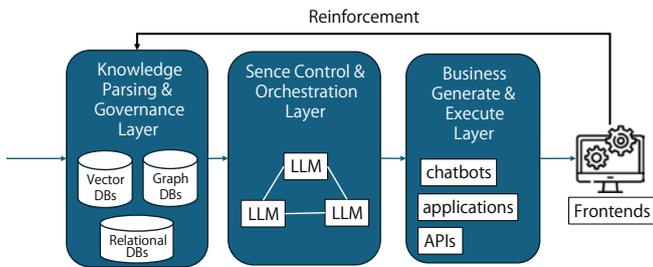


Fig1 Configuration of the system

2.2 前処理

前処理は、RAGとLLMを基盤とするAIエージェントシステムの処理精度、安定性、業務適合性を左右する基盤的な工程である。本システムが対象とする知識検索、翻訳等の業務シナリオでは、処理対象となる文書がPDF、Microsoft Word、Microsoft Excel<sup>注1</sup>といった多様なフォーマットにわたり、単一の前処理プロセスでは全ての業務ニーズを満たすことが困難である。

本システムの前処理では、「共通基盤+シナリオ特化」のアプローチ (Fig. 2) を採用し、各業務の特性に応じた前処理戦略を設計した。すなわち、文書の構造解析、意味抽出といった共通的な工程を標準化するとともに、知識検索のための戦略的な文書フォーマット変換、原文構造要素の保全と意味整合性確保のためのコンテンツの構造と再構成など専用の前処理プロセスを構築した。

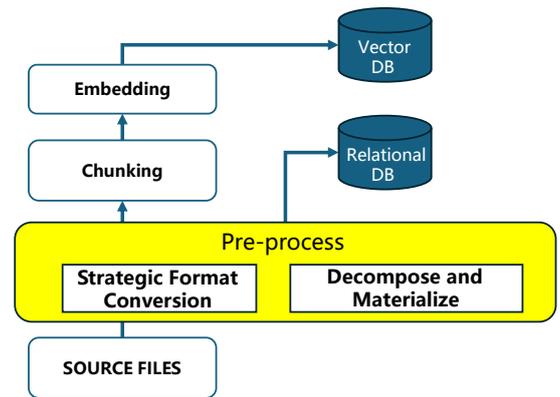


Fig.2 Pre-process of RAG

2.2.1 戦略的なフォーマット変換<sup>[8]</sup>

RAGシステムにおいて、文書解析の品質は検索効果に直接的な影響を与える。従来の方法では、元のフォーマット (特にExcelファイル) の文書を直接解析すると、深刻な「意味の断片化」と「構造情報の損失」という問題が発生する。その理由は、Excelファイルの分割が主にワークシートや固定行数に基づく機械的な処理に依存しており、セル間の行列関係やデータの意味的連携を把握できず、情報の断片化と構造的つながりの喪失を招くことにある。さらに、Excelの解析には多くの場合、軽量の処理ライブラリが用いられ、PDF解析で普及しているレイアウト分析や意味理解による多段階処理メカニズムが欠如しているため、複雑な表の階層構造やデータ間の論理関係を保全することが困難である。標準化された処理基盤を構築するため、本システムでは、まず各種文書を統一的にPDFファイルに変換することとした。PDFは、レイアウトが固定され、構造が安定しているため、文書の元のレイアウトと意味の完全性を保持する上で優位である。したがって、複雑なExcelソースファイルを直接処理するよりも解析精度が大幅に向上する。

注1 Microsoft Word, Microsoft Excelは、米国Microsoft Corporationの登録商標です。

## 2.2.2 コンテンツの構造分析と再構成

ビジネス運営のデジタル化において、知識検索と翻訳はいずれも重要なアプリケーションシナリオである。両者は文書解析と意味理解のレイヤーで技術基盤を共有するものの、Excelファイルに含まれる複雑な表、グラフ等の構造要素はフォーマットに敏感であり、汎用的な翻訳プロセスではレイアウトの崩れ、意味のゆがみ、コンテキストの断絶といった問題が発生しやすい。

この課題の解決策として、以下2つの重要なデータ処理がある。

- 1) Excelファイルにおけるテキスト要素（セル、コメント、グラフラベル等）は孤立して存在するものではなく、その位置と結合状態が重要な意味的コンテキストを構成している。そこで、Excel要素を自作Extractorにより抽出して、データベースへ格納することとした（Table 2）。なお、ここに示したものは模式的な構造であり、実際の業務シナリオにおける完全な表構造ではない。

Table 2 Structured tokens of the text

Idx	File name	Sheet name	Element type	Location	Row	Contents	Result
1	File 1	Sheet 1	Cell	\$A\$1	1	Test 1	
2	File 1	Sheet 1	Cell	\$A\$1	2	Test 2	
3	File 1	Sheet 2	Shape	Shape name	2	Test 1	

- 2) 本Extractorは、運用場面に応じて、抽出パターンのカスタマイズや、文字のフィルタリングを設定できる機能を持たせた。例えば、LLMを活用する際、翻訳精度と翻訳スピードを確保するため、単なる「○」、「ー」、数字等の翻訳不要な要素は対象外とすることが可能である。

## 2.2.3 プロンプトエンジニアリング

バッチ翻訳の対象となる原文を入力し、プロンプトを作成する。

プロンプト例：

1. Input format : `{{Idx}}和文原文;{{Idx}}和文原文;...` (Idxはインデクスの数字)
2. Output format : `{{Idx}}中文翻訳;{{Idx}}中文翻訳;...`

翻訳する文章例：

.....  
`{{1}}テスト1;{{2}}テスト2;{{3}}テスト1;`  
 .....

AIの回答例：

.....  
`{{1}}测试1;{{2}}测试2;{{3}}测试1;`  
 .....

なお、ここで示したプロンプト内容は、Table 2に基づいて作成された模式的な例であり、実際の運用シナリオにおける設定案ではない。

上記のAIによる回答から訳文を次の処理で入力するため、まず正規表現（例：`\{(\d+)\}(.*)\}`<sup>[6]</sup>）を用いて「`{{Idx}}訳文`」の固定構造をマッチさせる。この過程で、LLMのハルシネーションによる構造化破壊（例：`{{1}}测试1`（閉じ括弧欠如）、{{A}}测试1`（Idxが数字以外）、{{3}}测试3`（内容一致性の崩れ）等）を初歩的に検出する。`

次に、正規表現分析の結果に基づき「構造化破壊検証」を実施する。具体的には、検出された破壊箇所の数と種類を判定し、閾値（例：破壊箇所1件以上）を超えた場合には、LLMに対し「構造を遵守した回答の生成」を再指令する（再実行設計）。これにより、LLMのハルシネーションによる構造化破壊の悪影響を回避し、訳文抽出の精度を保証する。

用語の不一致問題を解決するため、本研究では翻訳指令生成フェーズに動的用語強化メカニズムを導入した (Fig. 3)。本システムは用語辞書をリアルタイムで検索し、識別された用語とその社内での標準的訳語を強制的な制約条件としてLLMのプロンプトに埋め込むことで、ソース側からキーワードの翻訳を固定し、全文における用語の一貫性を保証する。

最適構造したプロンプトをLLMリクエストに埋め込み、膨大な翻訳業務などに安定した出力を確保する。

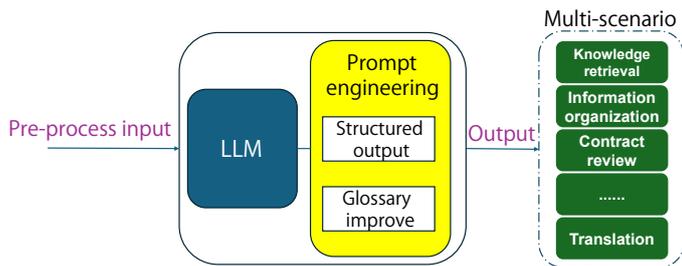


Fig. 3 Prompt engineering

## 2.3 後処理

### 2.3.1 デュアルLLMの競合最適選択<sup>[7]</sup>

単一モデルのバイアスを解決するため、本研究はデュアルLLMによる並列翻訳と自動最適選択メカニズムを採用した (Fig. 4)。本システムは訳文の「位相整合」を完了した後、用語の一致度、コンテキストの正確性、言語の流れの自然さといった多次元基準に基づいて、1項目ごとに比較・スコアリングを行い、総合的に優れた訳文を自動的に採用する。当該最適選択メカニズムは、異なるモデルの長所を効果的に集約し、翻訳品質の顕著な向上を実現した。

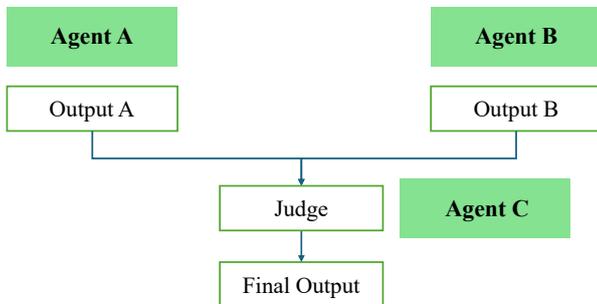


Fig. 4 Optimal selection of dual LLM outputs

### 2.3.2 画像の高精度処理

本研究はさらに文書中の画像翻訳課題を解決した。一般的な画像翻訳ツールまたはAPI (Application Programming Interface) では、元の画像の視覚的レイアウトを保持することが通常困難である。そこで精緻化された画像処理サブプロセスを設計した。そのコアステップと技術的課題は以下の通りである。

#### 1) 光学文字認識<sup>[4]</sup>、<sup>[5]</sup>及びテキストと座標の構造化

まず、Excel内の埋め込み画像を一括抽出し、OCR (Optical Character Recognition) でテキストコンテンツとそのバウンディングボックス座標 (A, B, C, D) を認識する。これにより、画像情報をテキストと座標を含む構造化データに変換し、後続の高精度処理の基盤を構築する (Fig. 5)。

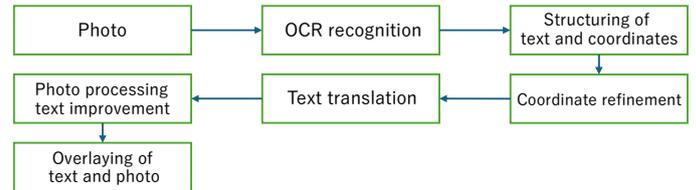


Fig. 5 Use of position information in text OCR

2) テキスト座標の高精度化キャリブレーション

OCRの認識座標に誤差が生じるという課題に対し、本研究は自動座標補正アルゴリズムを設計した。「デュアルスライダー」メカニズムを通じてピクセル単位の座標微調整を実現するとともに、アルゴリズムによってテキスト認識の有効性（例：文字分割の正確性、空白領域の合理性）を判定することで、高精度化後の最終座標（A1, B1, C1, D1）を確定し、訳文テキストボックスが画像背景と完全に一致することを保証する（Fig. 6）。

Coordinate refinement

1. Coordinates after OCR recognition: A, B, C, and D



2. Refinement process: By moving the slide pixel by pixel, the refined coordinates can be determined, based on whether the text can be recognized or not.



3. Coordinates after refinement: A1, B1, C1, and D1

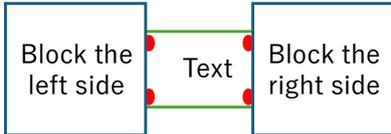


Fig. 6 Refinement of text position

3) テキスト翻訳, リライト及び画像合成

高精度座標を取得した後、本システムは認識テキストに対してLLMによる翻訳を実行する。続いてキャリブレーション後の座標に基づいて、適合するフォントで訳文をレンダリングした上で元画像に重ね合わせ、高品質な翻訳画像の生成と書き戻しを完了する。

3. 結果

2.2, 2.3で述べた前処理と後処理を本研究のAIエージェントシステム（RAGとLLMの連携）に組み合わせることで、回答の精度が大きく向上することを確認した。

3.1 前処理効果

特定のファイル形式の正規化において文書をPDFに統一した効果を評価するため、ベクトル検索を用いた比較実験を実施した（Table 3）。キーワード類似度は同一クエリ下で一致しており、これはチャンクの内容におけるキーワードの一致内容が同じであるためである。PDFチャンクは大半のケースでベクトル類似度がより高く、その語義構造をより良好に保持できることが実証されている。

Table 3 Effect of file format conversion

File Format	Average Hybrid Similarity [%]	Average Keyword Similarity [%]	Average Vector Similarity [%]
Excel	71.86	73.95	66.98
PDF	74.8	73.95	76.78

本実験では、標準プロンプトと用語強化プロンプトにおける2つの代表的なLLM AとLLM Bの性能を比較し、各グループで複数回繰り返して平均値を算出した。Fig. 7に示すように、用語を埋め込むことで用語使用率は、LLM Aで32%から98%、LLM Bで22%から86%にそれぞれ向上し、平均65%の大幅な改善が見られた。

Terminology usage before and after terminology injection

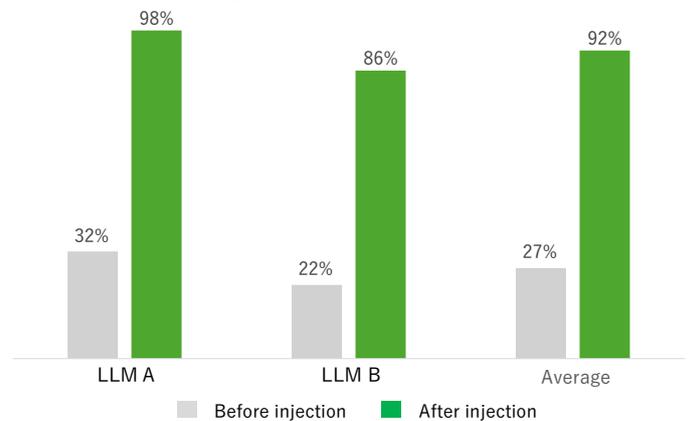


Fig. 7 Terminology injection effect

### 3.2 後処理効果

同一データセットに対し、2つの異なるLLMで個別に翻訳を実施した結果、出力間に差異が生じることを確認した。単一LLMの翻訳精度は平均96%であるが、2つのLLMの翻訳差異を判断する第三のLLMによる最適選択機構を導入することで、全体の翻訳精度を97%に向上させることが可能となった。この結果、複数のLLMによる協調メカニズムによる翻訳精度の向上が実現可能であることを示唆する。

画像高精度処理に関しては、Fig. 8に例示する中国語から日本語への翻訳において、座標精確化処理により整合性が大きく向上していることが分かる。



Fig. 8 Coordinate precision effect

最後に開発における技術文書の大量翻訳を通して、独立した精度重視型専門技術ルートを設計し、「デュアルLLMによる並列処理+意味強化（RAG）+構造化制御+結果最適選択」の混合戦略を採用し、総合テストの結果、社内用語の翻訳における一貫性が大幅に向上し、文書の構造要素が効果的に保全されると同時に、翻訳の総合的品質が顕著に改善された。従来の業務プロセスに比べ、翻訳にかかる総工数も大幅に削減された。特に画像内テキストの翻訳においては、座標ずれといった従来の課題が効果的に解消され、技術情報の正確な伝達と共有に必要な高品質な翻訳成果を安定的に提供できるようになった。

上記の手法の適用により、本システムは自動車部品製造会社の多岐にわたる事務業務において「精度向上」と「効率改善」の両方を達成した。

### 4. 考察

本システムの効率化で得られた成果は、RAGとLLMの協調メカニズムによって実現されている。具体的には、RAGが外部データベースから用語や社内知識をリアルタイムに検索し補強することで、LLMのハルシネーションを抑制し、翻訳における用語の一貫性や契約審査の精度を向上させた。同時に、知識検索業務においてはRAGによる高品質な知識リコールが、従来の手動検索に比べて大幅な時間短縮を実現した。このように、知識検索能力はAIシステムの業務適合性と処理精度を両立させる中核的な支えとなり、AI協調とその応用の実用化に不可欠な要素である。

そして、正確なセマンティック・チャンキング（Semantic Chunking）はRAGの識別精度を向上させる中核要因である。精度向上手段の一つとして、Excelの翻訳プロセスにおけるShape要素の枠内内容抽出や画像要素に含まれる内容の抽出などとデータの前処理の構造化アプローチは、他のファイル形式におけるセマンティック・チャンキング処理に新たなアイデアを提供している。

### 5. 結論

本研究の成果は、RAGによる知識検索能力とLLMによる言語処理能力の協調モデルが、自動車部品製造業の多岐にわたる事務業務に対して高い適応性と効率化をもたらす効果を発揮することを実証した。このモデルは、複雑な文書処理の自動化と品質保証を両立する解決策を提供し、自動車産業のスマートファクトリー化におけるAI技術の実用化に新たな視点を与える。

### 6. まとめ

本研究では、RAGとLLMを連携させたAIエージェントシステムによる事務領域の業務デジタル化・工数削減を実現した。当社におけるAI技術の活用は事務部門に限られず、生産現場においても同様に、画像処理と深層学習を応用して加工工程の検査自動化に取り組んでいる。

AIエージェントシステムの1年間運用結果を検証したところ、Table 4に示すように、翻訳・通訳（722時間）、知識検索（381時間）など8つの業務シナリオにおける工数削減を合算し、年間総計1,837時間の工数削減を達成した。

Table 4 Labor hour reduction effect

Core business scenarios	Reduction in man-hours (h)
Translation and interpretation	722
Knowledge retrieval	381
Application proxy service/PC kitting	243
Contract review/request and submission for approval	240
Program coding	231
Information organization	20

今後の課題としては、以下の4点を挙げる。

- 1) 本研究の前処理・後処理アーキテクチャについて、より多くの業務シナリオへの更なる応用を検討する。
- 2) 各モジュールの出力安定性と処理速度の更なる向上を図る。
- 3) LLMの技術発展に伴い、随時アーキテクチャと実装戦略の調整を行う。
- 4) 高度応用を目指し、事務代行実行のアシスタント、会議司会者、教育講師、設備予知メンテナンスアシスタント、部品設計アシスタント等のロール型エージェントを開発する。

## 7. 参考文献

- [1] Alammar, J. & Grootendorst, M. (2024). Hands-On Large Language Models. O'Reilly Media. (Official code repository: [https://gitcode.com/GitHub\\_Trending/ha/Hands-On-Large-Language-Models](https://gitcode.com/GitHub_Trending/ha/Hands-On-Large-Language-Models)).
- [2] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W. T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. *Advances in Neural Information Processing Systems*, 33 (NeurIPS 2020). arXiv Preprint arXiv:2005.11401. <https://doi.org/10.48550/arXiv.2005.11401>.
- [3] Johnson, J., Douze, M., & Jégou, H. (2019). FAISS: A Library for Efficient Similarity Search and Clustering of Dense Vectors. Research Facebook, arXiv Preprint arXiv:1702.08734. <https://doi.org/10.48550/arXiv.1702.08734>.
- [4] Cheng Cui, Ting Sun, Manhui Lin, et al. PaddleOCR 3.0 Technical Report [Online]. 2025. arXiv:2507.05595 [cs.CV]. <https://arxiv.org/abs/2507.05595>.
- [5] Cheng Cui, Ting Sun, Suyin Liang, et al. PaddleOCR-VL: Boosting Multilingual Document Parsing via a 0.9B Ultra-Compact Vision-Language Model [Online]. 2025. arXiv:2510.14528 [cs.CV]. <https://arxiv.org/abs/2510.14528>.
- [6] Jeffrey E. F. Friedl: *Mastering Regular Expressions*, 3rd Ed., Sebastopol, CA, O'Reilly Media, Inc., 2006, 515 pages. ISBN: 0-596-52812-4
- [7] Zhang, Y., Li, J., & Chen, W. 2023. JudgeLLM: Evaluating and Selecting LLM Outputs via a Specialized Judge Model. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8924–8938, Singapore. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.emnlp-main.521>
- [8] Zixuan Ke, Yucheng Wang, Yichao Zhou, Jia Li, Yutao Zhu, Yujiu Yang, Yi Liu, Ninghao Liu, Haiyun Jiang, Wayne Xin Zhao, & Zhicheng Dou. (2023). *A Survey on Table Understanding in Natural Language Processing*. arXiv Preprint arXiv:2305.13095. <https://doi.org/10.48550/arXiv.2305.13095>

### ■ 著者 ■



何 若文



郑 春慧



黄 剑波