



株式会社 LayerX



LayerX、機械学習モデル作成のアノテーション基盤に Amazon Bedrock を採用。アノテーション作業にかかる時間とコストを大幅に削減

導入効果

- **1.5 時間**：1,000 枚の書類に対するアノテーション作業の時間
- **週 3 回**：機械学習モデルの検証サイクル
- **アノテーションコストの大幅削減**
- **機械学習モデルの早期作成**
- **パーソナライズド AI-OCR の早期リリース**

ご利用中の主なサービス

- Amazon Bedrock
- AWS Batch
- Amazon S3

企業名：株式会社 LayerX

業種：IT サービス

国名：日本

従業員数：325 人（2024 年 7 月末日時点）

ウェブサイト：<https://layerx.co.jp/>

概要：法人支出管理や人的資源管理などの業務効率化クラウドサービス『バクラク』を提供する株式会社 LayerX。同社は AI が請求書や領収書の内容を読み取る AI-OCR 機能を強化した『パーソナライズド AI-OCR』の開発に Amazon Bedrock を活用することで、人力の作業では約 165 時間も要する機械学習モデルに必要なアノテーション作業を、わずか 1.5 時間にまで短縮。大幅な効率化を通じて、作業に必要な人件費を大幅に削減するとともに、データセットの早期作成を実現しました。

ビジネスの課題：新たな AI-OCR 開発におけるアノテーション作業の効率化

LayerX が提供する『バクラク』シリーズは、請求書処理や経費精算といった企業の支出管理業務を効率化するクラウドサービスです。同社は 2021 年 1 月に『バクラク 請求書受取』をリリースして以来、継続的にサービスを拡大。現在は『バクラク 経費精算』『バクラク 申請』など 7 つのサービスを提供しており、シリーズの累計導入社数は 1 万社を突破しています（2024 年 2 月時点）。

バクラクシリーズは、リリース当初から AI が領収書や請求書の内容（支払期日、支払金額、取引先名など）を読み取る AI-OCR で書類の読み取りした箇所を学習させることで、入力の手間やミスを削減する機能を拡充させてきました。しかし、利用したい書類上の情報は企業や用途によって異なるため、どうしても意図する情報を正確に読み取れないことがありました。そこで同社が着手したのが、企業や用途ごとのニーズを AI が学習し、より高い精度で情報の読み取りを可能にする『パーソナライズド AI-OCR』の開発でした。

バクラク事業部 機械学習・データ部 AI-OCR チームの深澤良介氏は、次のように話します。「例えば、企業によっては請求書の“税抜き金額”を読み取りたい場合もあれば、“税込み金額”を読み取りたい場合もあります。そこで、1 つの書類から AI がより

多くの情報を読み取り、それぞれの情報の意味を理解して、必要と考えられる情報を提案する機能を開発しました。AI が学習を重ねることで最適化が進み、より高精度な提案が可能となり、特殊な項目を読み取りたいお客様のニーズにも対応することができそうです」

バクラクの AI-OCR では、項目推定のタスクに利用する機械学習モデルを自社開発しているため、学習と検証用のデータセットが必要になります。データセットの作成には正解ラベルを付与する「アノテーション」が必須ですが、同社ではこれを手作業で行っていました。しかし、既存の AI-OCR よりも多くの項目を読み取るパーソナライズド AI-OCR では、アノテーション項目も必然的に増えるため、ここで検討したのが作業の効率化に向けた LLM（大規模言語モデル）の活用です。バクラク事業部 機械学習・データ部 AI-OCR チームの島越直人氏は、次のように振り返ります。

「従来の AI-OCR ではアノテーション項目が 4 個程度だったので、人力でもそれほど時間はかかりませんでしたが、パーソナライズド AI-OCR では項目が増えるため、時間のほかに難易度も高まります。データセットの作成だけで数か月、それらをもとにした機械学習モデルの評価にはさらに数か月の時間が必要です。そこで、LLM を活用すれば短時間で 1 次アノテーションを得ることができ、その後の検証プロセスを短縮できるのではないかと考えました。」





深澤 良介 氏
株式会社 LayerX
バクラク事業部
機械学習・データ部
AI OCR チーム



島越 直人 氏
株式会社 LayerX
バクラク事業部
機械学習・データ部
AI OCR チーム

企業概要

株式会社 LayerX

LayerX は「すべての経済活動を、デジタル化する。」をミッションに掲げ、SaaS+ Fintech を軸に、AI を中心としたソフトウェア体験を社会実装するスタートアップです。法人支出管理や人的資源管理などの業務効率化クラウドサービス「バクラク」を中心に、デジタルネイティブなアセットマネジメント会社を目指す合併会社「三井物産デジタル・アセットマネジメント」、大規模言語モデル (LLM) 関連技術を活用し企業や行政における業務効率化・データ活用を支援する「AI・LLM 事業」などを開発・運営しています。



「Amazon Bedrock を活用し、高度なパーソナライズド AI-OCR の開発を迅速に実現しました。引き続き AWS を活用し、バクラクのお客様に Wow を届けていきます。」

中川 佳希 氏

株式会社 LayerX バクラク事業部 CTO

ソリューション:高精度な LLM を利用できる Amazon Bedrock を活用

LLM の選定においては、LayerX のプロダクトの多くがアマゾン ウェブ サービス (AWS) 上で稼働していることを踏まえて、サービス連携が容易な Amazon Bedrock (Claude 3.5 Sonnet) を採用。また、島越氏は「Amazon Bedrock は日本国内の AWS リージョンで処理されるため海外へのデータ越境がないこと、LLM に Claude 3.5 Sonnet が利用できること、JSON 形式の出力でも正しいキーが抽出できることも、採用の決め手となりました」と採用の理由を説明します。

Amazon Bedrock を用いたアノテーション環境の構築は 2024 年 6 月からスタートし、最初の 1 か月ほどはアノテーション項目の検討に費やしました。

「お客様が利用するシチュエーションや書類上での解釈を考慮しながら細分化や統合を積み重ね、最終的に数十項目に絞りました。さらに、Amazon Bedrock にサンプルデータを入力し、その結果を検証しながらアノテーション部門やプロダクトチームの担当者とも議論を重ね、ラベルに落とし込む流れをチューニングしていきました」(島越氏) その後、アノテーション項目の検討と並行して、Amazon Bedrock による新しいアノテーション基盤構築にも取り組みました。「目的に合ったプロンプトエンジニアリングを実現するため、膨大なデータでもスキーマが崩れないようにスキーマ定義も投入しています。スキーマ管理には JSON Schema を採用し、階層構造を指定したうえで正しい構造化データが取得できるように工夫しました。さらに、LLM によるハルシネーションを抑制するために、妥当性をチェックするバリデーションの仕組みも導入し、推論結果でバリデーションエラーが発生した際はリトライしてからデータを投入するようにしています」(島越氏)

アノテーション基盤のアーキテクチャは、Amazon S3 上のデータを AWS Batch で Amazon Bedrock に投入するシンプルな

構成で、社内のプラットフォームチームの協力を得ながら内製で開発しました。

導入効果:1,000 枚の書類に対するアノテーションが 1.5 時間で完了

Amazon Bedrock を活用した新たなアノテーション基盤では、1,000 枚の書類に対する 1 回のアノテーションはわずか 1.5 時間ほどで完了します。仮に 40 項目のラベルを手で付与するとすれば、1 枚の書類に対して約 10 分かかり、それが 1,000 枚分で約 165 時間(仮に 4 人で、ひとりあたり 1 日 8 時間で分担すれば約 5 日間)かかる想定になります。

「結果として、人力で 165 時間かかるアノテーションを 1.5 時間にまで短縮することができました。機械学習モデルの検証も、人手の場合は 1 か月に 1 回程度行うところ、1 週間に 2~3 回できるようになりました。アノテーターの件費を考慮すると、これはコスト面での大きなメリットです」(深澤氏)

こうして開発されたパーソナライズド AI-OCR は、2024 年 11 月から『バクラク電子帳簿保存』を利用する一部のユーザーに向けてリリースされています。そして、2025 年 1 月には『バクラク電子帳簿保存』に対する提供範囲を広げ、『バクラク申請・経費精算』を一部のお客様に対しても提供を開始しました。

「リリースしてまだ間もない段階ですが、すでに大きな反響があり、従来の AI-OCR をお使いのお客様からは早く使ってみたいといった声が、新しいお客様からはバクラクシリーズを導入してみたいといった声が寄せられています」(深澤氏)

初期開発のアノテーションを終了した現在も、データ追加時のアノテーションで Amazon Bedrock を活用していますが、今後は LLM と人手のハイブリッド化を検討しています。さらには自社データを用いたファインチューニングに向けて、Claude 3 Haiku の国内リージョンでの提供にも期待を寄せています。



アマゾン ウェブ サービス ジャパン合同会社

〒141-0021 東京都品川区上大崎 3-1-1 目黒セントラルスクエア <https://aws.amazon.com/jp/>

Copyright © 2025, Amazon Web Services, Inc. or its affiliates. All rights reserved.