

2026年09月28日

各位

株式会社フィックスターズ
代表取締役社長 CEO 三木 聡
(コード番号: 3687 東証プライム)
問合せ先 取締役管理本部長 堀 美奈子
(電話番号: 03-6420-0751)

フィックスターズ、組込み向けAIモデルの移植・高速化を 従来比 1 / 3 の開発期間で実現する受託サービス開始

独自のセキュアAIエージェントを活用し、品質と性能はより高く、価格はそのまま

パフォーマンスエンジニアリング技術のリーディングカンパニーである株式会社フィックスターズ（東証プライム：3687、代表取締役社長 CEO：三木 聡）は、組込みデバイスに搭載されるAIモデルやソフトウェアの処理高速化と、各種デバイスへの移植を従来より大幅に短い期間で受託する「納期1/3 組込みデバイス向け AIモデル移植・高速化サービス」の提供を開始することをお知らせします。本サービスは、フィックスターズが独自に開発し、自社運用するセキュアなAIエージェントを使用して、高速化・移植の開発期間を大きく短縮することで実現しました。

組込みデバイス向け
AIモデル移植・高速化サービス



納期 1/3[※]

開発期間を大幅短縮。品質と性能は、より高く。

エンジニア × 自社運用のセキュアAI

無償アセスメント受付中

※ 期間・費用は開発内容により異なります。無償アセスメントにて個別にご提示します。

近年、AI推論や画像・信号処理といった計算負荷の高い処理をエッジデバイスや車載SoCへ移植するニーズが高まっています。一方でその最適化には専門知識と多くの検証工数を要するため、開発は長期化しがちです。

本サービスは、フィックスターズが20年以上にわたり培ってきたソフトウェア高速化・移植の技術と、社内AI基盤「Fixstars Vega」上でのエージェンティックエンジニアリングにより実現可能となりました。最新のオープンウェイトLLMを活用し、機密性と品質、性能を保ちながら開発期間の大幅な短縮を実現します。

「納期1/3 組込みデバイス向けAIモデル移植・高速化サービス」の特長

特長1 開発期間を従来の約1/3に短縮

高速化・移植の開発期間を、従来の約1/3に短縮します。例えば従来6ヶ月規模のプロジェクトであれば、約2ヶ月に納期を大幅短縮できる見込みです。開発期間の短縮は、製品の早期市場投入や次の開発テーマへの着手につながります。

特長2 特急料金なし、費用は従来水準

短納期でも特急料金は発生しません。開発費用は従来の受託開発と同水準です。

特長3 品質と性能は、より高く

網羅的なテストで品質を確保し、ハードウェアの限界まで性能を引き出します。テスト結果、ベンチマーク結果、設計ドキュメントまで含めて納品します。

特長4 万全の機密保持

開発に使用するAIエージェントは、フィックスターズが自社内で運用するAI基盤上で動作します。お客様から預かった仕様書、ソースコード、データが外部のAIサービスに送信されることはなく、外部AIモデルの学習に使われることもありません。

※ 短縮できる期間や費用は開発内容により異なり、無償アセスメントにより個別に提示します。

サービス概要

本サービスは、AIモデル・ソフトウェアの処理高速化と、各種組込みデバイスへの移植を対象として受託します。いずれも、実機での性能検証から納品後の継続改善まで一貫して対応します。

提供内容

- **処理の高速化**：既存システム上で処理速度を数倍に引き上げます。プロファイリングに基づき、マルチスレッド化・GPU活用・量子化などを適用します。対象は、画像・信号処理、AI推論・学習、シミュレーション、数値計算など計算負荷の高い処理全般です。
- **組み込みデバイスへの移植**：AIモデルやアプリケーションを実機へ移植し、精度を保ちながらハードウェアの性能を最大限に引き出します。移植先は、NVIDIA DRIVE / Jetson、Renesas R-Car、Qualcomm Snapdragon、Tenstorrent、各種FPGA・DSPなどです。

提供プロセス

1. **無償アセスメント**：対象処理・現状性能・ターゲット環境を確認し、短縮可能な期間と達成可能な性能を試算。
2. **見積り**：開発期間・品質基準（テスト方針・性能目標）・費用を提示し、合意のうえで着手。
3. **開発**：エンジニアとAIエージェントの協働体制で開発し、週次で進捗・成果物を共有。
4. **納品**：動作するソフトウェアに加え、テスト結果・ベンチマーク・設計ドキュメントを納品。

対象外のプロセッサや処理についても、1テーマ単位での相談・依頼を受け付けます。

詳細：

<https://www.fixstars.com/ja/products-services/embedded-software-development>

問い合わせ：<https://www.fixstars.com/ja/contact>

短納期を実現する仕組み — Fixstars Vega とエージェントックエンジニアリング

フィックスターズのエンジニアは、AIエージェントと協働して開発を行うエージェントックエンジニアリングを採用、実践しています。

また、エージェントックエンジニアリングのための環境として、社内AI基盤「Fixstars Vega」を構築し、その上で動作するパフォーマンスエンジニアリング・ハーネス[1]（高速化・移植に特化したAIエージェント実行環境）を開発し、日々改善に取り組んでいます。

Fixstars Vegaは、オープンウェイトLLM[2]を用いて社内インフラ・ネットワーク上でAI処理を完結できるように構築、運用しているAI基盤です。ハーネスには、チップごとの最適化

パターンや量子化戦略、性能計測の勘所といった、フィックスターズが長年にわたって蓄積してきた組み込みソフトウェア開発とパフォーマンスエンジニアリングのノウハウが実装されています。これにより、汎用のAIツールでは得られない、高い専門性をもった開発を高品質かつ短納期で実現することができます。

- 高速化・移植の専任スペシャリストがハーネスを日々開発・改善し、全エンジニアの開発効率を継続的に高めています。
- 設計・技術判断・コードレビューはエンジニアが実施し、AIの速度とプロの品質を両立します。
- ハーネスに実装しているのはフィックスターズ自身の技術ノウハウであり、顧客の秘密情報がAIの学習に使われることはありません。

超短納期 開発事例：自動運転AIモデルの推論高速化

従来は約2人月かかると見積られた AI推論最適化の課題に対して、エンジニアがAIエージェントを使用することで、わずか1週間で目標速度を達成できました。

工数削減効果としては、当社検証において最大となる約8分の1を実現しました。

- 開発期間：1週間（従来見積：約2人月） *1
- 開発内容：自動運転用 3D物体検出 AIモデルの推論高速化
- 処理速度：3.95倍（162.3 → 41 ms/iter）
- アプローチ：TensorRT化、PTQ[3]（重みの量子化）*2、カスタムCUDAカーネルのTensorRT Plugin 化 など

*1 所要期間は本モデルのデプロイ習熟者による実績です。

*2 精度は mAP 0.428 → 0.417（PTQ起因の微減。キャリブレーションで回復見込み。人手による最適化は0.428を維持）。

[1] ハーネス AIエージェントを実行するための環境。ここでは、高速化・移植に必要なツール群と最適化ノウハウを組み込んだ、フィックスターズ独自のAIエージェント実行環境を指します。

[2] オープンウェイトのLLM モデルの重み（パラメータ）が公開され、自社環境で運用できる大規模言語モデル。外部サービスに依存せず、社内のハードウェア上で推論を実行できます。

[3] PTQ（Post-Training Quantization）学習済みモデルの重みを、精度への影響を抑えながら低ビットに量子化し、処理を高速化・軽量化する手法。

株式会社フィックスターズについて

フィックスターズは、“Speed up your AI”をコーポレートメッセージとして掲げるテクノロジーカンパニーです。計算資源を最大限に活用するソフトウェア最適化技術を駆使し、AIモデルの推論処理と学習プロセスの両面で圧倒的な高速化を実現します。医療、製造、金融、モビリティをはじめ様々な分野で、次世代AI技術の進化を推進しています。

<https://www.fixstars.com/>

報道関係のお問い合わせ

株式会社フィックスターズ 広報担当

Email: press@fixstars.com

Tel: 03-6420-0751