

DX推進に向けたデータプラットフォームの活用—分析コード開発や機械学習プログラミングなどを汎用プログラミング言語により機能の最大化を図る

インターシステムズジャパンは、4月20日、「第4回 InterSystems 医療×ITセミナー ソリューション開発編Ⅲ」をオンライン上で開催した。同セミナーは、サービスプロバイダー向けに3回にわたって医療ITソリューション開発のための最新技術や情報を紹介するもの。その最終回となる今回は、「データ分析/データ活用」をテーマに、鳥飼幸太氏（群馬大学医学部附属病院 准教授）らが行ったが、同セミナーの講演内容の抄録を紹介する。



鳥飼幸太氏

開発しやすく、改造しやすく、保守しやすく、人材を集めやすいシステムを構築するために、システム環境に求められる条件とは何かでしょうか。

私は、システム環境に求められる条件として、①サービス提供までの時間が短い、②長期間にわたる保守・改造がしやすい、③事業継続人材を確保しやすいの3点を挙げます。

しかし、この3点は相反する傾向のある性質を持っています。例えば、コンピュータ言語（以下、言語）について考え

Embedded Pythonがオブジェクトデータベース InterOperability / Mの真価を引き出す

鳥飼幸太氏 群馬大学医学部附属病院 システム統合センター 副センター長 / 准教授

ると、①で良いと考えられるポイントとしては、システム開発の時間短縮を図るため、開発の速さを強調している新規言語を使いたいという要求が現場から上がります。

また、②については、長期間使用され続けるか不明な新規言語よりも、使い慣れた言語、すでに普及している言語など、古い言語を使いたいと考えます。

さらに③については、時代のニーズやユーザーが多く、人気のある言語、単純で扱いやすく独自性の低い言語が要求されますが、これらは必ずしも①、②での要望と合致しない場合もあります(図1)。

一方、データベースに関するシステム構築についてユーザー視点から見ると、時代の変化に対応する必要性があることから、改造することを前提としたシステム構築が求められます。また、データの抽出や分析を行う上では、快適な操作性と高速性を併せ持つことが求められますし、データの抽出や分析に関するスキルを有する人材がシステムの言語を理解する必要があります。

さらに、システムのアーキテクトを考えると、一般則として、システム構築・運用に必要な言語数が増えれば増えるほど、

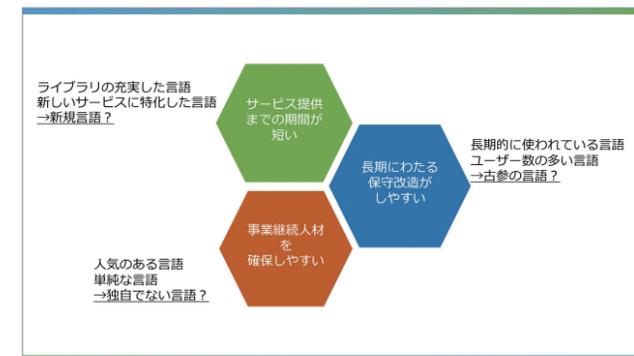


図1 システム環境には3つの条件が求められるが、それぞれが相反する性質を持っており、アプリケーション開発に用いる言語が問題となる。

「Embedded Python」が拓くデータプラットフォームの可能性

さて、データプラットフォームとして、インターシステムズの製品群が提供するサービスを大別すると、①オブジェクトデータベースの特徴を持つデータベース機能、②高いシステム間の相互運用性を実現するデータベース機能、③医療用のEMR/PHR機能、④FHIRサーバー機能の4つが挙げられます。

このうち「データベース機能である「Inter Operability」は、かつて「Ensemble」と呼ばれた機能で、ここに医療用の標準アダプタやWebテクノロジーを接続することができ、ログを取得できるのが大きな利点でした。

なお、インターシステムズのデータベース構築言語としては「ObjectScript」があります。それ故、インターシステムズの製品群を扱うためには、この「Object Script」を新たに習得する必要があります。このことが私自身「InterSystems IRIS for Health」の群馬大学への導入を躊躇させる要因でした。

しかし、新たな機能「Embedded Python」の登場により「Python」でデータベースに関するコードを記述できるようになったことで、アプリケーションもデータベースも「Python」で扱えるようになり、インターシステムズの製品群の性能を、さらに活用できるようになると考えています(図2)。

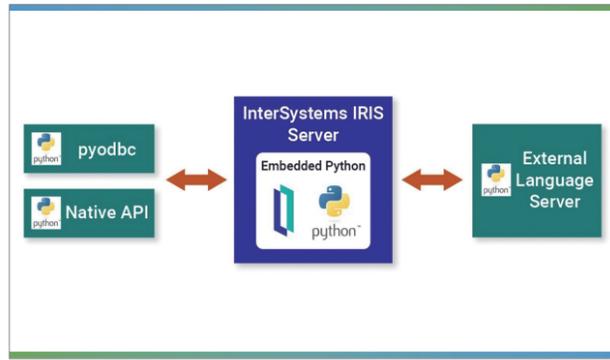


図2 「Embedded Python」機能の概念図。文法的にシンプルで扱いやすい「Python」言語で、アプリケーションもIRIS-DB操作も実現可能。

データプラットフォームとPythonの共存：データ活用の新しいアプローチ

堀田 稔氏 インターシステムズジャパン株式会社

データ分析の歴史を紐解くと、ITシステムの黎明期、データは業務を遂行するためのもので、データを分析することは当たり前ではありませんでした。しかし、時代を経てデータ分析の必要性が高まってきました。

DXという言葉が毎日のように聞かれる現在、爆発的に増加しているデジタルデータをリアルタイムに扱うには、アーキテクチャの設計や保守に高度なスキルが必要で、さらにセキュリティや情報保護などのデータガバナンス、そしてデータの正確性を担保することが求められます。このデータの生成から収集、分析・活用に至るまでの技術がデータエンジニアリングで、それを行う技術者がデータエンジニアです。

そして、データエンジニアをサポートするソフトウェアがデータプラットフォームであり、当社の「InterSystems IRIS Data Platform (以下、IRIS)」もその1つに当たります。

「IRIS」は、サーバサイドにプログラミング環境があるので、データ分析をより効率的に行うことができますが、「ObjectScript」という言語を新たに習得しなければならないという点が、問題

そのシステムでのサービスを維持することは困難になりますから、医療ユーザーが使うアプリケーションをベンダーが開発する際は、データの分析機能と、その機能を実装するためのアルゴリズムを、どの言語で記述するのが問題となります。

ECMAScriptはPythonと比べて「有用性」

その上で、コンピュ「ECMAScript」と

となっていました。

今回、新たに登場した「Embedded Python」機能は、「IRIS」のサーバサイド言語として汎用プログラミング言語である「Python」を利用できるようにする機能です。

「Python」は、統計や機械学習分野をはじめ、膨大なライブラリが公開されています。利用者の増加と、ライブラリの増加の相乗効果により、利便性の高いエコシステムが形成されますが、「Embedded Python」機能により、「IRIS」のデータプラットフォームとしての機能と「Python」のエコシステムが融合されることで、システムの大きなメリットを生み出すことができます。

また、組織づくりの観点からは、大勢存在する「Python」のプログラマーを「IRIS」のシステム開発に投入でき、人材確保しやすいというメリットが期待できます。

当社は、「Embedded Python」機能に関するセルフラーニングビデオを公開していますので、同機能の詳細を知りたい方は、ぜひアクセスしてください(https://jp.community.intersystems.com/node/520751)。