

# PIPELINE PILOT

Pipeline Pilot は高いパフォーマンスを誇るサイエンティフィック・インフォマティクス・プラットフォームです。研究における技術革新サイクルの最適化、業務効率の向上、研究およびITにかかるコストの削減を支援します。さまざまな場所に保存されている膨大な研究データへ総合的で、管理されたアクセスを提供します。また、データの科学的解析を自動化することや、組織内のユーザー全員が研究結果をすばやく検索、可視化して報告できるようにし、より広範な科学コミュニティでのコラボレーションを促進します。

Pipeline Pilot では、シンプルなデータ形式(文字や数値データ)と、複雑な科学的データ形式(画像や化学構造、遺伝子配列など)の両方を扱うことができます。幅広く奥深い科学知識とアクセルリスのパートナーエコシステムにより、医薬品、バイオテクノロジー、材料科学、一般消費財、石油およびガス、自動車、航空宇宙、エネルギー、防衛など、研究に重点をおくさまざまな業界の多様なニーズに応えます。

## PIPELINE PILOTを使用すると、次のことを実現できます。

- ・ 既存の研究や知的財産を有効利用し、定型的なデータ収集や解析作業を自動化することにより、研究開発プロジェクトを10倍以上の劇的な速さで完了
- ・ 複数の異なる研究領域における膨大なデータを、構造化の有無にかかわらず単一の環境ですばやく集計して処理することにより、データに隠された知識を引き出す
- ・ ベストプラクティスをカプセル化して配布し、コンプライアンスの徹底と研究開発組織全体におけるコラボレーションを実現
- ・ 予測科学に基づく創薬モデルの開発により、直接的な研究コストを削減
- ・ 標準的な技術と広く認められた科学知識に基づき、質の高い科学的解法をすばやく構築して展開



アクセルリスの Pipeline Pilot では、研究における技術革新のサイクルを最適化する科学的解析(緑で示した部分)用のツールを提供し、手作業による反復的なデータの準備や照合(赤で示した部分)を自動化、標準化します。これにより、科学者や技術者の能力や知識を、技術革新の促進に注ぐ時間により多く使うことができます。

- ・ リアルタイムのレポート・ダッシュボードを使用して実験結果をわかりやすく迅速に報告することにより、意思決定を支援
- ・ 互換性のない単機能のソフトウェアやデータベースを統合することにより、研究開発コストを削減

## 強力なデータ処理機能

アクセルリスのサイエンティフィック・インフォマティクス・プラットフォームである Pipeline Pilot の本質は、科学データの管理、解析、レポート作成を自動化するような、柔軟なプロトコルを、高速なアプリケーション開発環境で利用する強力なデータパイプライン・エンジンです。この優れた機能により、多くの人員とコストを費やす、非常に反復的で時に複雑な手作業による処理を削減できます。

Pipeline Pilot は、研究開発 IT 組織向けに標準化された技術を提供する柔軟なプラットフォームであり、総合的な維持コストを抑えながら科学技術革新を実現します。

強力なデータパイプライン・エンジンと科学研究に配慮した柔軟なアーキテクチャにより、PipelinePilotでは研究開発組織が直面する課題を的確に捉えることができます。Pipeline Pilotには、次のような機能があります。

- ・ 文字、数値、複雑な科学的データ(化学構造、遺伝子配列、画像など)といったデータの管理、解析、レポート作成
- ・ 要件の収集とプロトタイプ化を支援し、最終的な運用システムに直接結びつける、技術者、開発者、専門家向けの高速なアプリケーション開発環境
- ・ データの取得、操作、計算によるフィルタリング、表示を行う柔軟なコンポーネント
- ・ SOAP/Webサービス、Perl、Java、コマンドラインアクセスなどの標準的な技術を使用した、コンポーネントの「独自開発」機能
- ・ 柔軟性のあるサービス志向のアーキテクチャ
- ・ さまざまな科学的専門分野を網羅した、実績ある科学コンポーネントとベストプラクティス・ワークフロー
- ・ Microsoft SharePoint や Microsoft Office アプリケーション、化学物質登録データベース、画像データ管理システム、また、PubMed、PubChem、特許データベース、RCSB(遺伝子配列データおよびタンパク質構造)、BLAST(配列データベース)などのオンラインデータソースといった、サードパーティ製のアプリケーション、データベース、既存の科学インフラストラクチャの統合。PipelinePilotのオープンアーキテクチャは、ODBCやJDBC、Webサービスに準拠した、さまざまなサードパーティ製データソースへのアクセスを実現します。

## 異なるデータやアプリケーションの柔軟な統合

Pipeline Pilotを使用すると、異なる種類のデータを、組織全体でより適切な意思決定を行うために必要な、整理された情報に変換できます。多くの標準的な技術を使用して、社内のいたるところにある内部データベースやファイル、機器から、また、外部ソースからもデータを収集できます。一般的なフォーマットで作成されている各種データ(化学、配列、文字、画像、数値)を読み取ることができ、複数のソースのデータをリアルタイムに解析できます。効率的なオンデマンド型のデータ処理やレポート作成を行うために、Pipeline Pilot の抽出、変換、ロード機能を使用してデータマートを動的に作成できます。

## ベストプラクティスの獲得と展開

PipelinePilotでは、組織のベストプラクティスをカプセル化し、注釈をつけてバージョン管理できるため、特定の結果を得た過程を記録、再現することができます。作成したプロトコルは公開して他の人と共有し、共同開発や知識の共有に役立てることができます。自動化したプロセスをさらに大きなコミュニティで利用するには、ブラウザベースのPipeline Pilot Web Portインターフェース、SharePointなどのポータル技術など、さまざまな展開環境を使ってPipeline Pilotプロトコルを公開します。

## カスタマイズしたレポート作成とWEBアプリケーション

データ解析やデータマイニングの手続きをさまざまなテキスト、表、グラフ、画像を使用してまとめた、カスタムのレポートを作成できます。レイアウトやコンテンツを自由にコントロールできるため、研究結果の説明や伝達を簡単に行えます。単一のレポートに複数の表、グラフ、画像を表示できるので、異なるソースのデータや異なる方法で処理したデータを並べて比較するなど、データを異なる視点から見ることができます。また、レポートは、HTML、PDF、Word、Excel、PowerPointなど、さまざまな形式で配布できます。

リンクをレポートの内部、レポート間、あるいはレポートの外部に設定して、対話型の動的なレポートを作成できます。また、複数のプロトコルをリンクして機能ユニットにまとめるWebアプリケーションを作成すると、対話性を更に高めることができます。このような操作を行うとプロトコルの汎用性が飛躍的に高まり、エンドユーザーがプロトコルを作成せずにデータを入力したりデータの表示方法を決定したりできるようになります。

## PIPELINE PILOTクライアント

### Pipeline Pilot Professional Client

PipelinePilotProfessionalClientは、PipelinePilotの機能を余すところなく利用できる、データパイプライン・プロトコルを作成するオーサリングツールです。Professional Clientでは、自分で使用するためのプロトコルを作成したり変更したりできるほか、他の人が使えるようにプロトコルを公開することもできます。また、組織のさまざまなニーズに合わせて、新しいコンポーネントを作成することや、既存のコンポーネントを編集することもできます。統合処理の設定や、外部ツールとの統合を行うコンポーネントの作成により、サードパーティのデータソースやアプリケーションへのアクセスを標準化して、より広範なユーザーグループによる利用を実現できます。

### Pipeline Pilot Web Port

Pipeline Pilot Web Portは、より大規模なユーザーコミュニティでPipelinePilotを活用するための、ブラウザベースの環境です。Pipeline Pilot Professionalのユーザーが作成、実行するものと同じプロトコルをPipeline Pilot Web Portユーザーがアクセスできるので、シンプルなWebベースのインターフェースを使って、Pipeline Pilotの機能のすべてをより多くのユーザーに公開できます。Web Portユーザーは、利用可能なプロトコルを選択し、パラメータを設定することで、ニーズに合わせて各自のデータを処理することができます。その際、プロトコルがどのようにして作成されたかを意識する必要はありません。また、WebPort ユーザーはMicrosoftSharePoint、RSSフィード、Webリンク、Pipeline Pilot クライアントの SDK で作成されたカスタムのインターフェースなど、Webベースのインターフェースであればどのようなものでもプロトコルのアクセスや実行に使用できます。

## PIPELINE PILOTのコンポーネントコレクション

PipelinePilotのコンポーネントコレクションはサイエンティフィック・インフォマティクス・プラットフォームの構成要素であり、科学的なカテゴリや機能でグループ化されています。これらのコレクションには、研究者や開発者、IT担当者が行う科学的、一般的なデータ処理を支援する多数のコンポーネントが含まれています。コンポーネントをグラフィカルに組み合わせることで、データの検索、フィルタリング、分析、レポート作成のワークフローを作成できます。

機能	Professional Client	Pipeline Pilot Web Port
他の人が公開したプロトコルの使用	✓	✓
既存のプロトコルの実行	✓	✓
プロトコルパラメータの編集	✓	✓
プロトコルの実効状況の監視	✓	✓
プロトコルのレイアウトとロジックの表示	✓	
新規プロトコルの作成	✓	
既存のプロトコルの変更	✓	
ジョブと結果の管理	✓	
新規コンポーネントの作成	✓	
コンポーネントのインターフェースの編集	✓	
外部ツールの統合	✓	
スクリプト言語の導入 (PilotScriptなど)	✓	
他の人が使用するプロトコルの公開	✓	
スケジュール設定のためのコマンドラインアクセス	✓	

## 化学

### Chemistry Component Collection

化学的な知見に基づいたフィルタ機能と学習機能を用いて、データベースの化合物を分析、プロファイル、管理を行います。ベンダーのフラットファイルと直接化合物を比較し、計算値などの適切なプロファイルを用い、新しい分子を同定します。反応やMarkushライブラリをエニュミレーション解析します。ケミストリーファミリーから、共通する核や官能基のヒットリストを特定します。Pipeline Pilot とサードパーティのプロセスをシームレスに使用してすべての分析を自動化し、化学分析手順を実行します。

## ADMET Component Collection

合成候補物、ベンダーライブラリ、スクリーニング化合物のような分子コレクションに対して、吸収、分布、代謝、排泄、毒性 (ADMET) の特性予測を計算します。コレクションには、ヒト腸管吸収、水溶性、脳関門浸透性、血漿タンパク結合、チトクローム P450 2D6阻害性、肝毒性に対するモデルが含まれます。

## Cheminformatics Component Collection

Cheminformatics Collectionは、データベースの検索や閲覧を行うアプリケーションをPipeline Pilotで簡単に作成できるように設計されています。このコレクションは、データベースの検索と結果リストの管理システムであるList Management and Query Services (LMQS)と、化学構造や化学反応をOracleデータベースに保存したり登録、検索するための高性能データカートリッジであるPipeline Pilot Chemistry Cartridge という2つのサブシステムで構成されています。これにより、Pipeline Pilotの機能が強化され、さまざまな大規模ケムインフォマティクス・アプリケーションを作成する強固な基盤を提供します。

## 生物学

### Gene Expression Component Collection

個別の標的遺伝子に関する、遺伝子発現実験の表示、解釈、アノテーション、レポート作成を行います。コアとなる機能は、オープンソースのゲノム解析のBioConductorに基づいています。Pipeline PilotのGUIを使うことで、プログラムを書くことなく、複雑なワークフローを作成できます。配列解析、テキスト解析、レポート作成など、他のPipeline PilotベースのプロセスとGene Expression Collectionの解析を連動できます。

### Sequence Analysis Component Collection

モジュール方式のツールをグラフィカルに結合して、実際のバイオインフォマティクス・ワークフローを構築し、配列解析、アノテーション、配列比較を行います。配列解析用の100を超えるコンポーネントとして、BLAST、ClustalWなどの使い慣れたアルゴリズムを使用します。JavaやPerlで新しいコンポーネントを作成したり、EMBOSSと統合するなど、システムを拡張できます。

### Mass Spectrometry for Proteomics Component Collection

Mass Spectrometry for Proteomics Collectionは、カスタムの

プロテオミクスおよびメタボロミクスのワークフローを作成して自動化するための包括的なコンポーネントセットとサンプルプロトコルを提供します。このプロトコルでは、処理ステップを把握し、コンポーネントに実験データを流して、結果の読み取り、書き込み、可視化、操作、解析、比較、公開を行います。データモデルは大容量ファイルの迅速な処理をサポートし、複数のレベルで詳細情報を保持します。Pipeline Pilotのグラフィカルなプロトコル作成機能を利用すれば、ペプチドとタンパク質の同定や差異のプロファイリングといった複雑なデータ分析処理を、コードを書くことなく簡単に作成できます。

## ライフサイエンスのモデリング およびシミュレーション

### Catalyst Component Collection (Pharmacophore)

ファーマコフォアモデリングと3Dデータベースの管理を行う包括的なツールセットです。Catalyst®の著名な技術をPipeline Pilotプラットフォーム上で、ファーマコフォアモデリングと解析を効率化する、自動化した使いやすいワークフローを作成します。Catalystの洗練されたアルゴリズムを使用したコンフォメーションの概算3Dデータベースの作成、ファーマコフォア仮説の作成、仮想スクリーニングなどを実行します。

### CHARMm Component Collection (Simulations)

十分に実証されたCHARMmエンジンを基盤とする強力なコンポーネントセットを使用して、生体分子のシミュレーションを行います。このコンポーネントコレクションには、タンパク質、核酸、低分子、タンパク質-リガンド複合体に対する、安定して精度の高い分子力学と分子運動のシミュレーションを行う機能がPipeline Pilotの標準機能に追加されています。

## 材料のモデリングおよびシミュレーション

### Materials Studio Component Collection

Pipeline Pilot データパイプライン環境で、Materials Studio Component Collectionを使用すると、Reflex や Reflex QPA などの分析機能、QSAR Plusパッケージの主要な機能、Materials Studio®のスクリプト記述APIで作成されたスクリプトをPipeline Pilotプロトコルに統合するユーティリティなどMaterials Studioの主要なモデリング機能を使用できます。

## Polymer Properties Component Collection (Synthia)

Pipeline Pilot Polymer Properties Collectionは、大量のアモルファスホモポリマーやランダムコポリマーを、繰り返し単位情報、分子量、温度から迅速に推定する手段を提供します。このコレクションには、J.Biceranoの Prediction of Polymer Properties (Marcel Dekker, NY, 2002)の公開モデルと、自社開発の化合物または訓練セットのための拡張可能なモデルが含まれています。Polymer Properties Collectionは、対象の重合体を示す単量体の繰り返し単位を指定し、予測する物性を選択するだけで使用できます。

## レポート作成と可視化

### Reporting Component Collection

Reporting Collectionは、データ解析やデータマイニングの結果を表示するレポートやWebアプリケーションを作成してカスタマイズするための包括的なコンポーネントセットです。コンテンツやレイアウトを自由に制御できるため、同僚と知識を共有するための効率のよりコミュニケーションツールを作成できます。Webページ、Microsoft Officeアプリケーション、PDFなどさまざまな形式で作成でき、テキスト、表、グラフを使用できます。Webレポートに対話機能を追加すると、データの検討に役立つ強力な環境を構築できます。作成したWebアプリケーションは他の人と共有できるため、自ら所有するデータのほか、同僚のデータも参照することができます。Webアプリケーションの展開は、Web Port、SharePoint、RSS経由、または一般のWebブラウザに対して行えます。

## データベースとアプリケーション統合

### Integration Component Collection

Integration Collectionで、Java、Perl、VBScript、SOAPIに基づいた社内やサードパーティの既存プログラムをPipeline Pilotと統合します。ODBC、JDBCを利用してOracleからの分析・レポート生成用にデータを抽出し、また結果を直接データベースに格納できます。

## 画像処理

### Imaging Component Collection

Imaging Collectionでは、画像データの高精度化、処理、解析、統合、カタログ化、検索、レポート作成に関する高度な機能を提供しており、Pipeline Pilotのワークフロー自動化機能と視覚的なプログラミング機能を企業の画像情報処理に活用することができます。画像データと数値データ、化学データ、グラフィックデータ、テキストデータをまとめて、統一的な計算フレームワークに統合します。テンプレートプロトコルにより、カスタムのドリルダウン・レポート作成機能、画像のリンク機能を迅速に開発できます。高度な学習コンポーネントは教師有り型と教師無し型があり、クラス最高レベルの処理機能を備えています。Pipeline Pilotのオープンプラットフォームでは、標準的な画像ファイル形式のほぼすべてに対応しています。

## 解析と統計

### Data Modeling Component Collection

ラーニング、データモデリングツール、統計フィルター、大規模なデータに最適化されたクラスタリングコンポーネントを使用して、効果的なデータモデリングを実行します。Pipeline Pilotの標準機能を拡張し、統計的な予測のモデリングを行って、データマイニングを適用します。高速なデータクラスタリング、ベイジアン学習、主成分解析、線形回帰、部分最小二乗回帰など効果的な手法を使用します。構造活性相関モデル、化合物のクラスタリング、および共通する最大の基礎構造探索を実行します。

### Advanced Modeling Component Collection

Recursive Partitioning (RP法) および Multiobjective Pareto Optimization法の高度なデータモデリング法を利用します。RP法はSingle TreeおよびForest of Treeの学習モデルのコレクションが利用できます。この方法では1つまたは複数の応答変数 (Response Variables) を学習できます。Pareto Optimization componentは、多目的最適化問題に対する手法を含んでおり、部分的に矛盾する複数の目的間で、トレードオフする基準を解決する方法を提供します。

## R Statistics Component Collection

R Statistics Collectionは、知見に富んだ解析、情報に富むグラフ表示、そしてデータの裏付けされた意思決定を行います。データ操作、クラスタリング、モデル学習、古典的、予備的なデータ解析に対する統計手法を実行するコンポーネントがあります。統計解析エンジンは、広く使用されているR統計パッケージです。R 統計解析とデータ操作法を Pipeline Pilot のデータフローに適用します。また、R からの出力結果を Pipeline Pilot に組み込み、さらに解析できます。学習法のサンプリングには、部分最小二乗、ニューラルネット、サポートベクタマシンなどがあります。独自のRスクリプトをR Custom Scriptコンポーネントに追加すると、コレクションを簡単に拡張できます。

## ドキュメント検索と分析

### Text Analytics Component Collection

テキストマイニングを利用して、実験結果の補完に不可欠な知見を得ることができます。文献検索、特性解析、分析によって新しい情報で処理ルーチンを強化できます。探索、抽出、特性解析の手順を自動ルーチン化することで、最も困難なテキストマイニングの目的を実現します。テキストマイニングを既存のプロトコルに統合し、そのプロトコルを自動的に実行します。

### ChemMining Component Collection (Chemical Text Mining)

知的所有権(IP)の状況を調査し、公開資料と内部資料を化学的な知識を持ってテキストマイニングすることで、重複した研究を回避します。ChemMining Collectionでは、テキスト文書の中に化学名を見つけると、操作できるように実際の化学構造に変換します。同時に複数の文書を検索し、検出された化学構造を示すレポートを作成できます。また、大量の文書を処理して、検出した化学構造を検索可能なデータベースを作成できます。疾患プロセスや生体分子に関連する用語など、重要な専門用語に対して化学構造検索とテキスト検索を組み合わせ、完全に文脈に依存する化学的な知識に基づいた検索を行えます。

## ラボラトリ向け

### Plate Data Analytics Component Collection

プレートデータの読み書き、レポート作成、表示、編集、計算ができます。また、プレートやウェルの情報をデータパイプライン上で扱うことができ、様々な操作が可能です。さらにPipeline

PilotのGUIを使うことで、プログラムを書くことなく、スクリーニング結果の分析に必要な複雑なデータ解析手法が使えます。Integration Collectionを使うことで、プレートデータをデータベースに登録したり、呼び出したりできます。

### Analytical Instrumentation Component Collection

Analytical Instrumentation Collectionを使用すると、分析実験室で生成したデータにPipeline Pilotからアクセスし、データの処理や共有を行えるようになります。また、分析データの処理を合理化し、レポートを生成して、結果を他の部門と簡単に共有できます。このコレクションでは、NMRの機能として、ピーク同定、ピーク化の面積特定、線幅分析、バックグラウンド検出と除去、スペクトルの補完、トランケーション、スケールリング、スムージング、スペクトルの減算、汎用的フーリエ変換といった、一般的なデータ処理をサポートしています。

Pipeline Pilotの詳細については、次のURLを参照して下さい。

<http://accelrys.co.jp/products/pipeline-pilot>