

## タイトル:ファーマ 4.0 に備えよ

クラウドにおけるデータの完全性とビッグデータの応用

著者: Toni Manzano, PhD, and Gilad Langer, PhD

(Pharmaceutical Engineering, 2018, Vol 38, No5, 72-79)

翻訳:京都大学大学院医学研究科薬剤疫学分野 大学院生 服部 雅優 (Masakatsu HATTORI)

典型的な医薬品製造業務で収集されるデータの量は驚異的だが、この情報の多くはコンプライアンス以上にはほとんど使用されていないことを調査は示している。ビッグデータ、人工知能、機械学習、ディープラーニングなどの新技術は、リアルタイムデータの前例のない分析を可能にし、工程や運用の傾向を予測することさえできる。製造者は、工程を理解し、改善するために、これらの技術とその技術が与えてくれる情報を利用することができる。

機械学習 (Machine learning; ML) は、”粗暴な力” - ソリューションを探す際に可能な全ての選択肢を評価する、超高速で情け容赦のない計算を通じて AI を実現する方法と表現される。ディープラーニング (Deep learning; DL) は、一連の処理を経てデータをカスケードおよび変換するために、人間の学習の構造と機能に基づいたアルゴリズムを使用する。ビッグデータの技術は、大量のデータを高速に格納および取得できる。そして人工知能 (AI) は、計画、言語理解、問題解決などの人間の活動を学ぶことができ、これらのデータセットから、有意義な結論および予測を生み出すための高度な分析を提供する。あわせて、これらの技術は、ファーマ 4.0<sup>1</sup> の約束事の一つである継続的な改善に、必要な情報と知能を提供する。しかし、1つの疑問が生じる:いかに、特に GMP の運用を維持するための、データの完全性を保つことができるであろうか?

ビッグデータ環境とそのアルゴリズムは、データ完全性のガイドラインに沿うように設計されなければならない。これには、システム構成、データの取り込みと保管、およびデータ消費を含んだシステム設計にベストプラクティスを適用しようとする、明確かつ非常に協調的な努力が必要となる。しかし、これらの技術が、組み込みの、自動化された監査証跡を可能にすることは、非常に興味深い。監査証跡によって、AI が取り込まれたデータを分析し、不履行またはデータの完全性の問題に関しての警告を引き起こすことができる。

現在の医薬品製造は、分析とリアルタイムの製造インテリジェンスに必要なデータ取り込み機能だけでなく、GxP タスクを管理するためのさまざまなシステムを備えている。しかし、調査では、収集された全ての製造データのうちの 70%は使用されておらず、医薬品製造データ収集業務は、一般的にかなりの無駄があることが示された。最適化への圧力に

直面して、多くの製造者は、工程や作業についての洞察を得るために、すでに取り込まれたデータを使用することの利点を理解している。しかし、このデータを使用する能力は、それぞれに独自の形式を持った複数の融通の利かないシステムがあることを考えると、些細なものではない。インダストリー4.0のパラダイムは、ビッグデータや産業におけるモノのインターネット（IIoT）などの技術を用いて、これらの欠点の多くを解決することを見込んでいる。

## インダストリー4.0

第四次産業革命と称されるインダストリー4.0は、2011年にハノーファーフェアで初めて発表された。この新しいパラダイムは、ドイツ政府の技術戦略として始まったもので、科学とリスクに基づいたアプローチを、製造と工程のインテリジェンスに適用するものである。ISPEの特別分科会（Special Interest Group）では、それを医薬品製造のための”ファーマ4.0”として再定義した。図1は、4つの主な原則<sup>1</sup>を示しており、データと製造インテリジェンスの重要性を説明している。

現在の医薬品製造をファーマ4.0へ変換するには、製造と工程のデータ収集への新しいアプローチを必要とする。ビッグデータは、複雑な統合モデルを必要としない実用的なソリューションを提供し、人工知能向けのクラウドベースの高度な分析技術の使用を可能にする。十分な情報と質の高い内容があれば、これらの技術は、データを工程の最適化、継続的な改善、作業の卓越性、さらにはリアルタイムの公表などを含んだ、重要な活動を支持する知識へ変換することができる。ファーマ4.0はまた、AIアルゴリズムの結果に基づいたGxP決定の発展を可能にする。このことは、特に、米国連邦規則第21条第11章<sup>4</sup>、210章<sup>5</sup>および211章<sup>6</sup>やEudraLex Annex 11.7のようなその他の関連規制同様、米国食品医薬品局（FDA）<sup>2</sup>、欧州医薬品庁<sup>3</sup>の、データ完全性についてのガイドラインへのコンプライアンスを、非常に重要なものとしている。

他業界でのインダストリー4.0への発展は、完全に結びついた環境の価値を証明してきた。他業界での発展は、ビッグデータの分析がいかにより作業を最適化し、品質を向上させることができるかを示してきた。製薬業界やバイオテクノロジー業界は、データの完全性を確立し、準拠した方法で管理できさえすれば、同じ益を受けられるはずである。

## なぜビッグデータなのか？

インダストリー4.0では、製造環境が完全に結びついており、全ての作業と機器がリアルタイムでデータを伝達している。製造環境には、デバイスから作業システムに至るまで、製造作業全体（作業現場、工程区域、生産ユニット、生産ラインなど）にわたって、工場全体が含まれる。その結果、時系列の工程データからバッチレコードなどの複雑なデータセットに至るまで、収集されるデータの量は膨大かつ多様である。

中規模の施設でさえ、年間500テラバイトから10ペタバイトのデータを収集できる。（比

較の 1 例として、1 ペタバイトは米国議会図書館の所有量の 4 倍である)。これらの数字は膨大であるように見えるが、現代的な施設では標準的であり、新しい設備や、システム、機器が導入されるにしたがって増加する。そのため、ストレージと索引付けにはビッグデータのソリューションが必要であり、ビッグデータのソリューションによって、情報を履歴分析やリアルタイム分析に扱うことができるようになる。

この量および種類のデータを管理することは簡単な作業ではないが、従来の製造インテリジェンスシステムは、効果的な方法で行うことができない。この量の情報をデータ全体の GMP ワークフロー（取得、利用、バックアップ、廃棄）に沿って維持するために必要なコストと労力は、すぐに管理できなくなる可能性がある。これらのますます大量のデータは、データセンター、バックアップインフラストラクチャ、および IT サービスに多大な投資を必要とする。データ量が増加するにつれて、定期的なアップグレードで追加のコストが発生する。

しかし、これらのオンプレミスのアーキテクチャに対しての代替物がある：クラウドベースの大規模データサービスである。

### なぜクラウドなのか？

クラウドベースの選択肢は、必要とされるサービスによって異なる。

- サービスとしてのインフラストラクチャ (IaaS) は、物理的なハードウェアとそれを維持するために必要なロジックを外部調達するものである。サーバー、ハードディスク、およびそれらを管理するためのツール（地域立地、データ分割、尺度構成、セキュリティアクセス、記憶媒体のライフサイクル、およびバックアップ）などのコンピューティング資源は、サービスとして提供される。
- サービスとしてのプラットフォーム (PaaS) は、ソフトウェアを購入することなくソフトウェアの開発と展開が可能にするものである。Google、Microsoft、Amazon などの企業は、これらのタイプのすぐに使えるコンポーネントを提供している。
- サービスとしてのソフトウェア (SaaS) は、クラウドベースのプラットフォーム上でソフトウェアアプリケーションを実行する；ソフトウェアがローカルにインストールされることはない。このことは、従来のオンサイトアプリケーションとは非常に異なるものである。

これらの”XaaS”と呼ばれる代替製品は、大容量のデータストレージとコンピューティング資源の業界要件を手ごろな価格で満たすことができる新しい方法を提供する。それらの代替製品は、グローバルアクセス、組み込みセキュリティ、および固有の監査証跡でビッグデータを使用する新しい方法をもたらすものである。それらは、AI アルゴリズムで膨大な量のデータを処理するために必要な計算能力とデータストレージを提供する。なぜなら、大量の情報を利用するために開発されたものであり、大量の情報利用は、ファーマ 4.0 の重

要な要件となっているためである。

### ビッグデータと AI はどうなのか?

単にデータを取り込むこと自体は、私たちに知識をもたらすものではない。データは情報へと処理され、情報が今度は知識に変換されなければならない。変換にはコンテキストと分析が必要となる。

- ビッグデータはますます大きくなる。過去には、1 テラバイトがビッグデータとみなされていた;いまや、1 ペタバイトが標準である。
- 膨大な量のデータがあるため、スチューデントの T 分布や、カイ二乗、分散分析などの古典的な統計的方法は実用的ではない。
- データは構造化されておらず、従来の分析方法を不可能にさせる複数の形式 (数値型、カテゴリ型、ビジュアル型、ドキュメント型) がある。
- 従来のサーバーベースのシステムでは、テラバイトやペタバイトのデータを分析するために必要な処理能力を満たすことは、実質的に不可能である。

しかし、AI モデルは、K-Means、Random Forest、K-Nearest Neighbors などのアルゴリズムを使用して、膨大な量の非構造化データを分析できる。これらの方法は新しいものではないが (1960 年代に開発された)、現代のクラウドのコンピューティングと膨大な量のデータを組み合わせると非常に便利になる。これら 3 つの要素 (ビッグデータ、クラウドコンピューティング、AI) は、現代における製造の複雑な性質を理解するために科学を応用する新しい方法を提示している。これら 3 つの要素は、設備、装置、工程、システム、および作業者が連続したデータソースとなっている生産過程を監視する能力をもたらすものである。

### 「典型的な医薬品製造作業で収集されるデータ量は驚異的である」

これら 3 つの要素は、より実用的で、達成可能で、コスト効率の良いものにする一方で、継続的な工程検証も準備している。

現在のベストプラクティスの医薬品製造分析は、製品品質を監視するための特定の制御工程パラメータ (CPPs) および制御品質特性 (CQAs) に関する統計的手法を使用している。しかし、実際には、これらのパラメータは、自身の挙動同様、外部要因および内部要因にも依存している。製品品質を真に理解するためには、CPPs や CQAs だけでなく、工程における変数の本質的な相互依存性に加えて、製造環境の全ての側面を考慮した全体的な視点を取り入れる必要がある。

AI を用いたビッグデータは、この全体的な視点を支援することができる。伝統的な統計的手法が提供できるパターンや依存関係を越えた、パターンや依存関係を見つける能力は、多くの製造業で採用されているほど強力な手法を備えている。製薬業界はすでに AI を研究や、開発、製造の工程でデータを管理するための有効なツールと考えている。人工ニュー

ラルネットワークとサポートベクターマシンの両方が、高度な分析手法のための有効な計量化学の技術として、欧州薬局方によって近年提案されている。<sup>3</sup>

これまで、製造インテリジェンスソリューションのための、共通のコンテキスト統合された情報モデルをもたすために、多大なる労力が費やされてきた;これは例によって国際計測制御学会の階層モデル ISA-95 と ISA-88 に基づいたものである。しかし、前述のように、ビッグデータ環境に格納されたデータは構造化されていない。したがって、自動化および製造のシステムにどんなコンテキストがあっても、失われてしまっている。

しかし、非構造化データはまた、統合モデルを開発する必要がないということを意味する。実際、AI アルゴリズムは、非構造化データの方でより良く働くだけでなく、さまざまな形式の複数のソースからデータを処理できる。データを解釈して AI モデルを作成するためのコンテキストが依然として重要であるため、将来の製造インテリジェンスソリューションは、自動化および製造システムからコンテキスト階層を取り込み、それをビッグデータにオーバーレイとして構築できる必要がある。これらのオーバーレイは、(ストレージ内の) データを再構成するものではなく、単にデータの境界を指し示し、境界を設定するだけである。オーバーレイは、マスタレシピモデル、設備モデル、製品モデルなどのデータセットに対して複数の視点を提供できる。たとえば、因果関係を明らかにする AI アルゴリズムの実行が、いかに、従来の階層モデルを強固にするパターンを検出し、時には実行しなければ見過ごされていたであろう関係性を指摘することさえあるかは、注目に値するものである。

## 品質とセキュリティ

SaaS とクラウドベースのソリューションは、情報システムを製造するための新しいパラダイムを導くものである。ストレージ、サーバー、仮想化、オペレーティングシステム、ミドルウェア、データ、アプリケーションなどのオンプレミスの IT インフラストラクチャを展開して維持する必要がないため、IT インフラストラクチャの展開と保守に伴うコストはほとんどない。表 A は、異なるシナリオでの各コンポーネントの役割を比較している。

従来の IT アプリケーション、システムアーキテクチャ、データ、およびソフトウェアは、オンプレミス、つまり企業の IT 物理的なインフラストラクチャに展開されている。XaaS 環境には、物理サーバーやストレージ装置はない;サーバーのシリアル番号を取得することはできず、バックアップ処理は必要ない。さらに重要なことに、取り込まれ、保存され、および消費されるデータの品質は、古典的な IT システム手法を使用して測定または管理することはできない。クラウドコンピューティング技術は、必要とされるいかなるストレージとコンピューティング能力も瞬時に提供できる。この概念は、“エラスティックコンピューティング”と呼ばれ、無制限のストレージと自動スケラビリティを提供している。

XaaS は、ペイ・パー・ユース、ビジネス・フォーカス、堅牢なシステム、よく保護されたインフラストラクチャ、自動バックアップ、トータルワークフロー暗号化などの競争力

のある要素も導入している。現在の推定では、2021年までに、現在クラウドソリューションの採用において先導しているグローバル企業の半数以上が、全てのシステムをクラウドベースのインフラストラクチャに移行させると予測している。<sup>8</sup>クラウドベースのソリューションは、従来のハードウェアとソフトウェアの設計よりはるかに優れた堅牢性およびセキュリティを証明してきた。これは、技術ツールの高度な機密性を必要とする分野-金融、自動車、ヘルスケアなどのさまざまな業界での、クラウドベースソリューションシステムの急速な採用で確認されるものである。

しかし、他の業界と比較すると、製薬業界やバイオフーマ業界におけるクラウドベースのソリューションの採用は、主にセキュリティと品質に関する懸念から遅れている。これらのソリューション (IaaS、PaaS、SaaS、またはその組み合わせ) の認定には、リスクベースのアプローチとサービスプロバイダーの品質システムと工程への信頼が必要である。これはしばしば伝統的な認定の考え方に異を唱えるものである。

これらのソリューションには物性がない;検証するための、システムハードウェアもなければ、サーバインスタンスもない、シリアル番号もなければ、オペレータシステムもない。データや情報の保存やバックアップ工程、およびその保存の管理やコンプライアンスなどの運用面については、PaaSプロバイダーにより提供されるべきである。データストレージおよびコンピューティングシステムは、“サーバレス”環境で要求に応じて提供される。災害復旧計画は、地理的な位置にリアルタイムでデータと情報を同時に複製するアーキテクチャ (“geolocation replication”) により、よりシンプルかつ高速になった。

### サービス内容合意書

この種の環境では、ハードウェアと運用機能が XaaS サプライヤーに委任されており、サービス内容合意書 (SLA) はエンドユーザー-この場合、製薬会社-からの制御機構となる。品質は、サプライヤーの認証、監査、および定期的なレビューによって保証される。ライフサイエンス事業や関連するコンプライアンス実践を確立した XaaS サプライヤーを使用することは、大きな強みをもたらすものである。これはコンプライアンス業務の大半を移すだけでなく、より潜在的な品質リスクが軽減されていること示すものである。

コンピュータシステムのリスク評価分析に関する ISPE GAMP®のガイドラインは、これらの話題のいくつかを扱っている。<sup>10</sup> 世界保健機関<sup>11</sup> もまた、クラウドベースのサービスプロバイダーの品質契約を扱っている。英国の MHRA<sup>12</sup> は、クラウドベースのサービスを管理にガイダンスの1セクション全体を割り当てた最初の機関であった。これらの組織は全て、特定のライフサイエンス認定を取得したサプライヤーの使用を推奨しており、SLAによってそれらを管理することを推奨している。当然ながら、製薬企業がコンプライアンスを確証するための確かなプログラムを持っていることも重要である。

サイバーセキュリティは、重大な懸案となる別の観点であり、しばしば企業がクラウドベースのシステムを信用するのを渋らせるものである。しかし、クラウド技術は専心的な

開発を経ており、現在はオンプレミスのインフラストラクチャよりも安全性が高いと考えられている。<sup>13</sup>また、彼らは最新のセキュリティ技術を採用してきており、継続的な改善のためにセキュリティの専門家を雇い入れている。これらの進歩の多くは、クラウドベースのシステムを採用した際に、金融やヘルスケア分野などのデータのセキュリティとプライバシーを非常に重要にとらえる業界によって推進されてきた。

### クラウド XaaS モデルにおけるデータの完全性

従来の認定工程がクラウドベースの XaaS モデルで再検討される必要があるように、データの完全性に関するアプローチも再検討する必要がある。これは、ALCOA と GAMP の原則に基づいた全体的な品質アプローチを必要とする。

ALCOA—データ資産の帰属性、可読性、同時性、独自性、そして正確性に関する略—は、規制産業においてデータ完全性を保証するための枠組みとして使用されている。ALCOA は、文書化の規範(Good documentation practice; GDP) にきわめて重要であり、いかなるクラウドベースのソリューションでも考慮されるべきである。

**帰属性:** 保護されたプロトコルと証明書キーは、クラウド技術の柱である。暗号化アルゴリズム（キー、プライベートキー、公開キー、または最近のブロックチェーン技術の組み合わせ）は、メッセージソースの信頼性を保証するように設計されている。Hypertext Transfer Protocol Secure (HTTPS) や Message Queuing Telemetry Transport (MQTT) などのプロトコルは、セキュリティ証明書やピア・トゥ・ピアの構造を使用して帰属タスクをさらに容易にしている。

**可読性:** ビッグデータは、データをカプセル化する構造上に組み上げられている。可読性とは、これらの構造が解釈しやすいことを意味する。一般的な実装は、キーと値の組み合わせに基づいてプレーンテキストでデータを整理する JavaScript Object Notation (JSON) である。これらの構造はデータ型の定義を含むため、XML でよくあるような追加のタグ付けは必要ない。JSON は、可読性をより容易くし、計算能力をより使用しないような、より無駄の少ない解釈機序をあたえるものである。

**同時性:** 典型的な製造工程データはリアルタイムで収集され、ミリ秒単位の頻度で測定されている。計測について、高い入手可能性、精度、同期、および時間局在性を保証するために、多くの計測器が使用されている。ネットワークタイムプロトコル、ローカル環境設定、および高速インターネットインフラストラクチャが、測定値が適切な”包み”の中にあり、正しく印がつけられていることを保証する。生データは瞬間的に記録され、収集された時間を刻印されて、クラウドに送信される。接続が失われた場合、MQTT などの通信プロトコルが、元のタイムスタンプを維持しながら保管を遅らせることができる。

**根源性:** いくつかの電子ツールは、暗号化された証明書を使用して、データの根源を証明することができる。元の記録は、コンテキストを提供する特定のメタデータ（例えば、測定単位、地域設定、ジオロケーション）で包まれることができる。ピア・トゥ・ピア通信やソケット通信などのデータ転送モデルでは、データが根源であることを保証できる。

**正確性:** 論理ルールは、取り込み時点とデータレイクに保管する前のデータ品質を調べることができる。課題は、全てのビッグデータ環境で利用できるわけではないデータの変更を制御することである。変更管理は非常に重要であり、ビッグデータによるソリューションに組み込まれた機能は、完全な監査証跡、電子署名、およびセキュリティを含むべきである。

ビッグデータ環境で ALCOA 原則を適用し、リスクプロファイルを定義する場合、この環境で高度な分析の核となるデータフローと対応する決定フローを理解することが重要である（図 2）。

規制された環境でのビッグデータソリューションを支援するためには、ALCOA の原則は、ALCOA +として知られている組み合わせ、つまり、*完全な、一貫した、耐久性のある、そして利用可能な追加原則*で強化する必要がある。

**完全性と一貫性**は、システム設計の一環としてデータ定義スキーマと検証プロトコルで実装すべきである。

**耐久性:** 自動バックアップ機能と地理的なデータ複製機能を備えたクラウドプラットフォームは、“永遠に極めて耐久性があること”を実質的に保証するものである。データの永続性はデータ管理システムの一部として保証される;各バイトは、3 つの異なる場所および 3 つの異なるデータ記憶システム上で 3 つに複製される。

**利用可能性:** 既定の保存期間内は、バックアップまたはデータ管理を必要とせずに、データを常に利用することができる。

従来のチェックリストと評価は、クラウドベースのサービス環境では実用的ではない。表 B に示されるように、データ技術は、計画的に ALCOA +の原則（暗号化、セキュリティ、難読化、監視されたストレージ、アクセスルール、権利など）に準拠した固有のメカニズムを提供するためである。これらのシステムは、データの完全性がアーキテクチャの固有の機能である、GMP 環境に適合できる。この自己監視システム上のあらゆる挙動と変更は、GMP データレイクに保存される。そして、躍進的な見込みがある:データに AI アルゴリズム

ムを適用することにより、システムはコンプライアンスレベルを検出し、あらゆる不適合について警告を発することができる。これは、継続的な工程検証と同じ原理であり、データ自体に単純に適用されるものである。

### クラウドとビッグデータの実装

ビッグデータやクラウドベースのシステムへの切り替えを検討している製薬企業やバイオファーマ企業は、これらの新技術の経験やベストプラクティスが不足していることにより努力を強いられている。また、企業は、AIなどの、より破壊的な知識管理機能の専門的知識が不足しているかもしれない。結果として、GMPとGDPの適格性確保と維持のための新しい戦略が採用される必要がある。

- システムの品質とコンプライアンスの信用の確立を小さく始めなさい。これは新しい技術であり、おそらく製造組織のほとんどには知られていない。それは、採用と学習のカーブを急にするものである。

### 「現在の医薬品製造をファーマ 4.0 に変形させるには、製造および工程データの取り込みの新たなアプローチが必要である」

重要なステークホルダーは、新しいソリューションがもたらす価値に賛成することが肝要である。グッドプラクティスは、GMPの影響が最小限である摩擦の少ない展開から始めることである。予防的保全、環境モニタリング、または省エネルギーのためのアプリケーションが適切な対象である。<sup>1</sup>その他の選択肢は、再現性とスケーラビリティを必要とするタスクである。たとえば、FDAは、レポート文書を99.7%の精度で、クラウドベースのソリューションでデジタルデータに変換し、1ページあたり29ドルから0.25ドルへコストを削減した。<sup>14</sup>

- SaaS サプライヤーと親密かつ効果的な関係を構築しなさい。これは多くの組織にとって新しい技術であるため、全ての関係者が共通の目標を念頭に置いて働くことが肝要である。サプライヤーは、提示されたソリューションに関しての成熟度、強み、弱点について十分に率直でなければならず、医薬品ビジネスを理解し、インフラストラクチャ、プラットフォーム、サービスを適合するために必要なタスクを容易にすることを、工程を検証すること同様に行わなければならない。SaaS サプライヤーとの良好な関係を支えるSLAが重要である。コンプライアンスは依然必須であるが、実装責任はプロバイダーに移っていくものだ。
- 特定のデータ計算を必要とする人工知能アルゴリズム、索引付け、または集約工程に関連するサービスの場合、プロバイダーはデータ運用を認定するために使用された検証工程を共有する準備ができていなければならない。システムの認定とALCOA+原則の使

用のための新しい革新的な方法を検討することは価値があるかもしれない。

- リスクベースのアプローチは、ビッグデータ、つまりクラウドベースのソリューションを設計する際の貴重なツールである。信頼関係を構築するには、サプライヤーと組織の良好な仕事上の関係が必要である;これは QA 組織にとって特に当てはまるものである。
- AI、ML、および DL ツールは、非常に大きなデータセットの分析の標準とみなされなければならない、そのうちのいくつかはすでに欧州薬局方によって認識されている。これらの分析手法の結果は、従来の統計に適用されたのと同じルールで検証できるモデルである。
- クラウドとビッグデータインフラストラクチャに組み込まれた自動化ツールとデータ完全性ツールを必ず活用しなさい。

### 産業での適用

クラウドベースの大規模データ分析技術は、バイオテクノロジー、医薬品、医療機器業界で急速に受け入れられている。インダストリー4.0の技術は、ISPEに導かれたファーマ4.0を生じさせた。この新しいパラダイムで見込まれるものは、実践的な科学を基にしたアプローチを用いることによる一貫した品質である。2020年以降の予測では、ライフサイエンスと製造活動に関する以下の筋書きが記述されている:<sup>8</sup>

2020年までに:

- ライフサイエンスのトップ50の研究開発組織における現在のAIの概念実証(AIによる概念の実証)のうち、20%が生産段階に移るであろう。
- 主要技術プロバイダーは、独立したデジタルヘルスサービスを患者に直接提供するであろう。
- ビジネス機能主導のMLの構想は、企業IT主導のプロジェクトの少なくとも2倍の成功率となるであろう。

2021年までに:

- トップライフサイエンスの研究開発IT組織の4分の3は、従来とは異なるエコシステムとのデータ接続および業務提携を作り出すであろう。

2022年までに:

- 前競争的な、共有され、開かれた研究のITプラットフォームは、革新的環境に置かれたライフサイエンスの少なくとも25%の組織が、絶え間なく成長する研究開発エコシステムと結びつくのを支援するであろう。

**「XaaSという代替製品は、大容量のデータストレージおよびコンピューティン**

## 「データ資源の業界要件を、手ごろな価格で満たす新しい方法を提供してくれる」

適用は急な速度で進んでいる。アプリケーションではなく、データに基づいた業界に新しい活力がある。これらの新しい技術は、適合が容易であるばかりでなく、業界の現在のデータ完全性の課題の多くを解決できる。図 3 は、米国内外の FDA の警告書に、データ管理とデータ完全性についての言及が、ますます含まれていることを示している。

### 結論

いつもながら、進歩は止まらない。インダストリー4.0の技術は、製造における世代交代を容易にするであろう。ビッグデータ、AI、クラウドコンピューティング、継続的な接続性、高度な分析、そして IIoT によるデジタル化がここにあり、デジタル製造工場が現実のものとなりつつある。より控えめなバイオテクノロジーや製薬業界でも、これら画期的な技術の採用が慎重に試されている。このようにして、ファーマ 4.0 の概念は、特にこの新しい現実の中での GMP の実用適応化のために勢いを増すであろう。クラウド、ビッグデータ、AI はデータを資産にする;これは工程最適化、科学ベースのアプローチ、継続的な工程検証など、大きな可能性をもたらす。反応的から予測的へ、厳格な工程制御から適応的な工程制御に移行する能力は、高度な情報と製造技術が物理的工程に柔軟性をもたらす”スマートな”製造を生み出すものである。<sup>16</sup>

現在のデータの完全性のベストプラクティスとガイドラインは、固有のデータ管理技術が自動的にコンプライアンスを監視する XaaS のビッグデータ環境に適用されるものである。製造データを分析する AI 技術は、疑わしいデータを見つけて分離する高度な分析を利用した、絶え間ない監視とリスク分析も可能にしている。継続的なデータ取り込みは、全体的な構造やコンテキストを持たずに、製造または生産のソースからデータを収集する。ニューラルネットワークとパターン認識アルゴリズムは、正誤のシナリオについてシステムを訓練するために使用される。AI アルゴリズムは、不適合な動作、予期しないエラー、矛盾を検出することでコンテキストやパターンを発見する。これらのソリューションの例は、予防的保守分野ですでに発展したものである。<sup>17</sup>データの完全性の問題を検出して防止するために、同じ予測手法が現在適用されている。

ファーマ 4.0 のパラダイムは、製造インテリジェンスに新しい意味を与えた。我々は、科学に基づいた製造工程を管理する予測的かつ適応的なシステムを真に実現することができる。未来はここにあるのだ;採用曲線の急な部分を登ることが、私たちの照準になるはずである。

本文以上

<図表の説明>

表 A 古典的 IT および XaaS 概要に含まれる要素

表 B データの完全性を容易にするものとして機能する ALCOA+技術の概念

図 1 ファーマ 4.0

図 2 ビッグデータ環境におけるデータと意思決定の流れ

図 3 2013 年～2017 年に米国内外で出されたデータの完全性の問題に関する警告書