

# SHIONOGI DATA SCIENCE FES 2023

## データサイエンスとのコラボが生み出す新しい価値

### ーやりたいが繋がる場所ー

本記事では、前半に引き続き後半 3 つのトピック（データサイエンス×解析環境/DB 活用、データサイエンス×スポーツ、データサイエンス×感染症）の内容を振り返ります。

前半 3 つのトピック（データサイエンス×人工知能、データサイエンス×デバイス、データサイエンス×教育）の講演の記事も併せてご覧ください。

#### ▼目次

##### Topic 4 : データサイエンス×解析環境/DB 活用

- ・ 石尾 千晶氏/アマゾンウェブサービスジャパン合同会社 技術統括本部 ソリューションアーキテクト
- ・ 雑賀 恵美/塩野義製薬株式会社 データサイエンス部
- ・ 福永 真一/塩野義製薬株式会社 データサイエンス部

##### Topic 5 : データサイエンス×スポーツ

- ・ 那須 大毅氏/NTT コミュニケーション科学基礎研究所・柏野多様脳特別研究室 主任研究員
- ・ 馬場 崇充/塩野義製薬株式会社 データサイエンス部 / 解析センター

##### Topic 6 : データサイエンス×感染症

- ・ 【基調講演 3】西浦 博氏/京都大学 大学院医学研究科 社会健康医学系専攻 教授
- ・ 藤田 智紀/塩野義製薬株式会社 データサイエンス部

#### Topic4:データサイエンス×解析環境/DB 活用

製薬企業に限らず企業活動の中では様々な場面でデータが生まれ、その活用が求められています。それら膨大なデータをどのように収集、管理そして活用するのか、データ活用の戦略や展望について紹介されました。

## 「AWS で実現する協創のためのデータ解析基盤」

石尾 千晶氏 (アマゾンウェブサービスジャパン合同会社 技術統括本部 ソリューションアーキテクト)

石尾氏： 企業間の“協創”には、各企業がこれまでに蓄積してきたノウハウを持ち寄りつつも、機密データなどのプライバシーを守り安心して使える、柔軟で拡張性のある実験環境を用意することが必要です。本講演では、AWS を活用した“協創”の事例をご紹介します。まず創薬領域の事例として、MELLODDY という欧州の製薬企業などによる機械学習を用いた創薬の共同研究プロジェクトをご紹介します。ここでは、各社保有の化合物ライブラリは秘匿したまま、AWS 上で新たな共通モデルを作成し、新化合物に対する活性予測精度の 10%向上を達成されています。次に、新規事業開発での AWS の活用事例として、下水サーベイランスによる COVID-19 のモニタリング事業をしている AdvanSentinel (塩野義製薬と島津製作所による合併会社) の事例を紹介します。下水サーベイランスでは、各下水流域からサンプル採集、データ解析・解析結果の活用が行われますが、外部委託企業など複数のステークホルダーが関与する上でサンプル ID の管理、データの授受の権限管理などが煩雑、また国際基準データ形式へ追従のため大幅な運用変更が必要という課題がありました。関連する企業がそれぞれ AWS のアカウントを保有し、AWS 上で各企業の権限をコントロールしながらデータやログを管理し、システムを再構築することで、課題解決に繋がりました。

このように様々な AWS 上のサービスはユーザーのニーズに合わせた自由な組み合わせが可能であり、様々な企業や業界の“協創”を促進するためにもニーズに合わせたサポートやご提案をさせていただきたいと考えています。

## 「SHIONOGI における次世代データマネジメント」

塩野義製薬 データサイエンス部 雑賀 恵美

雑賀： 塩野義製薬をはじめとする創薬開発を行う製薬企業では、薬の研究開発をして終わりではなく、患者さんの手に届くまでに様々な領域の業務を経て、各領域で多様なデータが発生しています。データサイエンス部ではそのデータをひとところに集約・管理し、データの資産価値を高めるためにセントラルデータマネジメント (CDM) を行っています。CDM には大きく 2 つに分けて、業務システム間のデータ連携におけるデータハブの役割と、データを活用しやすい形で準備しておくデータウェアハウスの役割があります。私たちが考えるデータマネジメントのあり方として重要なポイントは、ニーズを先取りし能動的であること、そして中央集権型を脱却し分散型データコミュニティを形成することです。SHIONOGI では、データを収集したその先の活用シーンと要件を提案することや、データから導き出した仮説をビジネスサイドに提案することもデータマネジメントの役割と考えています。また、分散型データコミュニティとは、分析者自身がデータ管理者となって、データを持ち寄る仕組みです。データ活用が進めば進むほど各分析者の目的が多様化するなかで、データウェアハウスに各分析者の視点を素早く、的確に取り入れることが可能になります。SHIONOGI のデータマネージャは、この分散型データコミュニティを運営管理することで企画提案につながるアイデアを得られるとも考えています。

## 「ビジネス課題解決のためのデータ活用 ～データ蓄積から分析・可視化まで～」

塩野義製薬 データサイエンス部 福永 真一

福永： ビジネス領域において人の経験、勘に基づいた意思決定ではなく、データに基づいた意思決定が求められています。データドリブン型ビジネスを行うための全体像として、目的を明確にし、仮説検証サイクルを高速で回していく仕組み作りが重要だと考えています。ビジネス部門におけるデータ活用を進める上で、データの利用のスピード面、情報クオリティ面での課題があるのではないのでしょうか。

塩野義製薬では、このような課題を解決するには、どのような意思決定がビジネス上重要なかを特定し、どのようなデータが必要かという検討が重要だと捉え、その上でその意思決定に必要なデータを収集、蓄積を行っています。その上でデータの可視化やビジネス領域ごとに意思決定に必要なデータを分類・整理し、活用しやすいデータテーブル構造を設計して蓄積を行っています。これらを提供するだけでなくその効果検証も実施しており、アンケートによると週あたり約 120 時間の業務効率化の効果が得られたという結果が出ています。さらに一度提供して終わりではなく、可視化データの質とユーザビリティを継続的に改善させるサイクルを確立し、ユーザーの満足度向上を図っています。

現在は医薬品情報提供領域におけるデータ活用支援を行っておりますが、今後、人事や資産経理データ等について意思決定を検討するためのデータ活用支援にも拡大したいと考えています。

## Topic5:データサイエンス×スポーツ

製薬企業のデータサイエンス系イベントのテーマとしては珍しく、イメージが湧かない方もいらっしゃるかもしれません。本セッションでは、スポーツにデータサイエンスを適用した事例を通じて、意外なヘルスケアとの繋がりが紹介されました。

### 「スポーツ脳科学研究の取り組みについてのご紹介」

那須 大毅氏（NTT コミュニケーション科学基礎研究所・柏野多様脳特別研究室）

那須氏： スポーツ脳科学研究の取り組みについて、日本ソフトボール協会との共同実験の事例を紹介します。

私たちはソフトボール競技の中でもメインとなるピッチャーとバッターの対戦に着目し、対戦中の選手の全身の運動計測を行いました。ストレートとチェンジアップでは球速が異なるのでタイミングをずらして打撃する必要がありますが、日本代表選手はタイミングをずらすことができている一方で、若手選手はずらすことができていることが分かりました。日本代表選手はこの球種の判断に必要な情報を投手の投球フォームから読み取っているのではないかと仮説を立て、VR 上でバッティング環境を再現し、実際の打撃と同じように選手にバットを振ってもらい、投球フォームと球種を入れ替えることでバットを振るタイミングが狂わされ

るかを検証しました。その結果、タイミングは狂わされ、打者はその入れ替えに気づいておらず、打者は無自覚的、潜在的に投球フォームからの情報を使っていることが明らかになってきました。私たちはスポーツの現場にある素朴な疑問や課題に対して科学的にアプローチをして、そこから得た知見、アイデアを基に脳のメカニズムに立脚した、現場で使える技術や知識を提供することを目標としています。アスリートを対象とすることで人間の多様な能力と技を深く理解することが可能であると考えています。

## 「VR デバイスを用いた無意識下におけるバッティング技術の向上 ～ソフトボールとヘルスケアの意外な繋がり～」

塩野義製薬 データサイエンス部 馬場 崇充

馬場：私たちは社会、IT 技術、個々人の価値観などの変化と共に変わっていく、治療を含めたヘルスケアへのニーズに応えるため、異業種の方々と協業し HaaS (Healthcare as a service) の実現を目指しています。本取り組みは、本日のご発表者でもある那須さんら NTT の方々と協業して進めたものです。

ヘルスケアサービスを考える上で、まずはスポーツ選手というご自身のコンディションコントロールに関心の高い方たちのデータを取得し、パフォーマンスを向上するための打ち手を考えることを試みました。パフォーマンスの指標として、脳機能の一つである潜在的な認知判断能力に着目し、VR を用いた仮想空間内での視覚的トレーニングを行い、この機能の向上を図りました。研究のステップとして、まず VR デバイスによって潜在的な認知機能を正しく評価できるか、次に VR デバイスを用いたトレーニングを行った場合、潜在的な認知機能の強化が可能なのかを検証しました。本研究の結果、VR を用いたトレーニングで、時間や場所の制約を受けることなく、ソフトボール選手の潜在的な認知機能の評価と向上が可能であるという示唆が得られました。

本研究の対象とした潜在的な認知判断能力はスポーツ以外でも重要な能力であり、本研究の成果はデータサイエンス、スポーツと脳科学というかなり意外な組み合わせによって得られたビジネスの種だと思っており、今後の展開についても引き続き検討しています。

## Topic6:データサイエンス×感染症

コロナ禍と言われるようになり約 3 年が経ちましたが、まだ流行が完全に収まったといえる状況ではありません。COVID-19 を含む感染症に対し、どのような対策、制御ができるのかアカデミアと企業の観点から紹介されました。

【基調講演 3】

「感染症の数理モデルと観察データ」

西浦 博氏 (京都大学大学院 医学研究科社会健康医学系専攻 環境衛生学分野 教授) “

西浦氏：感染症の流行対策に数理モデルが使用されたのは COVID-19 が初めてだと思います。世界中でできるだけ死亡者数を減らすため hammer and dance という抑制政策が展開されてきました。流行が起こり医療が逼迫しそうになれば、ロックダウンなどで人の接触を低減して流行を遅らせピークを低く維持し、ワクチンや治療薬が開発されるまで時間を稼ぐというのが、医学的な対策がない段階での COVID-19 との付き合い方でした。この状況の中でデータを用いて政策にアドバイスをしなければならなかったメンバーの一人として今回の経験では反省点がたくさんあり、記録に残す必要があると思っています。例えば死亡者数のデータと向き合いきれていなかったことや、人流（人の移動）の影響が実際にはどうだったのか、一つ一つのデータをしっかりともう一回紐解いてモデル分析をしながら明らかにしていく必要があると考えています。

データは落ち着いて正しくハンドルすれば、うそをつくものではなく、そこがデータサイエンスの魅力だと思っています。感染症に関しては政策などの事後評価も大切です、将来新しいパンデミックが起こるかもしれないことも想定しながら研究を続けていきたいと思っています。本日聴講くださったみなさんの一人でもこの分野にご興味を持っていただけたら嬉しいです。

### 「データサイエンスを通じた感染症への関わり方」

塩野義製薬 データサイエンス部 藤田 智紀

藤田：パンデミック下において、感染症に関する様々な情報の定量化が求められています。これらに応えるため、データサイエンス部では、様々なデータを目的に合わせて活用して課題解決に取り組んでいます。ここでのアプローチはモデリング&シミュレーションの考え方をベースとしており、現実世界の現象を数式に落とし込み数学の言葉で理解することで、現象の解明につながるよう努めています。社会の状況の変化にもデータサイエンスの面から迅速に対応できるよう、即時性・追跡性の高い統一されたデータベースの構築や、より柔軟なモデリング・シミュレーション技術の開発が必要だと考えます。

その一つの検討例として、マルチエージェントシミュレーション（MAS）によるミクロレベルでのモデリング&シミュレーションを行いました。一人ひとりの自律的な行動をモデル化し、ワクチン接種シナリオの違いによって感染重症者数にどのような変化が見られるのかシミュレートし、経済への影響を最小限に抑える行動規制戦略の最適化について強化学習を組み合わせた検討も行うことで、仮想空間上を活用した、より効果的な意思決定の可能性が示唆されました。今後このようなモデル化に使用できる個人レベルのデータ収集が進めば、様々な AI・シミュレーション技術の活用により感染メカニズムの解明や感染拡大防止の意思決定への大きな助けとなることが期待されます。

### 【開催当日の様子】

本イベントの総合司会は、SHIONOGI 初のバーチャル社員であるシオノギカナデが担当し、ご好評いただきました。



本イベントは“協創”をキーワードに様々な業界からスペシャリストの方々にご登壇いただきました。

第 2 部の意見交換会では、登壇した塩野義製薬の従業員と参加者との質疑応答やディスカッション、情報交換の場としても大いに盛り上がりました。

当日は本当に多くの方にご参加いただき心より感謝申し上げます。

ご参加くださったみなさんにとって本イベントが有意義なものになっていましたら幸いです。