

# オープンイノベーションによる 次世代NDBデータ研究基盤の構築と発展

増加し続ける医療費と医療の質の向上といった課題を解決する鍵として、レセプト情報・特定健診等情報データベース(NDB<sup>※1</sup>)が注目を集めています。こうしたなか、日立は一般財団法人 医療経済研究・社会保険福祉協会 医療経済研究機構 (以下、医療経済研究機構)と東京大学生産技術研究所(以下、東大生研)とのオープンイノベーションにより次世代NDBデータ研究基盤を開発しました。膨大なデータの利活用による医療研究の進展に寄与しています。

※1 National Database of Health Insurance Claims and Specific Health Checkups of Japan

## ■ 超高速・超学際レセプト情報等ビッグデータ 解析プラットフォームシステム「SFINCS」

国民の共有財産であるNDBデータ<sup>※2</sup>は、2013年度から公益性のある研究を対象に第三者への提供が認められていますが、世界最大級のNDBデータを、医療研究などに役立てるため簡単・高速・精緻に解析可能とするのは容易ではありませんでした。

2017年、超高速・超学際レセプト情報等ビッグデータ解析プラットフォームシステム「SFINCS<sup>※3</sup>」は、医療経済研究機構が国立研究開発法人 日本医療研究開発機構から2016年度に受託した研究<sup>※4</sup>に基づき、三者のオープンイノベーションにより稼働しました。

本システムは現在6年分のNDBデータ(約2,000億レコード)を格納し、研究者に活用されています。さらに、本システムを発展するための三者の協業が継続しています。

※2 レセプト情報、特定健診・保健指導情報を収蔵。国民皆保険制度を導入している日本においては、医療機関で提供される医療の大部分が保険診療の対象となるため、レセプト情報の解析により、国民の医療の実態を明らかにすることが期待されている。

※3 Super-Fast super-Interdisciplinary japanese medical insurance Claim bigdata analytics platform System

※4 「エビデンスの飛躍的創出を可能とする超高速・超学際次世代NDBデータ研究基盤構築に関する研究」

## ■ 活用され続けるための日立の取り組み

このSFINCSを「将来にわたって活用され続けるシステム」にするためにはさまざまな課題がありました。日立が挑んだ課題解決のうち代表的な三つを紹介します。

### 1. 誰もが開発・運用できるよう

徹底的にオープンな技術にこだわり開発

「システムを容易に拡張できること」という課題については、NDBデータのような膨大なデータを超高速で解析処理する

ために、従来は特殊なハードウェアやプログラムを用意していました。しかしこれではシステムの増強や機能追加に高度な開発スキルを持つ人材の確保が必要になります。そこで日立は、汎用ハードウェアや一般的なプログラミング技術を使うなど、誰もが開発・運用できるよう徹底的にオープンな技術にこだわり、システム開発を行いました。例えば、SFINCS Appとデータベースとのやり取りには、取り扱える人材が豊富な標準規格(SQL<sup>※5</sup>)を採用しました。このような取り組みにより、安定的な人材確保を可能とし、将来の人材面のリスクも低減しました。

### 2. 高速かつ柔軟な解析基盤の実現

解析処理の速度は対象データの量に左右されますが、今後も増大し続けるNDBデータのようなビッグデータを高速に解析処理するため、一般的には必要なデータを抽出した目的別のデータマートを作成します。しかしそれにはプログラム開発やデータベース設計などの時間や労力、コストが必要です。そこで日立は、処理時間を大幅に短縮する超高速データベースエンジンHitachi Advanced Data Binder<sup>※6</sup>をSFINCSのデータ処理基盤に採用し、さらにデータベースのデータ構造を見直すことでデータマートに頼らず、高速な解析処理を実現しています。また、新たな解析ツールなどを拡充する際にもデータマートの構築などが不要であり、柔軟で機動的なサービスの提供が可能です。

### 3. データ解析の生産性を向上する共通機能

ビッグデータの解析では事前準備に多くの時間が割かれ、特にデータを解析しやすくするための加工処理(データ

クレンジング)には膨大な時間がかかっていました。そこで日立は、解析前のデータ加工処理をSFINCSの共通機能として実装することで利用者による処理を不要とし、利用者が解析作業に注力できるシステムを実現しました。

※5 Structured Query Language

※6 内閣府の最先端研究開発支援プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的社会的サービスの実証・評価」(中心研究者:喜連川 優 東京大学教授/国立情報学研究所所長)の成果を利用。

### SFINCS開発で活性化する NDBデータを活用した研究

SFINCSを核とした医療経済研究機構のプロジェクトには、2019年6月時点で15大学、6学会、60名を超える有識者が参加しています。NDBデータを活用した研究が活性化するなかで、これまでになかった成果が次々と生み出されていることは、SFINCS開発の大きな成果といえるでしょう。

一例を挙げれば、各種生活習慣病の実態に関するNDBデータの解析結果が、厚生労働省が個別に調査している結果と同様の傾向であることが確認されたことで、NDBデータの解析結果の有効性が検証されました。

また、他の例として、患者の社会的・経済的負担が大きく、医療費助成対象となる指定難病への認定が求められている1型糖尿病の患者数はこれまで把握されていませんでし

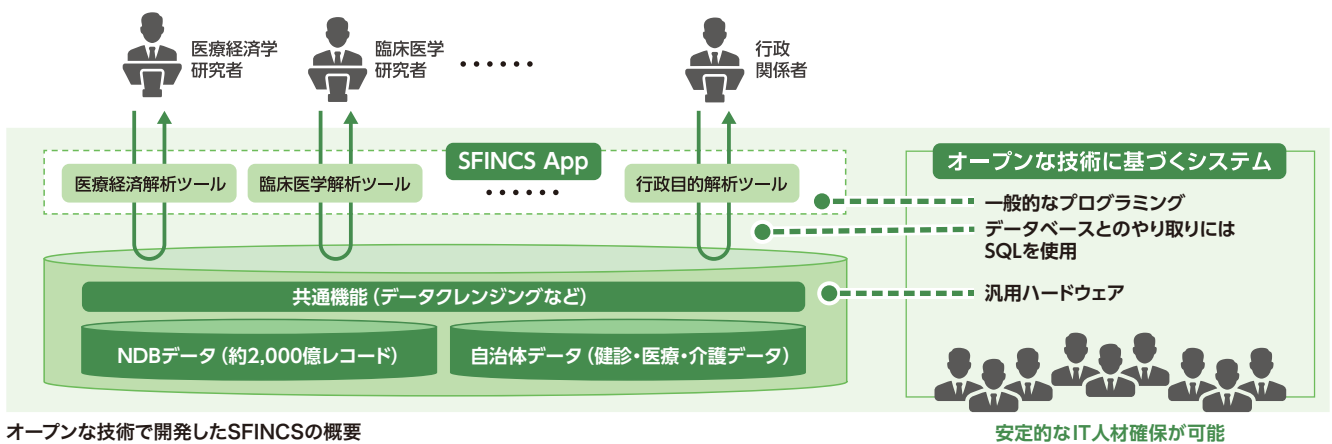
たが、NDBデータの解析調査によって、2014年時点の患者数が約11.7万人と推計されました。この結果は1型糖尿病が指定難病認定の基準となる「希少な疾患であること」を裏づける重要なデータとなっています。

### ビッグデータ時代が求める次世代NDBデータ基盤

現在、NDBデータをその他のビッグデータと連携させようという機運が高まっており、SFINCSはNDBデータと介護レセプトのデータ連携といった、次世代NDBデータの研究に先行して取り組んでいます。また併せて、クラウド利用を前提として設計する「クラウドネイティブ化」や、これに対応するプライバシー保護とデータの有用性を両立する高速な匿名加工処理の実装にも取り組んでいます。

このように、SFINCSはさまざまなビッグデータから新たなエビデンスの創出を支援する強力なプラットフォームとして今後も進化を続け、国民の健康と幸福に貢献していきます。

また、さらに本格化していくビッグデータ時代の次世代NDBデータ基盤を示すことができました。日立はヘルスケア分野にとどまらず、これからも幅広い分野でオープンイノベーションによる協創を加速させながら、直面するさまざまな社会課題の解決を追求し続けます。



オープンな技術で開発したSFINCSの概要

お問い合わせ先・情報提供サイト

(株)日立製作所 IoT・クラウドサービス事業部

<https://www.hitachi.co.jp/bigdata/platform/data-binder/>

