



# 医療情報の安心安全な研究利用を目指し 匿名化技術とHadoopで基盤を構築

## 安全性のカギを握る匿名加工エンジンにNSSOLの「匿丸」を採用

### 背景

次世代医療基盤法に基づき、医療情報を研究機関などに第三者提供するシステムの構築を検討していた。必要としていたのは、提供データの安全性を担保する匿名加工技術と数十億件規模の大規模データベース構築技術だ。



日本医師会  
ORCA管理機構株式会社  
代表取締役社長  
上野 智明氏



日本医師会  
ORCA管理機構株式会社  
事業推進部 部長  
プロジェクトマネージャー  
伊藤 伸昭氏



日本医師会  
ORCA管理機構株式会社  
事業推進部  
事業戦略アドバイザー  
工藤 憲一氏

日本医師会ORCA管理機構株式会社  
本社：東京都文京区本駒込6-1-21  
設立：2015年

### ソリューション

ITパートナーとして、これらの技術やノウハウを持つNSSOLを選定。Hadoopベースの大規模データベースを構築し、提供データの安全性と高い有用性を確保できる匿名加工ソリューションを開発した。

### 成果

国家プロジェクトの一環である実証事業が完了し、医療情報を第三者提供するための基盤は整った。今後、全国の医療機関や自治体から臨床、健診、介護などの情報を収集・蓄積し、2020年から第三者提供を開始する見込み。

### 医療情報を匿名化して第三者提供する基盤整備を検討

日本医師会ORCA管理機構は、日本医師会が医療のIT化と医療情報の標準化を目指して2001年に開始したORCAプロジェクトのもとで「日医標準レセプトソフト」などを開発・提供している。同ソフトを利用する医療機関は約1万8000に及ぶ。

医療情報の利活用を促進する次世代医療基盤法の成立を受け、日本医師会は2017年から同法に則った「生涯保健情報統合基盤」の構築を検討していた。国民の誕生から死亡までの医療情報（臨床・検査、健診、介護など）を全国の医療機関や自治体から収集して統合・蓄積し、匿名加工をして研究機関や民間企業（製薬・保険など）に第三者提供するためのシステムだ。数千万人、数十億件規模の医療情報を蓄積・提供するため、大規模データベースの構築技術と高度な匿名化技術を求めている。

### データの有用性と安全性を両立する匿名化技術を採用

このシステムを構築する準備として、日本医師会ORCA管理機構は2018年度上期に、日鉄ソリューションズ（以下、NSSOL）を含む3社のITベンダーと共同でデータの匿名加工に関する共同研究を実施した。ここでの成果を踏まえ、日本医師会ORCA管理機構はシステム構築のパートナーにNSSOLを選んだ。決め手は、ITベンダー各社の匿名加工技術を比較検討した結果、既にパッケージを有する他社と異なり、NSSOLはゼロからの開発となるため、柔軟な機能実装が期待できたことと、大規模データ処理の要となる分散処理技術（Hadoop、Impalaなど）に精通していたことだ。2018年10月に構築プロジェクトを開始。大規模システムの構築や匿名化に関するNSSOLの支援を受けながら、2019年3月に予定通り匿名加工エンジン（匿丸）を完成させた。

### 2020年の事業開始に向けた準備が最終段階に

日本医師会ORCA管理機構は、3000万人、数十億件に上る医療情報の蓄積に対応できるデータベースを構築し、第三者提供先のニーズに合わせて個別に匿名加工できるシステム基盤を整えた。現在は、医療情報の第三者提供事業を運営するために別途設立された一般財団法人が、次世代医療基盤法の認定事業者となるための申請を準備している段階である。

日本医師会ORCA管理機構は、日本医師会や事業運営を担う財団法人とともに2019年中にも医療情報の収集を開始し、2020年からの事業開始を目指す。これにより医療情報の安心安全な利活用が進み、医療の向上や新薬の開発が促進されることを期待している。同社は引き続きNSSOLと協力しながらシステムの強化を進める考えだ。

## Key to Success

日本医師会ORCA管理機構が目指したのは、安心安全な医療情報基盤の構築である。「一部の民間企業は、グレーゾーンで個人情報を収集し、販売しているようです。しかし、医療情報は取り扱いを間違えると、大きな病気や遺伝的な病気にかかっている人の差別につながる恐れがある。日本医師会は公益法人として、国民と医師の双方が納得する、透明で安心安全な医療情報の利活用を促進するために、自身の手で医療情報の提供事業に取り組むべきだと考えました。それが今回のプロジェクトの発端です」と代表取締役社長の上野智明氏は背景を語る。

この基盤を実現するには高度なノウハウが不可欠だった。事業推進部部長プロジェクトマネージャーの伊藤伸昭氏は、「最も重視したのは匿名加工技術です。個人が特定できないような加工が必要ですが、過度に加工すると研究目的でデータを利活用する時の有用性が下がってしまいます。そこで、次世代医療基盤法のガイドラインに則って十分な安全性を確保しながら、用途に合わせて有用性を最大限に確保するためのノウハウが欠かせませんでした。また、データ量が数十億件になっても確実に運用できるデータベースの構築ノウハウも必要でした」と話す。

同社がNSSOLをITパートナーに選んだ理由は、システム研究開発センターの高度なノウハウと開発の熱意を兼ね備えていたからだ。「プロジェクトの前に実施した共同研究の時から、NSSOLは当社の匿名加工に関する要望に熱心に耳を傾け、一緒に議論しながら柔軟に実現してくれました。匿名加工によって安全性と有用性がどう変化するかを定量的にレポートする指標

の開発や、NSSOLが独自開発した匿名加工アルゴリズムなどを特に評価しました。データベースの構築についても、Hadoopをベースに構築することも決めていましたが、問題はその構築技術でした。NSSOLはこの分野で豊富な構築経験と知見を持っていました」と事業推進部事業戦略アドバイザーの工藤憲一氏は語る。

### 世界的に価値の高い基盤が完成 NSSOLが期待以上の貢献

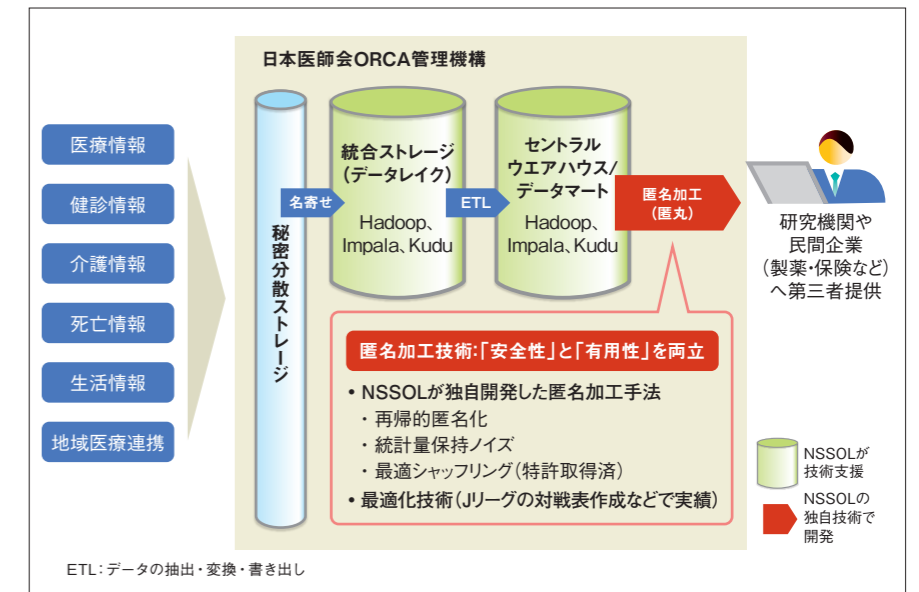
外部の要因によってプロジェクトが計画通りに進まなかったこともあったが、新システムの中核となる匿名加工エンジンは当初の予定通りに完成した。「生涯保健情報統合基盤は安心安全で確実な稼働が求められるシステムです。そこは製鉄のシステムと同様であり、NSSOLはその知見を生かし

て構築してくれました。プロジェクトのために、こちらの期待以上に人や技術を出してもらったので、頭が下がる思いです」と伊藤氏は評価する。

新事業を開始するのは少し先になるが、生涯保健情報統合基盤がもたらす成果への期待は大きい。「日本は医療情報の利活用で遅れていましたが、これで海外と対等な研究開発を進められます。また、この基盤は海外でも十分に利活用されていない軽症・安定期のデータや介護のデータを収集し、名寄せしたうえで統合します。このような生涯にわたる医療情報の提供は世界初とっていい取り組みであり、非常に大きな価値を生むと考えています」と工藤氏は力説する。

今後について上野氏は、「医療機関や自治体から粒度の異なる情報を集め、データベースに取り込んでいくという大変な作業が待っています。NSSOLには引き続き提案や支援をお願いしたい」と述べる。

### ■日本医師会ORCA管理機構が構築した生涯保健情報統合基盤の概要



### ■コアテクノロジー

データ匿名化に関する業界トップクラスの知見、データベース/Hadoopなどの構築・活用ノウハウ

### ■システム概要

- データベース：Hadoop、Impala、Kuduなど
- アプリケーション：匿名加工エンジン（匿丸）