

# 医療ビッグデータの利活用によるサービス創出の枠組み

## Framework for Service Creation through Utilization of Medical Big Data

星野 隆之

**要約** 医療ビッグデータを利活用し、ビジネスを連携させ、新しいサービスを創出するための枠組みを定義した。

これは、単なるデータの分析手順を示すものではなく、データの利活用目的を明確にし、分析を行うことで価値を抽出し、実際の事業やサービスを運用するまでのプロセス、基本アーキテクチャ、社会システムモデルまでをカバーしている。

この枠組みを参照し、社会システムをデザインすることで、医療ビッグデータを中心とした事業の連携と、サービスの創出が実現できると考えている。

**Abstract** This paper describes the framework which was defined to create new service by utilizing the medical big data and in cooperation with business.

This framework not only represents the procedure of data analysis but also clarifies the aim of data utilization, extracts value from data by analysis and covers the process before doing business and service, and the basic architecture and social system model.

With reference to this frame, I think that I can realize cooperation of the business mainly on medical big data and the creation of the service by designing a social system.

### 1. はじめに

日本ユニシスの研究開発部門である総合技術研究所では、2010年度より、事業部門とともに、疫学研究を題材として、健康・医療分野におけるビッグデータ（以下「医療ビッグデータ」）を利活用するための情報基盤構築に取り組んできた。

疫学研究は、地域住民などの集団を対象に、継続的なデータ収集を行い、疾患と体質、生活習慣、生活環境などの関係を明らかにし、個人に合った生活の改善方法、予防法、治療法の開発や適用に結びつけることを目的とした研究活動である。

医療ビッグデータは、一般的なビッグデータが持つ、量、多様性、迅速性、正確さという特長に、複雑性、階層性、不確実性が加わったものである。また、一般のビッグデータの世界では、因果関係ではなく相関関係を重視するのに対し、医療ビッグデータの世界では、相関関係から導き出された結果を意思決定につなげる際に、慎重な対応を必要とする。特に医療的介入を行う場合には、誤った判断は人の生死に関わる問題に直結するため、相関の背景（因果関係）への注目が欠かせない<sup>[1]</sup>。医療ビッグデータの利活用には、このような点に対応する必要がある。

現在、健康・医療情報のデータ収集・統合管理のための基盤や、データ解析基盤の構築を行っており<sup>[2][3][4]</sup>、さらに、事業やサービスにつなげるための分析モデルの開発や社会システムの検討も開始している。

本稿では、これまでの取り組みを通して得られた、医療ビッグデータを利活用し事業やサー

ビス創出につなげるためのプロセスと基本アーキテクチャをそれぞれ2章と3章で考察し、4章にて具体的な社会システムの例を示す。

## 2. データを利活用するということ

これまでのデータマネジメントは、ある品質のデータセットを、必要とするときに必要な量だけ適正なコストで取り揃え、要求元へタイムリーに供給するための管理活動だった。クラウドコンピューティングの発展につれて、データ発生源の重心が組織の外部に移り、個々の組織で一元管理できなくなりつつある中、総合技術研究所では、「膨大なデータから、未来を先読みし、個別に役立つ」情報を提供するため、従来から蓄積しているデータ分析技術に加え、データ統合環境やデータモデリングの研究を進めてきた。

大量に蓄積されたデータは、一般には、原データのままで得られる価値は少なく、適切な手法に基づく解析などを施し、解釈を加えることにより、広く受け入れられる価値を創出できると考えている。

本章では、健康・医療領域を例に、データを利活用することの意義について整理し、3章で医療ビッグデータを利活用して事業に適用するための枠組みについて定義する。

### 2.1 データを分析することによって得られる価値

「膨大なデータから、未来を先読みする」ためには、図1にあるように、まず、着目すべき結果（アウトカム）と、収集したデータの中から関係を持つと思われる予測因子を判別し、その関係（相関関係および因果関係）を定義して、予測因子を使ってアウトカムを説明する必要がある。そして、その関係式に、現在のデータを当てはめ、将来の状態を予測/判断する。

もし、何らかの手を打つ（介入する）ことにより、「今、予想される将来の状態」よりも「望ましい将来の状態」を得ることができれば、その差分が個人にとっての価値となる。つまり、分析し、予測や判断を行い、実際の行動を起こすことによって、価値が創出されるということになる。

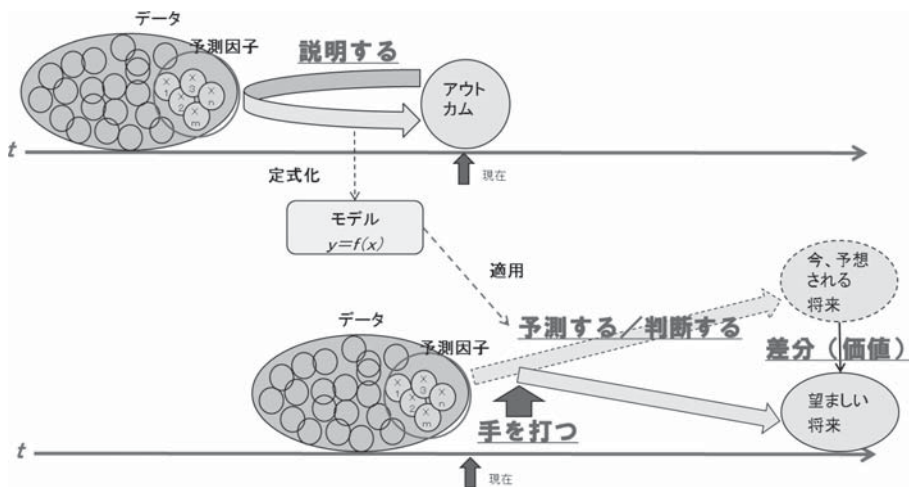


図1 データの分析と価値の創出

## 2.2 分析成果を事業に結びつけるためのモデルの適正化

データの分析により、価値を生み出す源泉となる関係式が定義できた場合、図2に示すような、事業化に向けた「モデルの適正化」が一般には必要である。

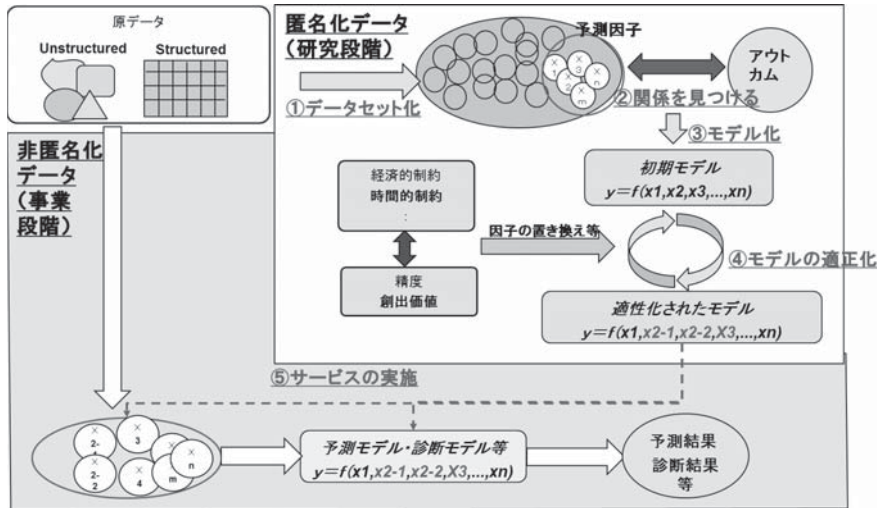


図2 事業に向けたモデルの適正化

アウトカムに着目して、原データから予測因子候補となる変数をデータセットとして抽出する。そして、データを分析して、因子とアウトカムの関係を見つけ、定式化する。これを初期モデルと呼ぶ。膨大なデータから、未来を予測し、個別に役立つ情報を提供することを考えた場合、初期モデルが事業化に適さないことが考えられる。具体的には、予測因子となるデータ値を求める際に、高価な計測機器が必要となりコストがかかるケース、計測自体に数日/数週間の時間が必要となるケースなどが考えられる。その際、簡易な計測手法を採用したり、他の指標で代用するなどの因子の置き換えを検討する必要がある。因子の置き換えにあたっては、予測/診断結果の精度や価値の低下を考慮する必要もあり、これを経てモデルを適正なものにする。実際のサービスは、各個人に対し計測されたデータと適正化されたモデルを使って提供され、予測結果や診断結果等の形で提供される。

データセットを抽出しモデルの適正化を行うまでの過程を「研究段階」、そして、適正化されたモデルに基づいて予測結果や診断結果を提供する段階を「事業段階」と呼ぶことにする。この二つの段階は、研究活動と事業活動といった活動性質の違いの他に、扱うデータが匿名化されたもので可能かという違いもある。研究段階では、アウトカムと因子の関係を、適切なサンプル数のデータを用いて求めることが重要であり、個々のサンプルの区別は必要ではあるが、個人の特定は必要ない。そのため、匿名化されたデータで十分実行が可能である。それに対し、事業段階では個人に予測結果や診断結果をフィードバックするため、個人の特定ができるようなデータ<sup>\*1</sup>を使用する必要がある。

## 2.3 データ利活用のプロセス

これまでの考察を踏まえ、サービス段階を含めて、データ利活用プロセスを整理したものが、図3である。

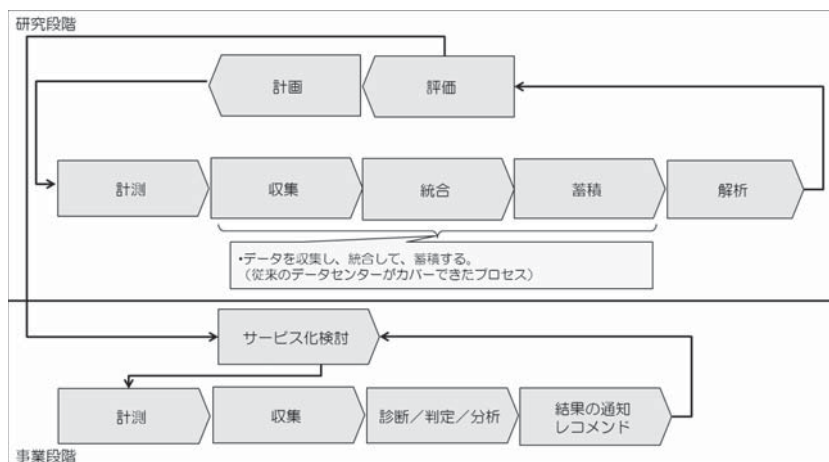


図3 データ活用プロセス

研究段階において必要となるデータは、計画に基づいて測定（計測）され、複数のデータが適切な頻度、タイミングで収集、統合され蓄積される。新たな技術やサービスの源泉を発見したり、アウトカムとの関係の有効性を検証するために解析が行われ、その結果の妥当性などを評価することにより、測定項目や測定方法の改良・改善を検討する。計画～評価までのサイクルを適切な回数実施した後、事業段階に移行する。

事業段階では、適正化されたモデルを使ったサービスを検討し、サービスを受ける個人に対するデータ測定（計測）を行う。測定されたデータを収集し、モデルを用いた診断、判定、分析を実施し、その結果を利用者に提供する。また、利用者からのフィードバックにより、サービス自体の改善などの検討も行う。

これまで、データの利活用においては、収集/統合/蓄積が、情報基盤が担う範囲として認識されており、いわゆるデータセンターがその役割を担っていた。しかし、利用者や顧客視点のサービス提供を考えた場合、上記のプロセスをサポートする基盤（データバンク機能）が必要になると考える。

#### 2.4 データ利活用を実現する社会システムモデル

健康・医療領域を題材として、データ利活用に関わるステークホルダーと、各々が享受する価値について整理した図を示す（図4）。これは、疫学研究を中心として、その成果がサービスとして展開され、個人に価値が還元される仕組みを示したものである。

個人（疫学研究に参加している住民等）は、医療機関からの診療データや健康診断結果等の健康・医療情報をデータバンクに登録する。研究機関は、データバンクより匿名化された情報を入手し、疫学研究などの健康・医療関連研究を行う。研究機関での成果は、新たな技術として医療機関に提供する他、生活習慣に関わる予測モデルやリスク評価モデルとしてサービス提供者にも提供する。それらの成果は、医療機関、行政からのサービスや、事業/サービス提供者からの個人向けの健康サービスとして、住民に還元される。

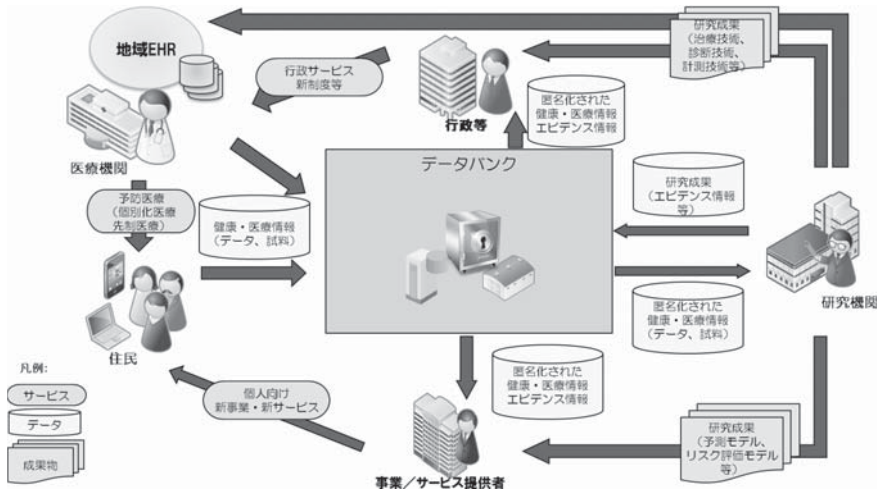


図4 健康・医療領域におけるデータ活用

このように、データの利活用のしくみは、業種を越え、複数の異なる業種の事業者が連携したビジネス・エコシステム（生態系）を形成することにつながり、人と社会に価値を提供することが実現できる。エコシステムを形成し、ビジネスをつなげて、単独の事業者ではできない新しいビジネスを創出するためには、ICTを上手く適用することが有効であると考える。

### 3. データを利活用するための基盤（データバンク）

本章では、2章で述べたビジネス・エコシステムを実現するための基盤（データバンク）の基本構造を整理する。基本構造を図5に示す。

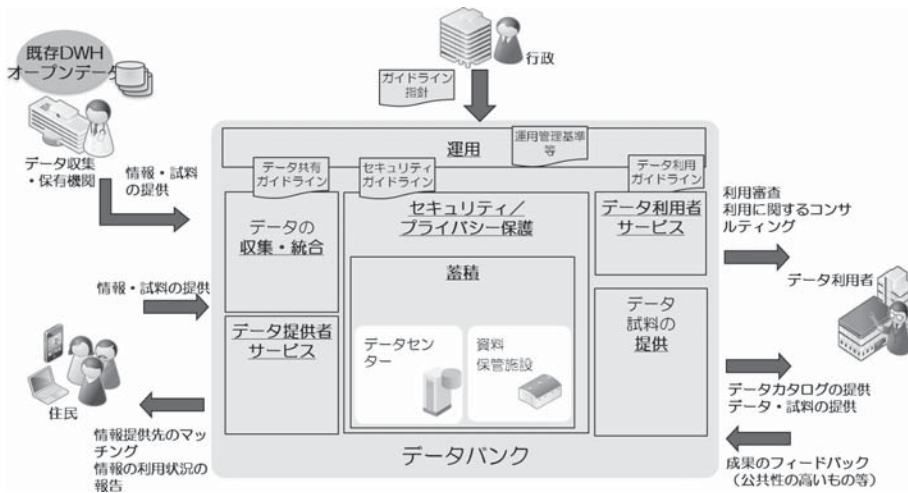


図5 データバンクの基本構造

#### 3.1 機能

データバンクは、住民およびデータを収集し保存する機関（既存のDWHやオープンデータ）から、情報や試料の提供を受け、収集・統合し、データセンターや試料保管施設に蓄積する。

研究機関やサービス提供者が、データバンクからデータや試料の提供を受ける際には、データバンクによる利用審査を経る必要があり、利用目的が適切と判断されれば、匿名化されたデータや試料を入手することが可能になる。

データ提供者に対しては、データの提供先や利用状況、提供によって享受できるサービスなどを提示する。また、データ利用者に対しては、データバンクが保有するデータや試料に関する

表1 データバンクの機能

機能	機能 (小区分)	概要
データ提供者サービス	マッチング	データ提供先（と享受できるサービス）に関する情報を提示し、データを提供する範囲を設定する。
	フィードバック	提供データがどのように利用されているかの情報を提示する。
	基本調査	統一された測定規格にもとづき、データ提供者に対する基本的な情報を採取する。
収集・統合	統合化	各データ源泉から提供されるデータについて、 <ul style="list-style-type: none"> <li>● 名寄せ（個人識別 ID の統合）</li> <li>● コード等の標準化</li> </ul> を行う。
	データ・キュレーション	収集したデータに対し、適切なレベルでの補正を行う。 収集したデータから、必要な項目を抽出する。 収集したデータの中から、不適当なものを除外する。
セキュリティ/ プライバシー保護	匿名化	収集、統合した情報において、個人を識別する ID を別の体系に変換する。 ※本機能は、他の機能分類とは独立した組織により実施する。 ※匿名化を行う組織は、ID 以外の情報は保持しない。
	セキュリティ	蓄積データに関するアクセス制御 データ漏えい防止（暗号化、分散化）等
提供	データカタログ提供	蓄積・管理しているデータ等についての概要（基本統計量、多次元分類、簡単な傾向分析などカタログ的な情報）を提供する。
	利用審査	利用計画を審査し、データの利用目的が適切であるかを審査する。
	データ提供	利用計画に基づき必要なデータ・試料を、匿名化を施した上で提供する。
	データ利用に関するコンサルティング	利用者に対し、利用計画とプラットフォームが提供できるデータ・試料の整合性などについてアドバイスやコンサルティングを行う。
蓄積	蓄積	収集、統合したデータ・試料を、データセンターや試料保管庫にて蓄積管理する。
運用	ガイドラインの作成と運用	データバンクの運営に必要なガイドラインの設定を行い、運用する。
	内部監査	データバンクが遵守すべきガイドラインに沿って運営されているかを内部的に監査する。
	外部監査	データバンクが遵守すべきガイドラインに沿って運営されているかを外部機関により監査する。
	倫理審査	データバンクの運営や利用者審査などにおいて、倫理的側面からの判断が必要な場合に、外部の有識者からなる倫理委員会により、その内容の審査を行う。

るカタログの提供や、利用にあたってのコンサルティングも行う。

データバンクの運営は、行政からの指針に基づくガイドラインに沿って行われ、データの収集・統合・蓄積には堅牢なセキュリティ対策とプライバシー保護が適用される。機能の詳細は、表1の通りである。

### 3.2 遵守すべき法令・ガイドライン

データバンクの運営にあたっては、セキュリティやプライバシー保護の対策が重要である。行政から提示される法令やガイドラインを遵守する形で、データバンクとしての運用ルール、ガイドラインを定義し、データ提供者、データ利用者に提示する。法令・ガイドラインに関するポイントを整理すると、以下のようになる<sup>\*2</sup>。

- 個人情報保護やプライバシー保護等を中心として、機能面以外での制度設計や組織設計などが必要となる。
- 「ICT」および「健康・医療領域」においては、図6のような法令、ガイドラインがあり、データバンクで扱うデータに応じて、同様に遵守すべきガイドラインを整理し、対応する必要がある。

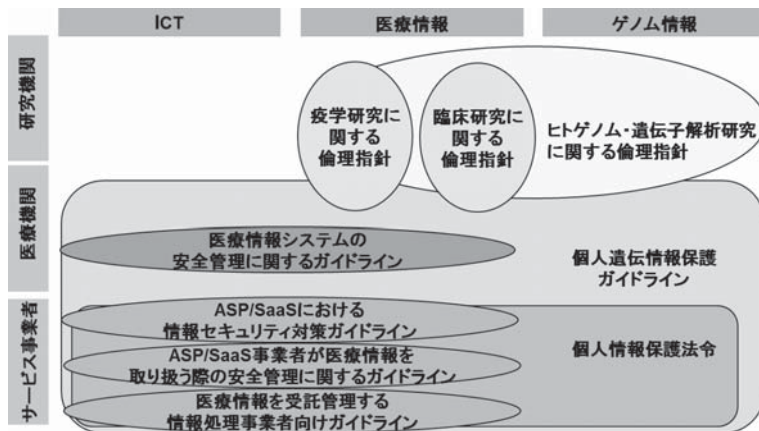


図6 遵守すべき法令・ガイドライン

## 4. スマートヘルシーシティ

本章では、データバンクを中核とし、医療ビッグデータを利活用して、サービスを展開する具体的取り組みを紹介する。

日本ユニシスグループは、社会課題を解決し、社会を豊かにするサービスを創造し、事業を推進するために、利用者密着/地域密着型のエコシステム「スマートヘルシーシティ」の実現に取り組んでいる。スマートヘルシーシティは、少子高齢化、独居老人の増加が進行している地域社会をささえる基盤であり、三つの分野（医療、食、決済）で地域、遠隔地（子供世帯）で暖かく見守るコミュニティを実現することを目指している（図7）。

健康支援機能は、健康情報と食の情報を基に、健康維持・改善の計画から実行・評価までをサポートする。健康情報と食の情報を活用することで、個人の状態に適したサポートを実現する。

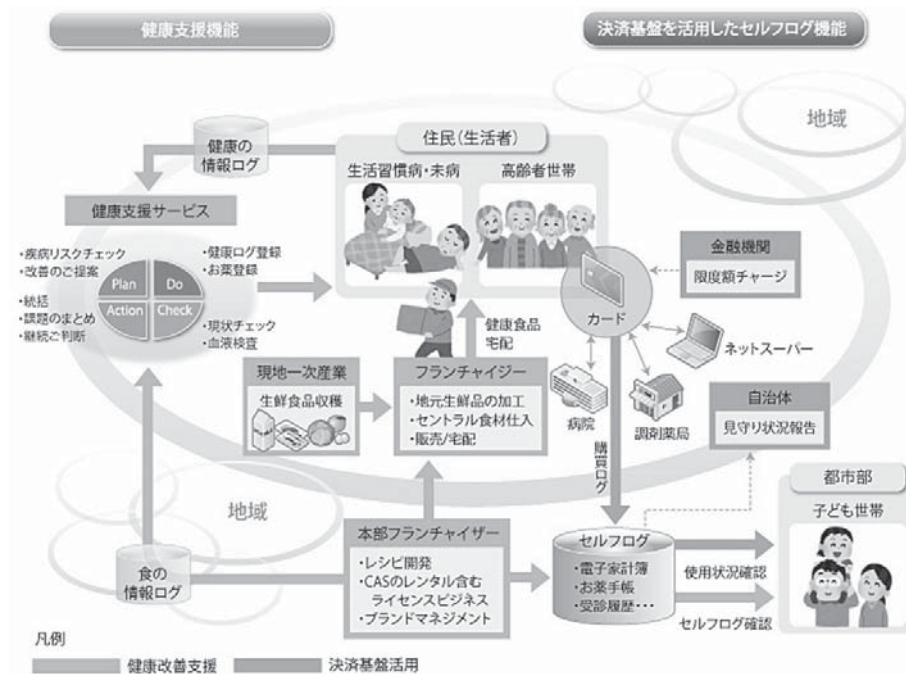


図7 スマートヘルシーシティ

健康支援機能の異業種連携の例の一つとして、冷凍技術事業者、食品提供事業者などとの連携がある。これは、食事による健康維持・改善を考えている住民に対して、CAS技術<sup>\*3</sup>を活用した冷凍食品を提供するというものである。CAS技術を活用した冷凍食品は、素材そのものの風味や触感が維持され、低塩分・低カロリーでも旨みを感じるため、味付けで塩分・カロリーが増すという問題を回避することが期待できる。

決済基盤を活用したセルフログ機能は、病院へ行った診察履歴や、健康食品の宅配サービスを利用した購買履歴、調剤薬局での薬の受け取りなど、生活に関わる各種決済情報をセルフログ化（見える化）することにより、健康支援サービスとの連携や、遠隔地（子ども世代）からも緩やかに見守ることを可能にする。

このように、複数の事業者のビジネスを連携させることにより、単一の事業者のビジネス範囲だけでは実現できない新たなサービスを、スマートヘルシーシティでは実現しようとしている。

## 5. おわりに

今後の高齢化社会を考えた時に、ICTで解決できる課題は、まだまだ多数存在すると思われる。特に、大規模データの利活用結果を、ICTを生かして利用者視点や顧客視点のサービスにつなげる取り組みは、今後も継続的に展開できるし、展開が求められていると考える。

農業の現場で生産調整という話があるが、先に紹介したCAS技術を応用すれば、せっかく収穫した農産物を捨ててしまうということは起きない。また、CAS技術と健康・医療系の取り組みや技術を融合し、医療分野における更なる応用も期待できる。



日本ユニシスグループは、2020年に向けた中長期的な成長領域として、「ライフイノベーション」を想定している。現在すでに、金融機関に新しい価値を提供するNEXT U's Vision、エネルギー管理の仕組みのHEMS/MEMS、地域医療連携ネットワークなどを提供しているが、今後、これらを利用者密着/地域密着型の視点でつなげることで、日々の買い物から、食とつながる医療や健康、家計やエネルギー管理のワンストップ化などを実現する地域統合サービスへ発展させる予定である。また、中長期では、次世代の技術として再生医療やセンサー技術などに早期に取り組み、これらを更に多面的につなげ、予防医療につながる食医農連携や地域生活視点のサービスを推進し、スマートヘルシーシティの実現を目指している。更に、今後、これらの活動で得た事業創出や運営のノウハウを蓄積することで、新たな事業の創出を加速させることも想定している。

本稿が、これらの取り組みのベースとなる考え方、枠組みとして、理解・活用いただければ幸いである。

最後に、医療ビッグデータの利活用、新サービスの創出に関する各種検討、議論に貴重なご意見やご指摘をくださった皆様、そして事例や取り組みに関する情報を提供いただいた方々に、心より感謝申し上げます。

- 
- \* 1 フィードバック対象となる個人が認識できれば十分であり、詳細な個人情報を必要とするレベルではない。
  - \* 2 2015年の個人情報保護法の改正にともなうパーソナルデータの扱いについては、グレーゾーンの解釈など、事業者が取り組むべき課題が存在する。詳しくは、参考文献[5]を参照されたい。
  - \* 3 Cells Alive System. 株式会社アビーが開発した細胞を生かす技術で、凍結装置と組み合わせることによって細胞破壊を最小限にとどめ、凍結時の鮮度・風味・旨味を維持する技術。

- 参考文献**
- [1] 中山健夫, 21世紀医療フォーラム「医療ビッグデータがもたらす社会変革」, 日経BP社, 2014年5月
  - [2] 星野隆之, 「大量かつ複雑な非構造化データを扱う解析基盤の仕組み」, ユニシス技報, 日本ユニシス, Vol.31 No.4 通巻111号, 2012年3月
  - [3] 福田健太, 「ヒト生命情報統合研究を支えるICT活用」, ユニシス技報, 日本ユニシス, Vol.34 No.2 通巻121号, 2014年9月
  - [4] 沖俊吾, 「大規模ゲノム疫学研究の統合情報基盤の構築事例」, ユニシス技報, 日本ユニシス, Vol.34 No.3 通巻122号, 2014年12月
  - [5] 八津川直伸, 「プライバシー・バイ・デザインに基づく適正なパーソナルデータの取り扱い」, ユニシス技報, 日本ユニシス, Vol.34 No.4 通巻123号, 2015年3月

**執筆者紹介** 星野 隆之 (Takayuki Hoshino)

1985年日本ユニシス(株)入社。人材育成部門、データ利用技術部門、電力事業部門で情報系のシステム開発プロジェクトに従事。2006年よりR&D部門において、データエンジニアリング関連の研究開発プロジェクトに従事。現在、総合技術研究所イノベーション室に所属。PMAJ公認PMレジスタード。

