

The **Beyond Disciplines** Collection

デジタルトランスフォーメーションに伴う 科学技術・イノベーションの変容



国立研究開発法人科学技術振興機構 研究開発戦略センター
Center for Research and Development Strategy, Japan Science and Technology Agency

はじめに

JST 研究開発戦略センター（CRDS）では、重要な研究開発戦略を立案しようとする過程で、その土台となる主要な科学技術分野を俯瞰する活動に取り組んできた。この結果は、「研究開発の俯瞰報告書」として隔年発行している。さらに、社会や政策等の動向を踏まえた上で分野を越えた全体像として捉えることを目指し、俯瞰報告書の「統合版」を 2019 年に発行した。

その中では、CRDS がこれまで俯瞰対象としてきた主要 4 分野（環境・エネルギー、システム・情報科学技術、ナノテクノロジー・材料、ライフサイエンス・臨床医学）の共通の課題として、「ビッグデータ活用・スマート化」、「データ駆動・デジタルトランスフォーメーション」、「研究環境・イノベーションエコシステム形成」、「オープン化」、「研究プラットフォーム化」、「グローバル化」、「研究関連人材」等のキーワードが抽出された。

ビッグデータや AI が世界規模で話題になり始めてから、おおよそ 10 年が経過した。これらのデジタル技術の活用がもたらす新たな価値と変革、すなわち「デジタルトランスフォーメーション」の観点から、様々な研究分野においてどのような科学研究、研究開発が行われているかを横断的に俯瞰したものが本報告書である。第 2 章では研究開発の現場における自然科学の側面の強い研究分野群を、第 3 章では、個人や社会・環境（法規制含む）の要素がパラメータとして大きい自然科学と人文・社会科学の共創が必要となる研究分野群を取り上げた。第 4 章はデジタルトランスフォーメーションを駆動する共通基盤技術を、第 5 章は関連する国内外の政策動向をまとめた。

科学や研究現場へのデジタル技術の浸透により、21 世紀の科学技術はそのあり方を大きく変容させており、いまは分岐点にある。現在はそれぞれの研究分野で独立して取り組まれている手法や研究成果・社会的価値がリンクして新たな価値・社会を創出していくことも展望される。

この報告書が、第 6 期科学技術基本計画を始めとする科学技術政策に資するとともに、科学技術と社会、経済・産業、国際協力等のより望ましい関係構築・発展へ向けた基礎的知見として活用されることを期待したい。

令和 2 年 4 月
国立研究開発法人科学技術振興機構
研究開発戦略センター

【目次】

はじめに	1
1 総論：デジタル技術の進展と科学技術イノベーション	5
2 研究開発の現場にデジタルトランスフォーメーションがもたらす変容	8
2.1 ライフサイエンス（X オーム・オミクス、脳神経研究）	8
2.2 創薬	14
2.3 バイオ生産	19
2.4 物質・材料	23
2.5 ものづくり	28
3 フィールド・社会に関する研究開発とデジタルトランスフォーメーション	34
3.1 医療	34
3.2 ヘルスケア	39
3.3 農業	44
3.4 生物多様性	48
3.5 気象・気候	53
3.6 エネルギー	59
3.7 都市	64
3.8 政策立案	70
4 デジタルトランスフォーメーションの基盤技術	77
4.1 データ基盤フレームワーク・システム	77
4.2 データ処理技術：AI・機械学習	81
4.3 データ取得技術：物理・化学的（量）計測	86
4.4 データ取得技術：センサー・ロボット（環境・社会計測）	90

5 国内・海外の関連政策	96
5.1 日本	96
5.2 米国	102
5.3 欧州連合（EU）	105
5.4 英国	107
5.5 ドイツ	108
5.6 中国	110
6 まとめ	111

1 総論：デジタル技術の進展と科学技術イノベーション

デジタルトランスフォーメーションとは、2004年にウメオ大学（スウェーデン）のエリック・ストルターマン教授が提唱したとされる「ITの浸透が、人々の生活をあらゆる面でより良い方向に変化させる」という概念であるが、本書では「新しいデジタル技術を活用することによって、新たな価値を生み出していくこと」と広く定義する。まず、この文脈で関連のある主な出来事を振り返る。

1990年代の情報通信革命以降、インターネットの出現や様々なハードウェア・ソフトウェア技術の進展により、人間が扱う知識、データの量は指数関数的な増加を見せている。

2008年にNature誌が「BIG DATA」（サイエンスはペタバイトのデータを扱う時代へ）と銘打った特集を組んだ。2009年にマイクロソフト社は、「Data-intensive computing」を、科学研究手法の第4のパラダイムとして紹介し、理論や実験、シミュレーションに続き、「データ科学」が第4の科学として脚光を浴びた。その後2011年頃の「ディープラーニング（深層学習）」の登場により、第3次「人工知能（AI）」ブームが起こった。その後もディープラーニングによる機械学習は着実に進化しており、人間の認識能力を超えるようなケースも生じてきている。こうした中で、世界のサイバー空間データの覇権争いを制したのは、米国大手IT企業の「GAFA（Google、Apple、Facebook、Amazon）」と呼ばれる企業群であった。

ビッグデータ分析を実現する情報システムとして、2006年にGoogle社CEOのエリック・シュミット氏が「クラウドコンピューティング」に言及し、コンピューター資源をサービスの形で提供する利用形態が注目を浴び始めた。「CPS（Cyber-Physical System）」は2006年頃、米NSFが提示した概念であり、「IoT（Internet of Things）」は身の周りのあらゆるモノがインターネットにつながる概念として、スマートフォンやセンサーの普及により、2010年頃から用いられるようになった。これらは、実世界（フィジカル空間）にある多様なデータをセンサー等で収集し、サイバー空間で大規模データ処理技術等を駆使して分析・知識化を行い、そこで創出した情報・価値によって、生活・産業上の効率化や新サービス、社会問題の解決を図っていくという点で共通する。2012年頃に提唱された米General Electric社を中心とした取組である「インダストリアル・インターネット」や、ドイツの推進する「インダストリー4.0」は、オープンイノベーションを通じて、製造業の新しい形態を先導しようとする戦略に他ならない。

リアルワールドデータ、フィジカルデータの収集・蓄積を担う様々なハードウェア、デバイスの動向はどうであろうか。MEMSセンサーやイメージング技術、次世代シーケンサーなど「計測技術」の進展は目を見張るものがある。デジタル化の進展により、ハイスループットに実測データが取得できるようになってきている。取得データの形態がスペクトルや画像等であるような計測技術は、機械学習と相性が良いと言われている。2010年前後には世界的にスマートフォンが普及し始め、人の活動情報を把握するセンサー、コンピューターとしての機能ももつようになってきている。2012年頃から「ドローン」が社会・産業の場に進出し始め、農業や測量、物流、災害現場の調査などで導入が進んでいる。なかでも、自動運転は究極の形態と言える。自動運転車は、「カメラ、レーダー、LIDAR、超音波センサー、GPS等」で周囲の環境を認識し、これらの膨大な取得データをGPGPU（General-purpose computing on graphics processing units; GPUによる汎用計算）やディープラーニング等により瞬時に処理して走行する。科学研究の現場では、「ロボット」などで研究活動そのものを自動化することにより、データを安定的にハイスループットに取得、解析、そしてフィードバックする試みも出始めている。

一方で、元々は軍事用や気象用での使用が主であった「スーパーコンピューター」が、半導体、並列化技術の進展により、2000年以降、大学や民間の研究開発に使われるようになってきた。次世代スーパーコンピューターは、一般的なコンピューターでは100年かかる計算が1年でできるという高速性だけでなく、複雑な計算ができることが特徴である。物理・数理的モデルに基づき、探索、最適化、シミュレーションなどができるが、これを使うことで各種の社会課題の解決が期待されている。例えば熱核融合炉のシミュレーションや、新薬の発見など多々ある。数年後、10PetaFLOPSの計算性能を個人レベルでも使用できる社会が訪れるかもしれない。また、「量子コンピューター」の1形態である量子ニューラルネットワークも、組合せ最適化問題の計算で重要な存在となる可能性が示唆されている。

通信ネットワークでは、2020年から日本でも第5世代移動通信システム（5G）が開始される。より円滑な通信やIoTを想定した多数の同時接続を前提に設計・開発されているため、多くの新サービスや産業のスマート化が実現するといわれている。

2018年からEUの「一般データ保護規則（GDPR）」が施行され、これまで以上に個人データの保護が重視され、取扱う者はその管理責任を負うようになった。EU域内でビジネスをおこなう日本企業も対象である。GAFAのようなプラットフォーマーに対して多くの国・地域が、域外データ移転の制限や独占禁止法による規制等に出ている。またフランスの上院は大手ITサービス企業に課税する「デジタル課税」を2019年7月に可決し、他国も追随する動きが見られる。

こうした流れの中、科学コミュニティにおいても、世界的に「オープンサイエンス（オープンアクセスと研究データのオープン化（オープンデータ）を含む概念）」の再検討が進んでいる。

日本では、第5期科学技術基本計画（2016-）や統合イノベーション戦略において、わが国が目指すべき未来社会の姿としてSociety 5.0の実現を提唱している。Society 5.0は、サイバー空間とフィジカル空間を高度に融合させることにより、経済的発展と社会的課題の解決を両立する、人間中心の社会を描いたものである。

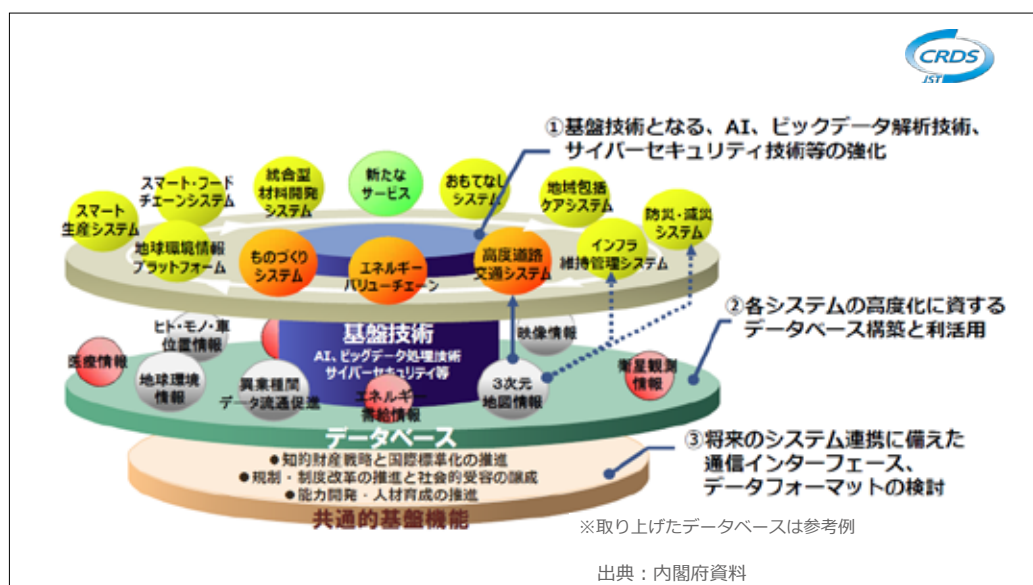


図 1-1 「Society 5.0」プラットフォーム構築のイメージ

ビッグデータや AI が世界規模で話題になり始めてから、おおよそ 10 年が経過し、これらデジタル技術の進展を受けて、科学研究や企業における製品・サービス開発のあり方も大きく変わりつつある。

デジタル技術が科学研究や研究開発等にどの程度浸透してきたか、何が課題（ボトルネック）となっていて、今後の展望はいかなるものか、様々な研究分野における相違や、世界と日本の相違を概観していく。

2 研究開発の現場にデジタルトランスフォーメーションがもたらす変容

2.1 ライフサイエンス（X オーム・オミクス、脳神経研究）

(1) 定義・意義

生命現象や疾患の理解、創薬を含めた治療法の確立において、個別の遺伝子や分子を一義的な原因、標的として解明する従来のアプローチはほぼ限界に達しており、多因子をシステムとして捉えることが望まれている。このように生体内の各種分子を網羅的に解析し、生命現象を包括的に理解しようとする手法をオミクスと呼び、ゲノム、エピゲノム、トランスクリプトーム、プロテオーム、メタボローム、マイクロバイオームなど、各階層におけるオミクスデータが大量に得られつつある。これら個々の階層におけるオミクス情報により生命現象を全体的なシステムとして理解・制御し、医療・創薬などに応用する未来像が浮かび上がっている。

ヒトの脳の基礎は 1000 億個に及ぶ神経細胞が 1000 兆個に達するシナプスによって相互に接続する神経回路網にある。脳機能は、神経回路ユニットがさまざまな脳領域において情報を処理することによって発揮される。脳の機能を理解するために、他のオミクスのアプローチに加えて神経回路の構造的・機能的結合状態の理解が必要とされることが、脳科学の特徴である。複雑なシナプス結合の情報を網羅的に記述するのがコネクトームであり、構造的・機能的コネクトーム解析により疾患の原因究明、治療法確立などに繋がると考えられる。喫緊の課題である高齢社会での認知症、社会的ストレスとうつ、ギャンブルやネット依存症の増加など、脳科学を基盤とした精神神経疾患の克服に向けて脳科学の成果を活用することが期待されている。

これらの期待に応えるために、学際性・多階層性に対応した技術開発と研究システムが必要となっており、ビッグデータに対峙する生命科学はデジタル化、コンピューテーショナル化が進んでいる。

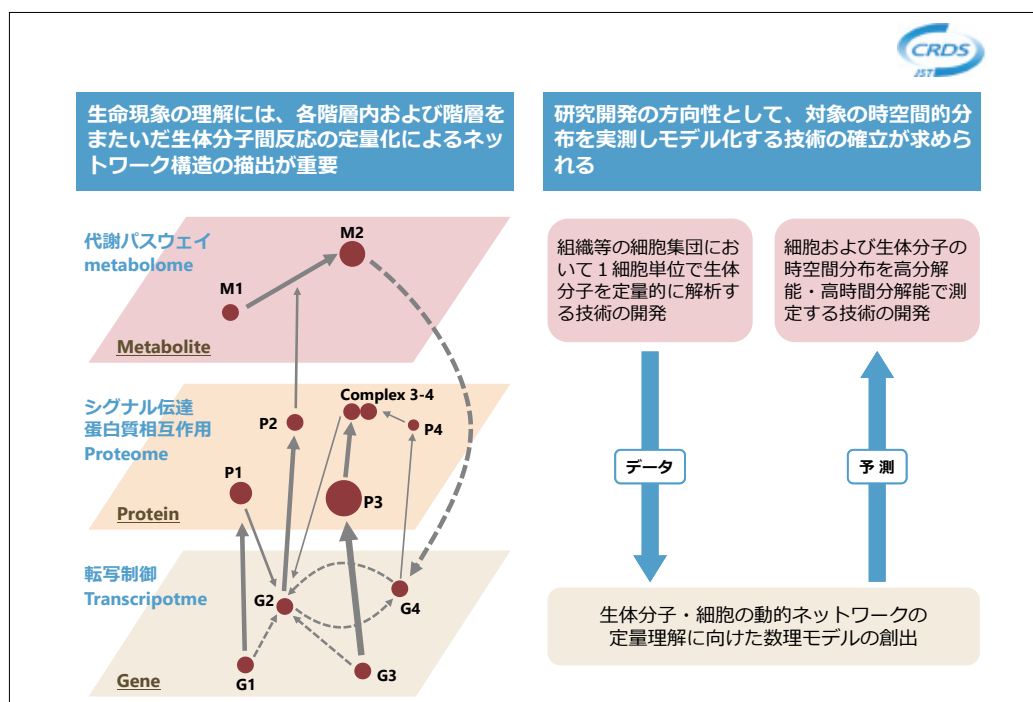


図 2-1 階層をまたぐオミクス情報の統合により生命現象を全体的なシステムとして理解・制御しようとするアプローチ

(2) 研究のトレンド

単独のオミクスとして遺伝子を対象としたゲノミクスが最初に発展し、2000年代前半にはヒトゲノム遺伝多型のカタログが作られ、2014年にはいわゆる1000ドルゲノムが達成された。

一方で、人体は推定37兆個の細胞により構成されていると言われていたが、その一つひとつの細胞のゲノムに生じた異常が、がん発症の原因とされ、結果として、腫瘍は、多様な遺伝子変異を持つ不均一ながん細胞集団から構成されていることが判明している。他にも、ヒト腸内細菌叢（マイクロバイオーム）は、約1000種類、約100兆個、約50万遺伝子（ヒト遺伝子は約2万個）と言われ、栄養供給（エネルギー（酪酸など）、ビタミン）、免疫系や代謝系の調節など、体内で様々な役割を担い、共生関係にあることが次々と示唆されている。こういった細胞の多様性やマイクロバイオームとヒトの健康、疾患との関係の理解は未だ途上にある。

近年のライフサイエンスと次世代シーケンサーや質量分析計などの計測技術の驚異的な進歩により、生物の遺伝子や酵素、代謝物等の相違や変化を包括的かつ網羅的に解析することが可能になってきた。単独階層のオミクス情報を複数階層にまたがるマルチオミクス、トランスオミクスに発展させることで、様々な分子の動的変化や個体差、生体ネットワークへの影響、薬剤応答などを予測する信頼性の高い方法を確立する試みが注目されている。ビッグデータを取り扱う数理・統計・情報学的手法の成熟にも後押しされ、ゲノム、メタボローム、プロテオーム、トランスクリプトーム、エピゲノム、マイクロバイオーム等、それぞれの階層に着目したプロジェクトが世界各国で実施されている。

プロテオミクスは、主に質量分析計を用いてタンパク質の発現、翻訳後修飾、局在、あるいは他の分子との相互作用を網羅的に計測する領域である。分離、分析技術の改良に伴って全細胞抽出物のような極めて複雑な試料から数千のタンパク質を一度に同定することが可能となっている。また、各種安定同位体標識による比較定量技術の確立によって、大規模な比較定量解析が行われている。タンパク質の機能に直結した情報取得もプロテオーム研究の重要な対象となっており（機能プロテオミクス）、タンパク質間の相互作用解析や、各タンパク質上で生じる各種翻訳後修飾の網羅的解析が盛んに行われている。

メタボロミクスは、生命活動によって生じる代謝物を網羅的に計測する領域である。表現型としては現れにくい生体内の変化を代謝物の変化として表現できる、高解像度のフェノタイピングシステムとして応用範囲の広がり期待が高まっている。代謝は多数の化学反応が相互に関連したネットワークを形成し、その相互作用を背景とした動的挙動を示すことが特徴のひとつである。そのため、パスウェイの制御を議論するには分子の量よりもそれを合成・分解する各反応の速度を知ることがより重要であり、近年メタボロミクスを含む各オミクス階層で速度を測定する技術が提案されている。

多階層にわたるオミクス情報を接続して理解するためには、数理情報学的手法が必須である。情報・数理・工学的手法を導入し、生命現象をシステムとして理解することを目的とした研究分野はシステム生物学と呼ばれ、2000年代から急速に進展した。大規模オミクスデータが得られるようになり、システム生物学を含むデータドリブンの解析はすでに世界的に用いられている。ネットワークそのものの構造を評価するような解析や機械学習を用いた解析など、新たな解析手法も開発されている。

個体の生命システム、そしてその破綻である疾患発症のメカニズムを解明するためには、細胞レベルの分子の振る舞いを包括的に理解することが必要不可欠であり、単一細胞レベルでの各種オミクス解析が行われている。それらのデータは生命現象の本質的な分子機構の解明に有用であるだけでなく、時空間的な分解能を持った細胞挙動の解析、臓器・個体レベルの機能と連動した解析へと応用されており、1細胞オミクス技術は極めて解像度の高い包括的な分子顕微鏡と考えることもできる。次世代シーケンサーを用いることで1細胞レベルのトランスクリプトームは実用化の段階に達している。特に、マイクロ流体技術と分子バーコードを利用してハイスループット化されたことにより大量の1細胞RNAシーケンスが可能となり、高出力型1細胞RNAシーケンスが盛んに実施されるようにな

りつつある。遺伝子と異なり増幅して分析することが困難なプロテオーム、メタボロームなど他階層オミックスの高感度化は、重要な課題となっている。

臓器において種々の細胞は綿密な空間的構造を構築して恒常性を保っており、空間情報を保持したまま遺伝子発現解析をすることで、疾患発症の分子病態の空間的なダイナミクスを解析できる。このような空間的関係性を理解する上で 1 細胞 RNA-seq 解析は大きな意義を持っており、1 細胞を単離、破壊することなく細胞や分子を「本来の場所で (in situ)」計測する手法の開発が盛んである。イメージングをベースとした 1 細胞レベルのトランスクリプトミクス技術 (in situ sequencing) として、組織切片上での細胞の空間分布を保持したまま、数百から数千におよぶ mRNA の発現量を 1 細胞レベルの分解能で計測する MERFISH や seq-FISH といった方法が開発されている。

脳科学分野では、ヒトと同じ霊長類であるマカクザルを対象にした光遺伝学・化学遺伝学技術の開発が進展しつつあり、霊長類が持つ高次脳機能の神経回路基盤についての理解が進むものと期待される。電子顕微鏡の自動撮像技術を用いたコネクトーム研究が進みショウジョウバエやマウス全脳の再構築が試みられており、ペタバイト規模の電顕画像情報から細胞構造を解読し神経回路構造を再構築するための AI を利用したソフトウェア開発がグーグル社を中心に進められている。

(3) 国内外のトピックス

• Cancer Genome Atlas

米国 National Cancer Institute (NCI) と National Human Genome Research Institute (NHGRI) の共同で開始されたがんゲノム解析プロジェクトである。合計 33 種類のがん腫、11,000 例を超える患者検体のゲノム、エピゲノムやトランスクリプトームなど統合的遺伝子解析が実施され、各がん腫において包括的に遺伝子異常の全体像が解明されただけでなく、それらの分子分類が提唱された。2018 年には、その成果をまとめて、細胞起源、共通のがん化プロセス、がん腫ごとの異常パスウェイの違いなどのテーマごとに Pan-Cancer Atlas として発表されている。さらに、治療に難渋することがある転移がんの全ゲノム解析も 2019 年に報告されている。

• プロテオゲノミクス

米国の Cancer Moonshot Initiative にサポートされて発足した International Cancer Proteogenome Consortium(ICPC) には日本を含む 13 カ国が参画しており、国際的にプロテオゲノミクスを用いてがん研究を推し進める機運が高まっている。特に、米国では 2011 年にスタートした Clinical Proteomic Tumor Analysis Consortium (CPTAC) によって、がん臨床検体におけるプロテオーム解析が先導されており、ICPC におけるプロテオゲノミクス研究も CPTAC チームが担当している。日本では国立がん研究センターが代表となり肉腫担当で参画している。

• Microbiome Project

米 NIH によって 2008 年に開始された Human Microbiome Project (HMP) は、鼻腔、口腔、皮膚、消化管、泌尿生殖器など、人体のさまざまな部位にまたがる 300 人の健康な個人の微生物群集を特徴づけるためのベースライン研究が行われた。2014 年からは、統合的ヒトマイクロバイオームプロジェクト (iHMP) として、微生物叢が人間の健康と病気になどどのように影響するかについての理解を進めるため、コホート研究から、ホストとマイクロバイオームの生物学的特性の統合縦断データセットを作成している。

• Human Cell Atlas

前述のように、単一細胞解析は生命現象の解析ツールとして確固たる地位を確立しており、技術開発によって空間情報を維持した遺伝子発現解析のハイスループット化が進展している。Human cell atlas は、ヒトの体を構成する全主要組織での 1 細胞トランスクリプトームにより、細胞種、細胞 3 次元位置、地理的、人種的な違いを考慮したヒト細胞のカタログ化を目指して 2017 年からスタート

している国際プロジェクトである。イメージング情報とオミクス情報が網羅的に集積されることで、ヒトの全細胞について、種類・状態・系統などを分類した細胞地図が構成される。

・脳機能の全容解明を目指した大規模プロジェクト

2013-2014 年にかけて、米国で The Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative、EU で Human Brain Project (HBP)、日本で Brain/MINDS と呼ばれる脳科学の大型プロジェクトが相次いで開始された。これらは、(1) 脳情報の読み出しや制御に関係する革新的な技術開発、(2) 脳全体の構造や機能のマップを作成し全脳レベルでの脳機能の解明を目指すこと、(3) 脳の理論や計算機科学を活用した融合型研究を目指すこと、(4) 得られた脳に関するデータを公開し、精神・神経疾患の克服に活用すること、といった点を共有する。トップダウン型で国際連携に重点を置いたプロジェクトを実施することが、脳の理解を飛躍的に進めるためには不可欠である、という共通の認識に基づいている。これらプロジェクトによりサポートされるプロジェクトの多くは生物学的神経科学領域でオミクスや局所・広域神経回路の構造的・機能的情報を網羅的に読み出すことを目的としていて、これらの技術革新により多階層にわたる膨大な計測データが今後産出されることになる。

・日本人多層オミクス参照パネル (jMorp)

東北大学東北メディカル・メガバンク機構 (ToMMo) は、ゲノミクス、プロテオミクス、メタボロミクス情報を網羅的に収載した日本人のリファレンスパネルを拡充、公開している。約 4,700 人を解析した全ゲノムリファレンスパネル 4.7KJPN は日本人のゲノムデータベースとして最大規模である。世界初となる経時的メタボローム変化データを含む、約 15,000 人分のメタボローム情報も公開され、疾患の原因探索、バイオマーカー発見など個別化医療、予防医療の進展に貢献することが期待されている。

(4) 課題と方向性

オミクス

マルチオミクス／トランスオミクスを実現するためには、それぞれの階層で統合を前提としたデータ取得方法の検討が必要である。例えば、ひとつのオミクスデータをもとに重要タンパク質と代謝物に対象を絞りそれらの定量を主として統合解析を進めるとともに、並行して従来のノンターゲットでの包括的な測定を行っておく。包括的な解析において重要分子と相関を示す分子が見つかった場合は、それらについて再度定量し、統合解析を行うといったように、それぞれのオミクス計測技術のスループット、データの性質を考慮したデータ取得が重要である。また、ゲノム、エピゲノム、トランスクリプトーム、プロテオーム、メタボローム、マイクロバイオームなどの階層間にわたる解析を、数学的／情報科学的な高度な手法をもって包括的に、人手や事前知識に頼らず、データ主導のバイアスの少ない方法で実行でき、そのためのアルゴリズム、ソフトウェアの開発を行い、さらにこれらを使いこなせる人材の育成が課題である。特に、階層内の組合せおよび階層間の組合せをいかに解析するかは極めてチャレンジングな課題である。

HCA で主に利用されている 1 細胞 RNA-seq では、非ポリ A 型 RNA や長鎖 RNA の全長が計測されず、RNA を標的とした核酸医薬、RNA 編集技術による創薬には貢献しにくい。1 細胞エピゲノム、プロテオミクス、メタボロミクスにおいても、感度、定量性、再現性、ハイスループット化が望まれている。1 細胞完全長トータル RNA シーケンス法である RamDA-seq、超高感度メタボローム分析法としてキャピラリー電気泳動・質量分析 (CE-MS) 分析法を高感度化する nanoCESI (Capillary Electrophoresis Sheathless Ionization) 法開発と代謝物を濃縮する LDIS (Large-volume Dual preconcentration by Isotachopheresis and Stacking) 法を組み合わせた方法の確立、高感度にエピゲノムを解析する手法であるクロマチン挿入標識法など、わが国はこの領域の技術開発に存在感を示している。

単一細胞解析において、位置情報との糾合および時間情報の取得も課題である。時間解像度に優れ

る（ライブセル）イメージングとの連携など異なるアプローチが求められる。情報科学的アプローチとしては、monocle や velocity など擬似的に時間情報を作り出す偽時系列解析、RNA velocity のようなスナップショットのトランスクリプトームデータから時間情報を抽出するような解析手法が提案されている。

高度な研究開発チーム（異分野連携）

従来よりもさらに生命医学研究におけるデータ解析、数理モデル解析、工学的デバイスの開発、低分子化合物の作成とスクリーニング等々、複数の分野の研究者の共同研究が重要になっている。現在、旧来の大学教育、大学院教育の枠組みのあり方、例えば理学部、工学部、薬学部、医学部といった古い枠組み自体がすでに時代の進歩においついていないのではないかという考え方もある。

また、オミクスは大規模データの解析に立脚しており、分野の進展には情報科学の研究者と技術者が必須であり、脳科学においても情報科学者との相互乗り入れができる仕組みが必要である。情報科学の研究者と医学・生命科学の実験系の研究者が研究デザインの段階から協力してプロジェクトを立案できる体制を整えることが望ましい。欧米ではそれらを複合し共同研究に主眼をおいた研究施設がすでに意味を持った形で機能し、成果を上げつつあるが、わが国にはほとんどない。そのため、大学や研究機関での組織編成を抜本的に見直すなど、計測・解析が表裏一体となって組織運営できるようにし、データ計測・解析のインフラ整備を行なうなど、より戦略的な取り組みが必要である。

データ基盤

生物・生命の理解の難しさとして、「非線形性：多階層生物ネットワーク」、「不均一性：細胞・疾患の不均一性」、「動的恒常性：疾患発症のダイナミクス」等が挙げられ、これに起因して、機械学習などを用いて、データ駆動型で生命を解析する際の生物分野のデータの取り方、蓄積の方法に難しい問題点が生じている。下記を考慮した上で高度に設計されたデータの取得をしなければ、研究に用いることは難しい。

1. 測定項目の次元数が非常に多いのに比べてサンプル数が少ない
2. 因子間の相互作用が非線形
3. 計測時刻が散発的で不規則、開始時刻が一定でない
4. 内部因子だけでモデルが完結しない（環境因子）

コアファシリティ

現状のオミクス研究は高額装置をその技術基盤としており、使用する装置の更新が国際的競争力の維持には重要である。さらに、ハイスループット化するには複数装置を並列で運用することが必要となる。しかし、これらの装置の導入費用やその維持管理費の捻出は大きな課題であり、大学等アカデミア研究機関の個別の研究室での対応は難しい。

また、オミクスやイメージング等の統合解析を目指した研究開発を進めるためには、単独の研究室で実施可能なレベルを越えている。海外主要国では、研究所内外のコアファシリティに技術移転を行い、トップラボは研究開発に集中しつつ国内のユーザーがデータを得られるように工夫されている。必要な技術の開発、その最適化、さらに上記のデータ取得に向けたプロトコル構築までの技術を一拠点に集約し実施することが望ましい。

人材育成

日本ではバイオインフォマティクスや脳科学分野での情報科学の重要性が高まっているにも関わらず、その人材が極度に欠如している。最大の原因は、大規模なバイオインフォマティクス研究所が存在せず、旧来の実験研究所や情報科学研究所の一研究室で大部分の生命情報科学教育が行われていることである。また、両分野を習得するには膨大な時間がかかり、米国の様にダブルメジャーの体制がわが国では不十分であることも課題である。

またウェットとドライの両者をバランス良く理解した人材も今後多数輩出される必要がある。オミ

クス技術の開発や利活用には、高度な情報処理が必須になっている。日本では、バイオインフォマティクス研究者がヘッドを務めるゲノム科学のラボや大型プロジェクトはほぼなく、国際競争力に乏しい。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 ライフサイエンス・臨床医学分野（2019年）」（CRDS-FY2018-FR-04）（2019年3月）
- ・ 科学技術振興機構 研究開発戦略センター「戦略プロポーザル “ライブセルアトラス” 多次元解析で紐解く生命システムのダイナミクス ～オミクス×イメージング×データ・モデリングによる基盤技術の創成～」（CRDS-FY2018-SP-09）（2019年3月）
- ・ 科学技術振興機構 研究開発戦略センター「戦略プロポーザル 微生物叢（マイクロバイオーム）研究の統合的推進 ～生命、健康・医療の新展開～」（CRDS-FY2015-SP-05）（2016年3月）

2.2 創薬

(1) 定義・意義

患者ヘルスレコードの電子化やゲノムをはじめとするオミクス情報の拡充などにより、医療健康ビッグデータは急速に拡張している。製薬産業では、バリューチェーンの全般にわたってビッグデータを利活用することによって、価値創造、生産性向上を果たすことに期待が寄せられている。このような状況の下、創薬プロセスの変革においてもビッグデータ、AIの活用が期待が高まっている。

低分子を中心とした druggable な創薬ターゲットが枯渇傾向にあると言われており、既存の手法の延長線上で新薬創出の成功確率が低下していることは大きな課題である。医薬品の研究、開発コストも年々上昇し、製薬産業のコスト体質を悪化させている。強化学習・自然言語解析等の AI における基盤技術は製薬産業の様々なビジネスプロセスに適用可能であると考えられることから、コストと時間短縮のため製薬企業で積極的な AI の導入が推進されている。また、この機を捉えて IT 企業が AI 技術を用いて創薬事業へ参入する動きが加速しており、製薬産業全体に大きな波が押し寄せている。

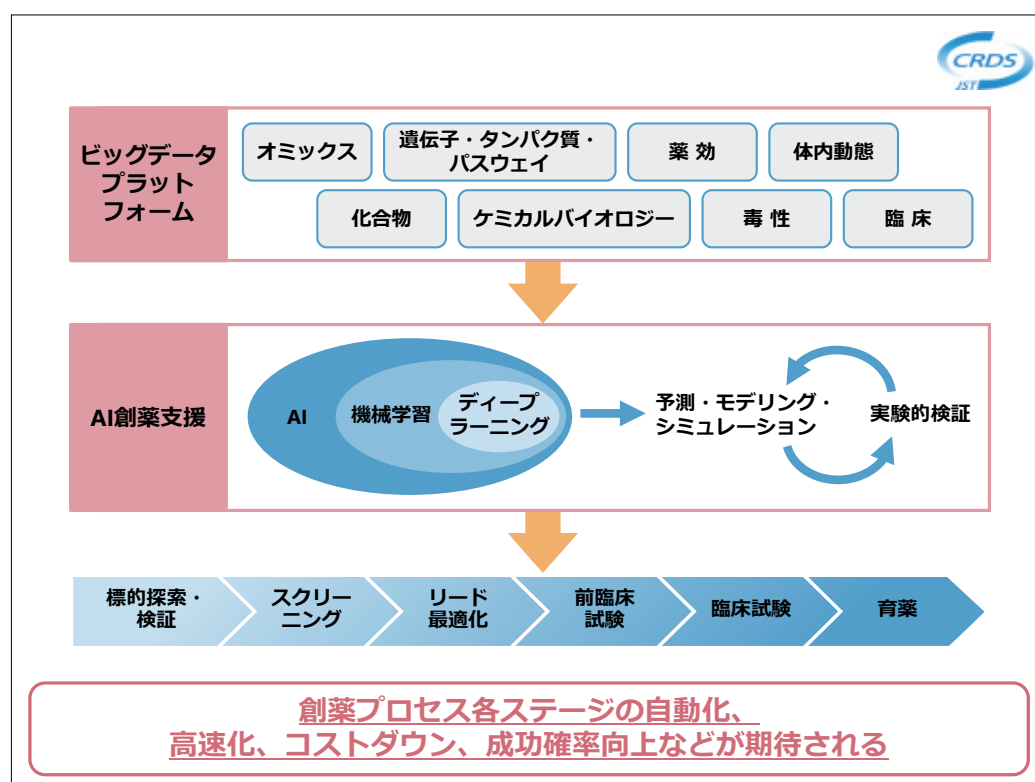


図 2-2 創薬プロセスの効率化にビッグデータ、AI の活用が望まれている

(2) 研究のトレンド

創薬の成功確率は 3 万分の 1 ほどとされ、開発には 1000 億円以上の費用と 10 年以上の期間がかかると言われている。医療健康ビッグデータの急速な拡張により医薬品の研究開発は標的分子主導型からデータ主導型に転換しつつあり、生命現象の精緻な理解と予測に基づく生産性の向上が期待されている。また、Precision Medicine を実現しようとする趨勢から、予防や患者への適切な医療の提供

が要求されており、個々の患者の疾患と関わりのあるデータ（遺伝子など）を創薬に利活用することが検討されている。

創薬のプロセスは、標的分子の探索からリード化合物探索、最適化を経て、臨床段階へと至る多岐に渡る専門領域が連結した工程であり、創薬の現場ではこの工程に沿って試行錯誤やフィードバックを繰り返している。医薬品の候補となる化合物を探索、絞り込む際の要素技術は、化合物の収集・管理、アッセイ系の構築・実施、評価データの利用である。コンピューターを用いてシミュレーションと AI（ビッグデータ解析・機械学習）を活用し、プロセスの自動・高速化、コストダウン、および大量情報処理による知識の抽出を経て創薬を効率化する潮流が生じている。

・ 創薬ターゲットの同定

創薬ターゲット探索は、疾患を引き起こす原因に関連する遺伝子を特定することから始まる。1980 年代から急速に発展した遺伝子配列解析技術は、創薬の可能性を大きく広げた。網羅的に遺伝子情報が得られるようになると、ゲノム中の 1 塩基多型情報 SNP（Single Nucleotide Polymorphism）と疾患の関連を調べることが可能になり、GWAS（Genome Wide Association Study）など疾患関連遺伝子の探索技術が発展してきた。例えば、高コレステロール血症を持つ家系の遺伝子解析で PCSK9（Proprotein convertasesubtilisin/kexin type 9）が原因遺伝子として同定され、PCSK9 を標的とする家族性高コレステロール血症治療薬アリロクマブ、エボロクマブ（PCSK9 阻害薬）の創出につながった。創薬ターゲットを明らかにすることは、医薬品を開発していく上での第一ステップであり、特にヒトゲノム配列が解明されて以降、各製薬企業は創薬ターゲット探索に巨額の投資を行ってきた。Roche 社は 2018 年にがん関連遺伝子解析プラットフォームを有する Foundation Medicine 社を 24 億ドルで子会社化した。2019 年 9 月には英 UK Biobank の 50 万人を対象とした Whole-Genome Sequencing Project に Amgen 社、AstraZeneca 社、GlaxoSmithKline 社、Johnson & Johnson 社が計 1 億ポンドを出資して、優先的にデータにアクセスする権利を得ている。

わが国でも「創薬標的の枯渇」問題を克服するための取り組みとして、官民研究開発投資拡大プログラム（PRISM）の枠組みで「新薬創出を加速する人工知能の開発」を目指したプロジェクトが 2018 年より開始されている。特発性肺線維症と肺がんにおける新規創薬ターゲット探索を目的に、臨床情報と文献など既存知識のデータベースを構築し、それを活用して創薬ターゲットを同定する AI の研究開発が行われている。

・ リード化合物の発見

リード化合物探索は、創薬ターゲットを特定したのち、想定される疾患メカニズムをもとに *in vitro* スクリーニング系を構築し、膨大な化合物ライブラリーを用いて目的の機能を持つ化合物のスクリーニングで行われる。しかし、膨大な化合物ライブラリーを無制限に評価することは事実上困難であるため、コンピューター上でリード化合物を予測するためのアルゴリズム開発も進んできた。創薬の現場では精度の高いリード化合物探索が必要とされ、スパコンを用いた高精度のシミュレーションを元にしたバーチャルスクリーニング技術が発展してきている。現在、バーチャルスクリーニングのソフトウェアは、有償、無償のものを含め非常に多くのものが存在し、創薬現場でも特別な技術ではなくなっている。バーチャルスクリーニングには、創薬のスピードアップ、精度の高い化合物デザイン、*de novo* ドラッグデザインなどの成果が期待されている。

バーチャルスクリーニングは、技術的な観点から、ケモインフォマティクス（化学情報学）に基づく方法と計算化学に基づく方法の 2 系統に大きく分類される。既知のリガンドや活性化合物の情報を有する場合は Ligand-based drug design（LBDD）の枠組みで定式化できるケモインフォマティクスに基づく方法が適用可能であり、標的タンパク質の立体構造情報を有する場合は Structure-based drug design（SBDD）の枠組みで定式化できる計算化学に基づく方法が適用できる。リガンド情報が

未知のタンパク質に対しても解析可能であるが、立体構造が分かっているタンパク質にしか適用できないという弱点もある。スーパーコンピュータ「京」(2007 年～)、ポスト「京」(「富岳」)(2014 年～)を用いた戦略分野の一つに創薬応用を掲げており、「京」の利用に向けて GENESIS、Modylas など様々なソフトウェアが開発され、利用が始まっている。

・ リード化合物の最適化

リード化合物をもとに化学構造変換と in vitro/in vivo アッセイによる活性評価を繰り返すことこのプロセスにおいても、分子シミュレーションによる化合物デザイン (SBDD: Structure-Based Drug Design) が大きな役割を果たしてきた。特に、キナーゼのような複数種類のタンパク質で活性中心の構造が似ている場合に、ターゲットとなるタンパク質に特異的な構造を設計する上で重要な示唆が得られる。リトナビル (HIV プロテアーゼ阻害薬)、イマチニブ (Bcl-Abl キナーゼ阻害薬)、ゲフィチニブ (EGFR 阻害薬)、オセルタミビル (ノイラミニダーゼ阻害薬) などは、SBDD により見いだされた薬の代表格である。

計算化学に基づく結合親和性予測法は、量子化学的手法と分子動力学計算的手法の2つのアプローチで開発されている。SBDD による化合物最適化プロセスは、前述のクライオ電子顕微鏡を用いたタンパク質構造の精密化により、精度の高い化合物デザインが可能になっている。

・ 前臨床薬効試験

非臨床薬効試験では、疾患モデル動物による化合物の効果が評価される。このとき、薬物動態 (PK: Pharmacokinetics) と薬効 (PD: Pharmacodynamics) の関連を解析することで、最適な化合物選択と臨床での薬効予測を行い、臨床試験に進む化合物の選択が行われる。薬物動態の評価では、薬効発現に必要な体内分布になっているか、適切に代謝・排泄されているかなど、薬効発現に必要な化合物の選択が行われる。非臨床モデル動物での実測値だけでなく、薬物動態に影響を与える化合物の特性 (化合物のタンパク質結合率、膜透過性、薬物トランスポーターとの親和性、血液脳関門透過性、薬物代謝酵素の阻害・誘導活性など) を測定し、これらのパラメータから非臨床・臨床の体内動態のシミュレーションを行う技術も研究されており、候補化合物の選択に貢献している。2015 年から医薬基盤・健康・栄養研究所が中心となり開始されたインシリコ創薬支援プロジェクトでは、薬物動態、安全性の予測システム開発が進められている。

(3) 国内外のトピックス

・ Merck Molecular Activity Challenge

2012 年に製薬大手 Merck 社が実施した、約 15 万の化合物とその活性を AI に学習させ、Merck がすでに実験済みの化合物について化合物活性値を予測するコンテストである。様々な機械学習の手法が比較、検討され、複雑で大量のデータを扱う創薬プロセスにおいて、AI の活用が機能し得るという認識を広める大きなきっかけとなった。コンテストで 1 位となったのはディープラーニングの手法を用いたチームだったが、2 位のサポートベクターマシンとは僅差であった。ディープラーニングは常に万能ではなく、特徴データの形式、学習データの質・量、解くべき問題に大きく依存すると思われる。生命医科学や創薬科学の関連データは欠損値が大部分で、バイアスが大きく、サンプル数が少ない傾向にあり、特有の難しさが障壁となる。

・ 産学連携の取り組み

▶ Life Intelligence Consortium (LINC)

わが国では、2016 年に製薬・化学・食品・医療・ヘルスケア関連のライフサイエンス分野のための AI・ビッグデータ技術開発を pre-competitive 領域で開発し、産業競争力を高めるコンソーシアムとして LINC (Life Intelligence Consortium) が京都大学 奥野恭史教授を代表として設立された。アカデミア、製薬・ライフサイエンス企業、IT 企業など 100 以上の機関が参画し、10 組のワーキング

グループがシームレスな AI 創薬プラットフォームの構築を目指した技術開発を進めている。

▶ Machine Learning for Pharmaceutical Discovery and Synthesis Consortium

マサチューセッツ工科大学が中心となり、医薬品発見と合成のための機械学習コンソーシアム (Machine Learning for Pharmaceutical Discovery and Synthesis Consortium) が 2018 年に立ち上がり、Pfizer、Novartis、Lilly など大手製薬会社も参画している。同コンソーシアムは、低分子の発見と合成に役立つソフトウェアのデザイン促進を目的としている。

▶ Machine Learning Ledger Orchestration for Drug Discovery (MELLODDY) project

欧州の製薬企業 10 社とアカデミア、IT 企業が参画して 2019 年から開始された MELLODDY (Machine Learning Ledger Orchestration for Drug Discovery) project は、機械学習による有望な化合物予測プラットフォームの構築を目指している。各企業はデータの秘匿性を保持しながら 1000 万を超える低分子化合物のデータを共同利用できる枠組みを構築し、世界最大のコレクションを活用した、より正確な予測モデルの確立、創薬の効率化が期待されている。

・ AI 創薬ソリューション

米国 IBM 社が開発した膨大な文献データの俯瞰から仮説を提示する Watson for drug discovery が先行している。p53 のリン酸化酵素や筋萎縮性側索硬化症 (ALS) に関係する RNA binding protein が新たに同定されたことが報告されており、創薬ターゲット探索に貢献することが期待されている。

また、AI 創薬スタートアップである Insilico Medicine 社は、DDR1 を標的とする fibrosis などへの治療薬候補となる化合物を、ディープラーニングシステムを利用することによりわずか 21 日間で設計することに成功したことを 2019 年に報告している¹。同社が開発した低分子のデノボ設計を可能とするディープラーニングモデル GENenerative Tensorial Reinforcement Learning (GENTRL) を使った分子設計により、従来 2 ～ 3 年を要すると言われるヒットからリード化合物を創出する過程を 2 ヶ月に短縮し、コストの低減も期待される成果として注目を集めた。

AI を活用して創薬に何らかのソリューションを提供しようとするスタートアップ企業は、すでに 100 社以上が誕生している。国内製薬企業も含めてビッグファーマと AI 創薬企業との提携、買収が盛んに行われている。例えば、AI を利用して自動化された医薬品化合物設計などを提供する英 Exscientia 社は、2012 年設立以来急速に発展し、GlaxoSmithKline 社、Roche 社、Sanofi 社、Bayer 社などと共同研究を行っている。大日本住友製薬社との共同研究では、AI を活用して平均 4 年半とされる探索研究を 12 ヶ月未満で終え、見出された化合物の第 1 相試験を開始したことが 2020 年に発表された。国内では、DeNA 社が旭化成ファーマ社、塩野義製薬社などと AI 創薬の検証を行っており、NEC が AI を活用した創薬事業に本格参入することも 2019 年に発表されている。

ビッグファーマは創薬に留まらず AI 技術をプラットフォームとして獲得してデジタルトランスフォーメーションを遂げる姿勢を強めており、Sanofi 社と Google 社が提携して Virtual Innovation Lab を設立することや Novartis 社が Microsoft 社を戦略的パートナーとして AI イノベーション研究所を設立することが 2019 年に相次いで発表されている。

(4) 課題と方向性

医療データを創薬へ活かすための大きな課題に、医療現場のデータが整備されていないことが挙げられる。各病院の電子カルテデータはフォーマットが統一されておらず、自然文で記述されることが多いため、標準フォーマットや オントロジーの整備が課題となっている。解決するために、共通のテンプレートを用意する方法、あるいは、自然言語処理技術を応用する方法の大きく 2 つのアプローチが試みられている。

1 Nat Biotechnol. 2019 Sep;37(9):1038-1040. doi: 10.1038/s41587-019-0224-x.

また、AI 開発にあたり、学習データセットの選択とその構築方法は大きな課題である。AI は一般的に学習データセットの規模が大きくなるほど性能が高くなる傾向にあるが、製薬会社が個々に独自のデータセットを確立することは非効率的であり、国内でいくつも似たようなデータセットが同時多発的に構築されることは望ましくない。さらに、AI は発展途上の技術であるため合目的的な学習データセット構築方法が確立されておらず、コンソーシアム的な取り組みでも利害関係者が多すぎる場合、かじ取りが困難である。

医薬品候補化合物取得の方法には、予め標的を定めないフェノタイプベースのハイスループットスクリーニング (PHTS) がある。目的とするフェノタイプに合致した化合物をヒット化合物として得るための重要な方法である。しかし、PHTS で見出された化合物は標的分子や作用機序が未知で、標的分子の同定に大きな労力を必要とするため、標的未知化合物の標的予測法が切望されている。化合物の化学構造、人体に対する副作用パターン、化合物応答遺伝子発現情報などを基に機械学習の技術を駆使した標的分子予測法が提案されてきているが、リガンド情報が未知なタンパク質に対する予測 (学習データが無い場合の予測) は困難なのが現状である。

創薬インフォマティクス研究は、医学薬学の高度な知識と計算科学の専門性が要求されるため、専門人材が産業界で不足している。また、学際領域の教育システムが定着しにくいわが国においては、総合的な創薬の教育体制を整備している大学は皆無に等しい。そもそも日本の大学において創薬インフォマティクスを専門にする教員のポストはほとんど無く、人材開発は大きな課題である。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 ライフサイエンス・臨床医学分野 (2019 年)」(CRDS-FY2018-FR-04) (2019 年 3 月)

2.3 バイオ生産

(1) 定義・意義

「バイオ生産」とは、微生物・細胞、植物、動物などの生物を用い、低価値な資源（糖、無機塩類、光、二酸化炭素、飼料など）から、育種・生産プロセス（培養、栽培、飼育養殖）を通じ、高付加価値の物質や生物体自体を目的産物（食料、燃料、化成品、素材、医薬品、生物的工具など）として生産することを指す。

バイオ生産は、遺伝子組み換えなどの手法により既存の生体分子や生命システムを改変する育種によって進められてきたが、近年は生体内の各種分子、細胞を網羅的に解析することにより得られた知見を活用し、より人工的な設計指針、構築プロセスを用いて要素構成的に生物を理解し、制御しようとする取組み、すなわち「合成生物学」的なアプローチが台頭してきている。この合成生物学的なアプローチは、これまでの「仮説検証型」のアプローチに対し、計測・分析技術によって得られたデータを定量化し活用することで大量の生命情報から法則を発見するという観点で、「データ駆動型」の新しいアプローチである。

(2) 研究のトレンド

バイオ生産に対するアプローチ変化の背景には、今世紀に入り、次世代シーケンサー、各種オミクス、クライオ電子顕微鏡などの分析・解析手法の登場に加え、IoT 技術などの情報技術も飛躍的に発展し、膨大な生物由来データが産出されるようになったことがある。得られた膨大なデータを AI で解析し、その結果をフィードバックしたモデルを構築するために必要な DNA 合成コストも低下した。加えて、CRISPR/Cas9 に代表されるゲノム編集技術が開発されたこともあり、自由自在に目的の DNA 配列を設計、構築できる時代は目前になってきた。このように、目的の DNA 配列を設計、構築し、目的の細胞に導入、評価を行い、その結果を再び設計にフィードバックするサイクルは Design-Build-Test-Learn の頭文字をとって DBTL サイクルと呼ばれる（図 2-3）。この DBTL サイクルの推進により合成生物学は加速度的に進展しており、微生物・細胞を用いたバイオ生産にはこの合成生物学的なアプローチによる育種が欠かせないものとなっている。

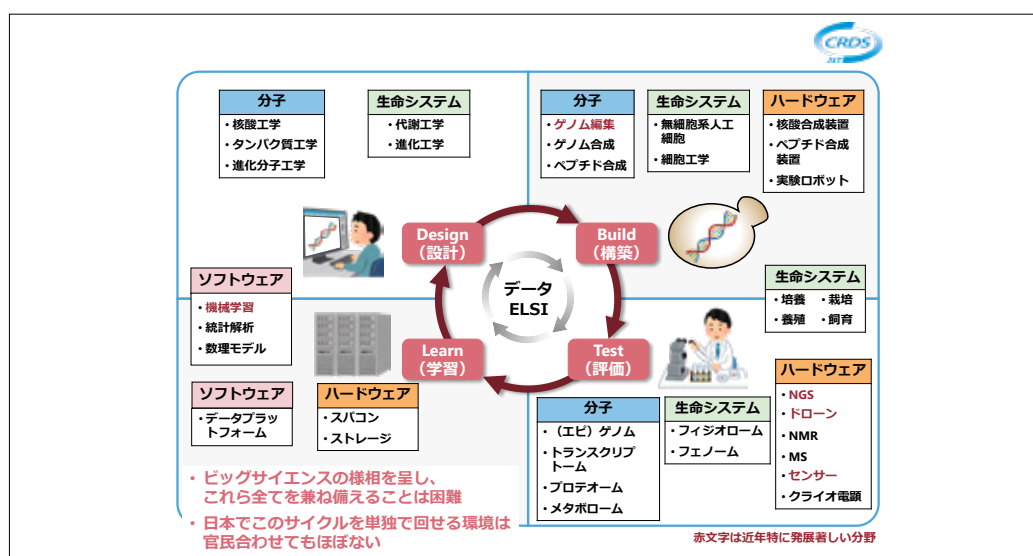


図 2-3 DBTL サイクルと合成生物学関連技術俯瞰

バイオ生産における DBTL サイクルにおいて、Design 領域では AI・シミュレーションを活用した分子・システム設計が行われている。Build・Test 領域ではルーチンワークの省力化、各種オミクスなどの解析データ収集の効率化・自動化や、産出されるデータの高品質化を目指し、数千万ドル規模の巨額の資金を投じて実験ロボットや IoT による機器制御が導入されている。Learn 領域では得られたデータ解析・機械学習結果をもとに、さらなる設計精度の向上に繋げる試みが行われている。DBTL の個々の技術（特にタンパク質の立体構造解析など）に関してはハイスループット化、高精度化に未だ検討の余地があるものの、全体としては Build・Test 領域は成熟段階にあると認識されている。そのため Design 領域研究へのビッグデータ活用に向けて、実験ロボットや IoT を積極導入、ルーチンワークの合理化及び省力化と、高品質データ収集の自動化が欧米諸国で先行して行われた。近年ではデータ取得そのものというよりは、得られたデータから Learn、Design 領域でいかに効率よく学習し、少ないサンプル、試行数で目的の機能を有する生体分子、代謝経路を創出できるか、というビッグデータ解析にシフトしており、ディープラーニング技術の応用等を含めた解析技術開発が進められている。

このような生物的生産プロセスは、化学的プロセスと比較してグルコース等の再生可能資源を用いて複雑な化合物を高い選択性、特異性で生産可能といった特長を有する。また一般的に常温、常圧下での反応であり、有機溶媒や重金属の使用量も少ないことから、エネルギー消費、環境負荷の少ない生産方法であり、SDGs、バイオエコノミー（石油資源代替）の観点からも世界的に関心が高まっている。そして、上記のような技術革新により、微生物学においては、メタゲノム解析によりこれまで培養できなかった難培養微生物に関する研究が大きく進展し、分類、進化に関して重要な知見が多数得られている。分離培養による研究でも新たな代謝経路を有する微生物が発見される等、物質循環や産業応用の点でも重要な発見が近年相次いでいる。そして、従来は偶発的な変異等に頼っていた育種が、人為的な設計に基づいて行われるようになり、さらには微生物・細胞への新規機能の付与、創出が目指されるに至っている。モデル微生物でのゲノム編集技術確立に加えて、産業用微生物・細胞を用いた高機能物質生産、微細藻類の脂質生産量の向上によるバイオ燃料生産など応用分野を志向した研究開発が進められている。

一方で、現在の技術では、生体分子・生命システムの設計精度が不足しており、生体内で目的の機能を発揮するかどうか、実際に生体内に導入して検証してみないとわからないなど、DBTL サイクルを用いた手法にはコスト面での効率が悪いという問題がある。そのためターゲットが限定され、現状は検討にかかるコストを回収する見込みが立ちそうなもの、大規模市場が予想できるものに限定されている。例えば、医薬品原料として高い需要のあるテルペノイド、フラボノイド、アルカロイドなどの高付加価値化合物などであり、オピオイド系アルカロイド生産では酵母、大腸菌双方で成功例が見られる。

また、手本、原型となるものが天然に存在しない生体分子・生命システムの *de novo* 設計は依然として困難である。特に *de novo* 設計は既存の分子・システムの改変の場合とは比較にならない膨大な組み合わせから、求める機能を発揮できるものを選択する必要がある。そのためにはさらに高精度の設計指針、そしてそれを導き出すための体系的なデータ産出体制、基盤が求められることから、今後はこのような観点での研究開発が重要である。

(3) 国内外のトピックス

DBTL サイクルを活用した微生物育種の成功例としては、米国のバイオベンチャー、Amyris による抗マラリア薬、アルテミシニンの生産プロセス構築が挙げられる。組換え酵母を用いて前駆体アルテミシニン酸を生産し、その後化学プロセスを用いてアルテミシニン酸をアルテミシニンへと変換する工業プロセスが確立され、合成生物学を活用した育種の成功事例として世界的に認識されている。

DBTL サイクルを支える個別の技術としては、米国ベンター研究所による合成ゲノムマイコプラズマの作製とゲノム移植、その後のミニマム細菌の作製、合成ゲノム酵母 Yeast2.0 などを発端として、長鎖 DNA 合成や、そのハンドリング（精製、細胞への導入など）技術、人工細胞系構築技術の開発が進んでいる。加えて慶応大学のグループが、光合成細菌ゲノムを丸ごと枯草菌に入れたシアノバチルスの作製を報告しており、国内からの研究成果も顕著である。

本分野では、前述の Amyris や、Ginkgo Bioworks、Zymergen をはじめとして、バイオベンチャーが特に米国に多数設立されており、高効率なバイオ生産を目指し、Build・Test 分野のハイスループット化と自動化に向けた IoT 技術の導入、ロボット化によるビッグデータ構築に関する技術開発を行うなど著しい進展がみられる。

海外で進行中のプログラムとしては、DARPA「Living Foundries」(2011 年～)や NSF「Understanding the Rules of Life (URoL)」(2018 年～)が挙げられる。DARPA のプログラムは、DBTL サイクルを迅速に回すための基盤技術（設計ツール開発、標準シャーシ株作製、評価）を開発し、ゲノム編集、機械学習などの技術も取り込み、化成品・医薬品など有用な 1000 の分子の生物生産法の開発を目指している。一方、NSF のプログラムは、遺伝子型から表現型の予測など、数学物理、コンピューターサイエンスを活用し、普遍的な生命のルール of 生命の解明を目的として研究が展開されており、NSF が掲げる、「未来に向けて投資すべき 10 大アイデア」の一つとしても取り上げられている。また、国際的な取り組みとしては、ボトムアップで酵母の染色体を化学合成 DNA で置き換えていくことを目的とした米国、中国、英国が参画する国際プロジェクト「Sc2.0」や、ヒトゲノム全体のデザインとアセンブルを目的とし、10 年以内にゲノム合成にかかるコストを 1/1000 にすることを掲げる国際コンソーシアム「GP-write」がある。

国内では、NEDO において、「植物等の生物を用いた高機能品生産技術の開発」(2016 年度 -2020 年度)が実施されている。このプログラムでは、遺伝子設計に必要となる精緻で大規模な生物情報を高速に取得するシステム、細胞内プロセスの設計、ゲノム編集等の技術等を利用して、植物・微生物等の生物細胞内に物質生産プロセスを内含する「スマートセル」を構築し、省エネルギー・低コストな高機能品生産技術の確立を目指している。また、内閣府の戦略的イノベーション創造プログラム (SIP) 第 2 期で実施中の「スマートバイオ産業・農業基盤技術」(2018 年度開始)においては研究項目の 1 つとして、データ駆動型の機能製品設計技術による「生物機能を活用したものづくり」が実施されており、前述の NEDO プロジェクトとも連携しつつ、日本における Bioeconomy の拡大を目指している。

(4) 課題と方向性

Design・Built 領域の課題は機能性分子、生命システムの設計指針に乏しいことである。成功例もほぼ全てが既存の分子、システムの流用、改変の域を脱していない。酵素などの機能性分子をゼロから設計したり、長鎖 DNA を安定的に機能発現させるには、十分な知識・データの情報処理、機械学習などの援用が必須であり、機能性分子や生命システムの設計に向けた体系的なビッグデータ構築が課題である。Test 領域では、現況は既存技術の自動化、ハイスループット化といった点に注力しているが、計測技術のライフサイクルを考慮すると、次世代の計測技術（非破壊計測やマイクロ流体等）を活用した少量サンプルで高感度に計測可能な技術の開発が望まれる。また IoT 技術とビッグデータ構築技術など情報処理技術との継続的な連携が必要である。Learn 領域では、公共データベースをもとに設計、学習を実施しているが、予測・推定をもとにした設計を行うには十分な体系的データが供給されていない。解決には合成生物学に特化したデータが必要だが、既存のデータベースは一般生物情報・データの域を出ず、改めて文献等からの知識抽出や計測データの補完が必要である。また、近年ディープラーニングの活用が顕著だが、あらゆる問題に適用可能という訳ではなく、現状は見極め

の段階であり、平行して次世代アルゴリズム開発も必要である。スマートセル設計システムの場合でも、代謝設計・モデル構築・遺伝子配列選択、設計・遺伝子制御関連解析など多様な問題があり、様々な統計解析・機械学習技術との併用や新しいアルゴリズム開発が求められている。DBTL サイクルの各領域に共通した問題として、各場面で先端技術を活用する一方で、技術のライフサイクルもまた早いことから、適切な問題設定とデータ・計測技術・解析技術・情報処理技術を適時選択することも望まれる。

このように合成生物学は非常に学際的な領域であり、その性質上ライフサイエンス全般に関する知見の統合、体系化に向けた多様な人材の参画や様々な実験、解析設備を要する。機材や設備などハード面からみると、個々の技術自体は成熟段階にあるものが多いが、実際に組み合わせて利用すると精度不足、技術同士がかみ合わないといったことが生じるため、要素技術のブラッシュアップ、すり合わせ、標準化を意識した技術開発が重要である。そのためには少量のサンプルでの検討を可能にする微細加工技術や、工程の高速、高精度、自動化に向けた制御工学の分野との連携が必要不可欠である。また、ソフト面では、得られた高品質なデータを蓄積し、次の手を模索するデータ駆動型科学を進める情報科学分野の人材の参画が不可欠である。効率的な研究にはこれらを集約、高度に連携するための技術開発、体制づくりが喫緊の課題となっている。特にわが国においては、以前より Build・Test 分野の個別の技術開発が注力され、DBTL サイクルを単独で機能できる拠点到に乏しい。また、バイオインフォマティクス人材が不足しており、情報処理技術の活用、という点でも後れを取っている。

バイオ生産では、生命システムの改変、創造が挙げられるため、常に ELSI、デュアルユースといった問題に向き合う必要がある。科学技術の視点のみならず、社会に与えるインパクトを考慮しながら研究開発を進めることが求められている。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター 「戦略プロポーザル 次世代育種・生物生産基盤の創成（第1部）～核酸、タンパク質、細胞を結ぶ、多階層横断的サイエンス推進による生体分子・生命システム設計ルール創出～」(CRDS-FY2018-SP-07) (2019年3月)
- ・ 科学技術振興機構 研究開発戦略センター 「研究開発の俯瞰報告書 ライフサイエンス・臨床医学分野（2019年）」(CRDS-FY2018-FR-04) (2019年3月)

2.4 物質・材料

(1) 定義・意義

新物質・材料の設計や創製に、データ分析や機械学習の手法を取り入れた「マテリアルズ・インフォマティクス（以下 MI）」に世界中から大きな注目と期待が寄せられている。背景には、物質・材料科学における理論・実験の進展に加え、コンピューターの演算能力の向上、計測・分析技術の進展、シミュレーション手法の発展など物質・材料科学の進歩とともに、ディープラーニングに代表される機械学習手法の急激な進展がある。MI は狭義には新物質・材料の予測・設計が目的とされるが、今後は新物質・材料の発見加速だけではなく、さらに、合成・評価の実験回数の低減、合成・製造プロセスの最適化（プロセス・インフォマティクス）への期待や、実験と測定評価の工程そのものをつないで自動化するロボティクスとの融合（オートノマス・サイエンス）がその先の将来として描かれ、世界の産学から注目が集まっている。

MI は、計算科学による物性予測とそれを実証するハイスループット材料合成・評価、それらのデータを統合管理する材料データベースや機械学習などを統合的に活用した物質・材料探索・設計の取組全般を指す研究開発領域である。実験、計算で得られた物質・材料に関する知識とデータを駆使して、統計的手法により物質・材料の機能を制御する規則を探り、それを通して具体的に新物質・新材料の発見を加速する。究極的には、規則の背景にある材料特性を支配する法則を発見し、自在な材料設計を可能とする系統的アプローチの構築を目指す取り組みである。物質・材料の状態や機能を記述する新概念や新形式などの開発や、物性や材料特性の新しい理解など、物質・材料科学にパラダイムシフトを起こすものとしても期待され、実際に物質・材料開発における基盤技術として定着しつつある。

研究開発の現場にデジタルトランスフォーメーションがもたらす変容

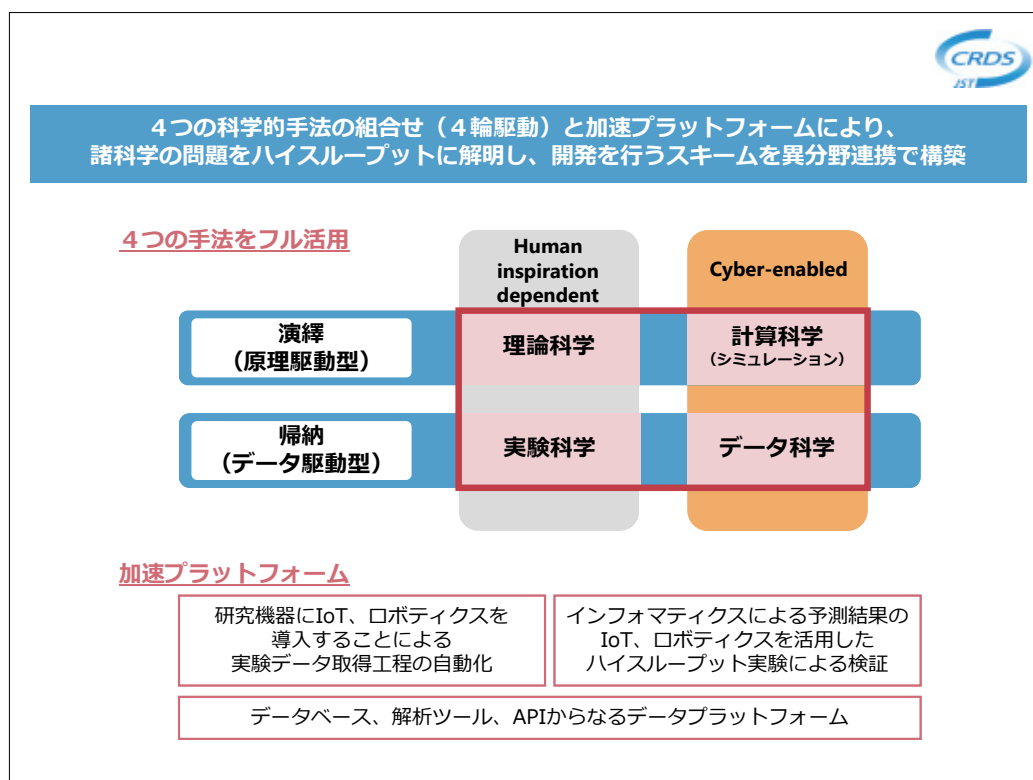


図 2-4 データ駆動型研究開発の概念図

(2) 研究のトレンド

計算科学に基づいたデータ駆動型物質・材料開発 (MI) の概念は、2000 年頃から提唱されていたが、当時は計算機で取り扱える原子数や精度が不十分であった。また、大量データを統合し材料科学の問題をマイニングする手法も未開拓であった。しかし、近年の計算機能力の著しい向上と大量データを取り扱える環境が整ったことで、状況は劇的に変わりつつある。2011 年に米国が「Materials Genome Initiative (MGI)」を発表したことをきっかけに、計算科学とデータ科学を融合させたデータ駆動型材料設計が注目され、世界中で精力的な研究がなされている。かつて、理論、実験に次ぐ「第 3 の科学」として計算科学が物質・材料研究における解析や予測の手段として重要な位置を占めるようになったことと同様に、「第 4 の科学」としてデータ科学の重要性が指摘されている。

わが国の第 5 期科学技術基本計画 (2016 年 1 月 22 日閣議決定) においては、超スマート社会「Society 5.0」の実現に貢献する 11 のシステムの 1 つである「統合型材料開発システム」として表現され、「計算科学・データ科学を駆使した革新的な機能性材料、構造材料等の創製を進めるとともに、その開発期間の大幅な短縮を期待する」と記載されている。将来的には開発期間・コストの短縮・削減だけではなく、人工知能が新規な物質を予測・予言したり、材料化のレシピを示したり、新たな物理法則を発見することができるのではないかと期待されている。

以下では、MI の注目される研究開発動向として、「新物質探索」「ミクロな組織構造と材料物性の相関」「結晶構造から材料組成までのマルチスケール統合」の 3 分類について、国内の研究グループによる最近の研究成果を事例に挙げる。

〔新物質探索〕

第一原理計算とデータ科学 (ベイズ最適化、ディープラーニング等) を併用することで効率的に候補物質をスクリーニングし、新規物質探索を行う手法が主流である。ターゲットは電池材料、半導体材料、触媒材料など様々である。

- ・ 京都大学の研究グループは、リチウムイオン電池の正極材料のサイクル寿命を改善する目的で、数千種類の固溶元素の組み合わせを考慮して網羅的に高精度な第一原理計算を実施した結果、従来のリチウムイオン電池の 6 倍以上の寿命をもつ新材料開発に成功した (2014 年)。
- ・ 東京工業大学の研究グループは、既知および仮想的な物質を含む約 600 種類の候補物質リストを作成し、基礎物性 (バンドギャップの大きさ、有効質量の大きさ等) や格子振動および競合相に対する結晶安定性などを条件に第一原理計算を用いたスクリーニングを行うことで、赤色発光デバイスや太陽電池への応用が期待できる新規窒化物半導体の合成に成功した (2016 年)。
- ・ 東京大学の研究グループは、光の吸収波長に着目し、これまで困難であった AI による有機分子の自動設計を可能とするためにディープラーニングと分子シミュレーション技術を組み合わせることで、望み通りの特性を示す有機分子の選別およびその合成に成功し、機能性分子設計に有用であることを実証した (2018 年)。

〔ミクロな微細構造と材料物性との相関〕

合成プロセスなどの違いにより、同じ組成をもつバルク物質であっても、無数の種類の界面や異なる微細構造が存在することで、電気伝導率やイオン伝導性、熱輸送特性などが大きく影響を受けることが知られている。そのため、ナノ構造や界面構造などのミクロな微細構造とバルクの材料特性との相関を解明することが必要となるが、物質から材料にする際には様々な階層構造が存在するため、それらの相関を人間の能力のみで解明するのは限界があり、データ科学を用いることが有効である。

- ・ 東京大学の研究グループは、資源探索分野におけるビッグデータ活用法であるクリギングと呼

ばれる空間補間法を用いて、物質の特性に決定的な役割を果たす界面構造を高速に決定することに成功した。これによって、これまで数千～数万回という膨大な数の構造緩和計算が必要であった物質界面の最適化を、わずか 100 回程度の計算で可能とした(2016 年)。また翌年には、転移学習を用いることで物質の界面構造を決定するための計算コストを約 1/3600 まで削減することにも成功している。

- ・ 熱マネジメントなどの応用を念頭に、材料内部のナノスケールの構造を作製して熱輸送を制御する技術が注目されている中、東京大学の研究グループは、ナノ構造の熱輸送を計算する手法とベイズ最適化手法を組み合わせることで、従来の数パーセントの計算量で熱抵抗を最大または最小にする最適なナノ構造を特定する計算手法を開発し、非直感的な新規ナノ構造を設計することに成功した(2017 年)。翌年には、ベイズ最適化を用いた機械学習によって、膨大な候補構造から熱電変換効率が最大になるナノ構造を高い最適化効率で決定する手法開発にも成功している。

〔結晶構造から材料組成までのマルチスケール統合〕

鉄鋼やゴムなどの構造材料のさらなる高性能化・高機能化には、原子・ナノスケールの微細構造と組織構造やマクロ特性、製造プロセスとの関係を明らかにすることが重要となっている。その場合の入力はナノ・メソ・マクロのそれぞれのスケールにおける組織情報(多くの場合は画像)であり、画像認識タスク(特徴抽出)との相性がよいディープラーニングによる MI が威力を発揮すると期待されている。

- ・ 名古屋大学の研究グループは、ディープラーニングを用いて材料組織画像を元に構造材料の微細構造を定量化し、その結果を用いて特性を推定するだけでなく、所望の特性を発現する材料組織を逆解析する材料情報統合システム(Materials Integration Phase Analyzer: MIPHA)の開発に成功している。これによって、複雑系を対象とした場合でも効率よく組織→特性(順問題)、特性→組織(逆問題)の解析を行うことが可能となる。
- ・ タイヤの性能にはゴム材料となるポリマー(ゴム)とフィラー(カーボンブラックやシリカなどの微粒子)の複雑な微細構造が設計因子となるが、横浜ゴムは自社で開発済みのシミュレーション技術と AI 技術を用いることで、求める性能を実現するために重要となる微細構造の設計因子とその閾値を客観的かつ定量的に短時間で導出することに成功し、MI によるゴム材料開発技術を確立したことを発表している(2017 年)。

(3) 国内外のトピックス

米国が 2016 年 8 月に公表した MGI の総括文書によると、DOE や DoD、NSF、NIST、NASA 等の機関に 5 年間で総額 500 百万ドルを超える国費を投入した結果、得られた主な成果として「材料開発におけるパラダイムシフトの醸成」「NSF による異分野横断型基盤的ファンドの充実」「次世代人材の育成」「データベースの構築」「コンソーシアムの設立」が挙げられ、ホワイトハウスで「米国製造業に革新をもたらす可能性があるため、政権交代後も引き続き支援すべき」とされている。

しかしながら、トランプ政権下では MGI は優先事項には含まれておらず、目下のところ、CNGMD、ChiMaD、PRISMS、SUNCAT の 4 つのセンターを中心に各省の予算の中で自主的に研究開発が継続されている。このような状況に対して、NIST は調査会社 RTI に依頼し、「Economic Analysis of National Needs for Technology Infrastructure to Support the Materials Genome Initiative (MGI を支持する技術基盤に対する国家的ニーズの経済分析)」というレポートを 2018 年 4 月に発行している。ここでは、米国材料業界の 100 名を超える専門家へのアンケート結果に基づいて、「実験、計算、ソフトウェアコードなどの高品質データへの容易なアクセス」「材料データ共有などの協働ネットワークの構築」「計算科学的アプローチを含む材料設計手法の開発」「試行錯誤から脱却し、高い費用対効

果を産み出す生産とスケールアップ手法の開発」「材料欠陥制御や製品ばらつき予測を可能とする品質保証制御とコンポーネント基準認証方法の確立」「モデル検証と不確定性の定量化」の6つの重要ニーズの特定、およびそれらのニーズを満足することで得られる経済的インパクトの見積りとして「年間経済利益 1230 億～2700 億ドル」「新材料開発に伴うリスクが約 50%削減」「新材料開発期間が約 35%短縮」を挙げている。

わが国では、情報統合型物質・材料開発イニシアティブ(MI²I)の取り組みが注目される。MI²IはJST「国立研究開発法人を中核としたイノベーションハブの構築支援事業」の2015年度の採択課題であり、NIMSを中核拠点として、最大5年度、4.5億円/年の規模で実施されている。ここでは、データ科学と物質・材料科学の融合によって新物質・材料科学研究を加速する取り組みの場として、Matnaviを中心とするデータベースの開発・整備、物質・材料から情報・数理科学にわたる産学官の協働作業の体制の構築、より広範な企業の参画を促進し、オープンイノベーションに繋がるハブ拠点化を目指している。具体的課題としては、画期的な蓄電材料、磁性材料、伝熱制御材料を開発し社会実装へ繋げると同時に、新しい物質・材料科学手法の開発・蓄積を進め情報統合型物質・材料開発手法のパッケージ化・システム化を目的としている。2016年4月からMI²Iコンソーシアムが設立され、法人会員、個人会員の受付を随時行っている(2016年7月末時点で参加企業31社)。コンソーシアム会員には、各会員が関心を持つテーマに対して情報統合型の研究手法を確立する取り組みを行う一方で、国家的課題である情報統合型物質・材料開発の研究手法を早期に実現するため、オープンイノベーションを体現する場の創成に積極的に関わることが求められている。

なお、NIMSには2017年4月に統合型材料開発・情報基盤部門(MaDIS)が発足し、MI²Iを推進する情報統合型物質・材料研究拠点とSIP「革新的構造材料」マテリアルズインテグレーションを推進するSIP-MIラボに加えて、新たに材料データプラットフォームセンターが設置された。材料データプラットフォームセンターが構築と運用を担う物質・材料データプラットフォームは、MI²Iと密接に連携し、またNIMSが形成を目指している研究基盤MRB(マテリアルズ・リサーチバンク)の柱のひとつになっている。データベース整備に関しては、世界最大級の物質・材料データベース「MatNavi」を有するものの、データ解析を前提として作られていなかったため、データ駆動型物質・材料開発に適したデータベースに再構築する必要があり、MI²Iの中にデータプラットフォーム委員会を設けて、今後のグランドデザインについて検討を進めている。

(4) 課題と方向性

科学技術的な課題としては、データ駆動型物質・材料開発を進める上で主要な科学技術的課題として物質から材料に至るまでの特性や性能のデータに関する課題がある。

物質から材料へ至るにはナノ、ミクロ、マクロの各スケールで多くの構造や組織の因子が関係する。例えば、構造材料においては、ナノスケールでは転位を含む欠陥の密度や分布、溶質原子の配列や分布、粒界や界面の構造、ミクロスケールでは、結晶粒のサイズや形状、応力やひずみの分布、マクロスケールでは欠陥や溶質偏析、残留応力など極めて多様な因子が相互に関係している。それらの因子を如何にデータベース化するか、またこれらの因子は製造や加工のプロセス条件によって変化するものであるため、そのようなメタデータをどのように付与するのか、さらに各因子の空間分布や時間変化などの情報をどのように付与するか等、課題は山積している。さらに、問題とする材料特性ごとに、その特性を司るクリティカルな因子はそれぞれ異なるのが普通であり、どの程度まで網羅的なデータベースを構築すればよいか、理想的なデータベースが構築できたとしてもデータから元の物質・材料が再現可能であるのか等についても検討が必要である。

また、MIでは候補物質は提示してくれても、合成方法までは提示してはくれない。「(2) 研究トレンド」で記載したように、いくつかの新規物質、新規材料はMIの力を得て短期間で合成することに

成功しているものの、現状の技術レベルでは物質合成に関しては依然として研究者の経験と勘によるところが大きい。したがって、今後は、候補物質を安定的に作製可能な合成プロセス条件まで示唆してくれるインフォマティクス、すなわち、「プロセス・インフォマティクス」を確立することが必要になると考えられる。

MI およびプロセス・インフォマティクスの研究開発の推進にあたり、やはり最大の課題は材料データやプロセスデータをいかに収集・共有するかであろう。材料は全ての産業の基盤であり、そのデータはアカデミア・産業界を問わず広く共有されるべきである一方、合成プロセスと材料組織構造、材料特性との関係を示すプロセスデータは産業界にとってはそれ自体がノウハウであり競争力の源泉でもある。そもそも材料分野は競争的性格の強い分野でもあるため、他者とのデータシェアリングが歓迎されない文化が醸成されている。そのことによって、データの相互利用を礎とするデータ駆動型物質・材料開発を推進する上で、材料分野における競争的性格が最も大きな障壁になっていると言えるため、それを乗り越えるオープン＆クローズド戦略をどう構築するかが重要である。

これらの課題への回答の一つの方向性として、最近、ロボティクス技術による自動実験の仕組みと解釈可能な機械学習（explained AI）を組み合わせた「解釈可能な MI」技術開発による新材料開発の効率化に成功したり（2019 年、NEC・東北大学・メリーランド大学の共同研究成果）、新たに開発したハイスループット触媒評価装置から得られた触媒ビッグデータを用いた触媒反応プロセスの超高効率化に成功（30 年の研究が実働 1 ヶ月未満で実施できることを実証）している（2019 年、JAIST・北海道大学・熊本大学の共同研究成果）。これはいわば、実験科学の自動化（オートノマス・サイエンス）へと向かう事例であり、オートノマス・サイエンスの取り組みは、海外では英国、米国、カナダが先行している。

さらに、データ提供のインセンティブを促進することも重要であり、例えばデータを提供することが研究者の次なる活動へのモチベーションや評価等につながっていくような、新たな仕組み・研究評価システムを構築する必要がある。同時にこのような物質・材料研究に関わるデジタルトランスフォーメーションを牽引する、産学それぞれにおけるリーダー層の出現や若手の活躍が最も重要である。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター「戦略プロポーザル データ科学との連携・融合による新世代物質・材料設計研究の促進（マテリアルズ・インフォマティクス）」（CRDS-FY2013-SP-01）（2013 年 8 月）。
- ・ 科学技術振興機構 研究開発戦略センター「ワークショップ報告書 データを活用した設計型物質・材料研究（マテリアルズ・インフォマティクス）」（CRDS-FY2013-WR-03）（2013 年 8 月）。
- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 ナノテクノロジー・材料分野（2019 年）」（CRDS-FY2018-FR-03）（2019 年 3 月）。
- ・ Materials Genome Initiative
<https://www.mgi.gov/>（2020 年 3 月 1 日アクセス）
- ・ 宮下哲、島津博基「わが国におけるマテリアルズ・インフォマティクスの進展状況」『工業材料』2019 年 2 月

2.5 ものづくり

(1) 定義・意義

ものづくりのデジタル化を牽引する中核技術として近年注目されている“デジタルツイン”は、サイバーフィジカルシステムの一つの形態であり、デジタルデータを基に物理的な製品をサイバー空間上で仮想的に複製し、将来発生する事象をデジタルの仮想世界で予測することが可能な先進的なシミュレーション技術である(図1)。

従来、現実世界の環境を複製するには、現実世界の情報を人が入出力していたため、サイバー空間に入力されるデータの量はある程度限定され、現実世界をそのまま再現することは不可能であった。しかしながら、IoTの普及によりリアルタイムのデータ自動取得が可能となった結果、ある一定の条件のもとでは、現実世界を仮想世界に再現することができるようになり、デジタルツインが着目されるようになった。

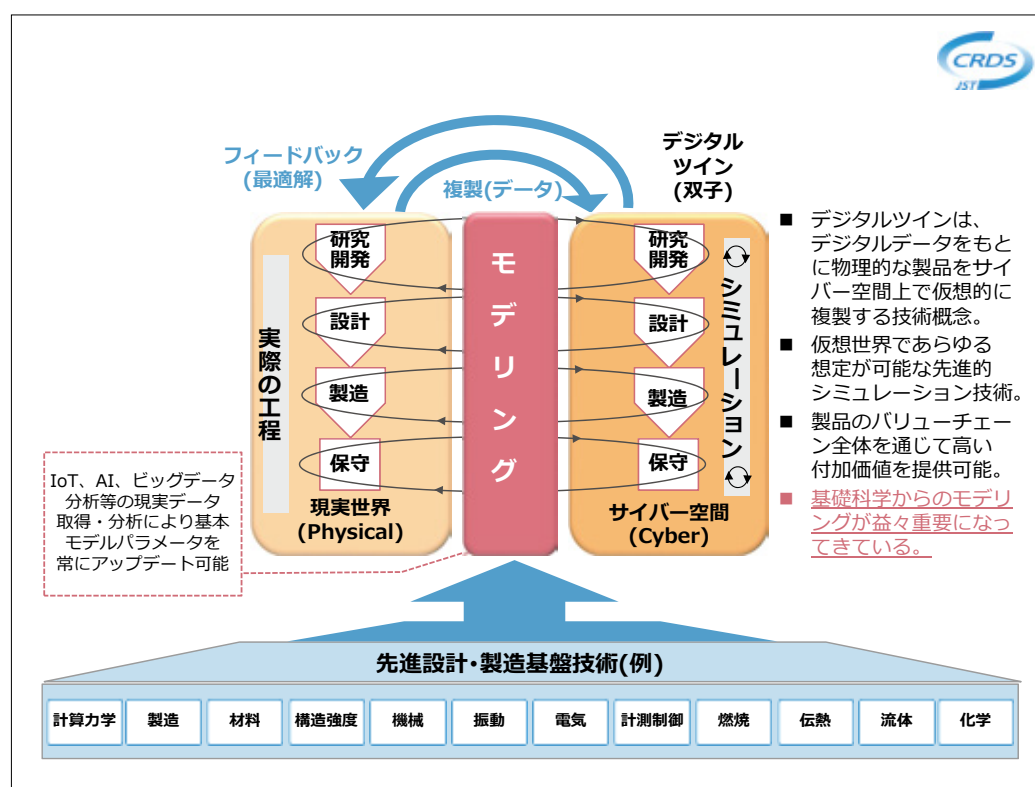


図 2-5-1 デジタルツインと、分野横断の複合現象モデリングを支える多様な基盤技術

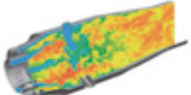
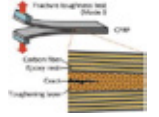
デジタルツインの環境を利用することで、製品・サービスのバリューチェーン全体を通じて高い付加価値が提供されると期待されている。環境・エネルギーや輸送に関する機器・サービスの開発・設計・製造や保守(寿命予測)、これらの生産性や品質向上に役立ち、Society 5.0や低炭素社会の実現に貢献する。例えば、風車、ガスタービン、蒸気タービン、自動車、工作機械、船舶海洋等の環境・エネルギー機器、輸送機器の開発・設計・製造や保守(寿命予測)を行う際に必要となるデジタルツインが出口として想定される。

(2) 研究のトレンド

デジタルツインの根幹は、製造、材料、構造・強度、機械、燃焼、伝熱、流体、振動、化学、電気などの技術分野を横断する複合現象モデル（例えば機械損傷プロセスモデル等）の開発、検証で、その開発と活用が注目されている。しかしながら、モデリングの基礎となる破壊などの物理・化学現象は十分に解明されていないため、求められる先進的な環境エネルギー機器・サービスの創出ができていない。これら現象の把握・理解に基づいた多様な基盤技術の統合化による、基礎的原理説明からのモデル開発・検証および計算効率化技術開発、基礎科学研究からの知識基盤構築とその活用により、低炭素社会に資する先進的機器・サービスの開発が期待される。

上述のようにデジタルツインの鍵は、材料、構造・強度、機械、燃焼、流体、振動等様々な分野を包含する未踏複合現象の解明とそのモデル化で、この差別化が競争力の源泉となる。表1は、モデリングの基礎となる、十分に解明されていない分野横断の未踏複合現象の例を示す。これらの現象解明とモデリングを進めるには、例えば、(i) 機械要素の損傷・亀裂進展メカニズムや、(ii) 急激な過渡現象に対応する燃焼素反応と乱流、伝熱の相互作用メカニズムなどの未踏複合現象モデル化において、(i) では構造・強度、材料の専門性だけでなく、製造、機械、振動、流体などの他技術分野の専門性と、(ii) では燃焼の専門性だけでなく、伝熱、流体、振動、化学、電気などとの他技術分野の専門性との融合・横断が必要となる。デジタルツインを用いた次世代設計・製造技術を構築するには、ものづくりの基盤技術を横断し、知を融合させる工学系専門人材が求められる。

表 2-5-1 モデリングの基礎となる分野横断の未踏複合現象

技術分野	未解明の物理・化学現象(例)	参考図
機械(機械要素・トライボロジー)／構造強度	・ 転がり軸受けなど機械要素の損傷メカニズムにおける亀裂進展プロセス ・ 潤滑膜の流れと構造体の運動、表面変形などの相互作用メカニズム 等	
燃焼／流体	・ 急激な過渡現象に対応するガス燃焼制御、空力制御に必要な素反応と乱流の相互作用メカニズム 等	
流体／構造／振動	・ 回転機械や輸送機器における流体、構造、音響の相互作用メカニズム ・ 複雑な流体関連振動現象における自励振動メカニズム 等	
構造強度／材料／流体	・ 流体中における炭素繊維複合材料(CFRP)の亀裂進展プロセス 等	

図出典：ジェイテクト、重工技報、龍谷大学理工学部、JST

(3) 国内外のトピックス

近年、デジタルツインやそれを支える先進設計・製造基盤技術において、時代の変化に対応した新しい視点での設計・製造・加工技術などのイノベーションの開発が国内外で進んでいる。ドイツでは、

Industrie4.0に代表されるように、シミュレーション基盤技術などのイノベーション振興施策が実施されている。EUは、Horizon2020の下、2016年「Smart Cyber Physical Systems」として、約22億円を製造の自動化に向けたサイバーフィジカルシステム技術(CPS、デジタルツインはCPSの一つの形態)への取り組みに投資した。英国では高付加価値製造カタパルトが既に始まっており、2017年3月には、デジタルカタパルトとイノベートUKが、中小企業に対し、The Digital Engineering and Test Centreとの共同研究によるデジタルツインを活用した先進製造技術テーマの公募を開始した。

米国では、2011年以降、国家先進製造戦略計画(National Strategic Plan for Advanced Manufacturing)や先進製造パートナーシップ(AMP: Advanced Manufacturing Partnership)などで、先進製造のためのデータ基盤を整備してきており、製造イノベーションの育成と商業化の促進に取り組む産学官パートナーシップManufacturing USAなども推進している。また、米国科学財団NSFが、2008年以来、計約300億円をCPS技術に対して投資してきている。一例としてERC(Engineering Research Centers)プログラムでは、産学連携プロジェクトにて、基礎科学の知見獲得からの基盤技術に基づいてデジタルツインの社会実装まで着実に実施しているものもある。さらに、General Electric(GE)などの企業で、デジタルツイン技術の導入事例が出てきている。

中国は、第14次五か年計画(2021～2025年)を視野に入れ、製造業発展のための施策として、「中国製造2025」を発表した。本施策では、製造工程のスマート化(「スマートマニュファクチャリング」および「環境配慮型製造工程」)やイノベーション強化に関する5つのモデルプロジェクトが選定されている。2019年9月には、北京の北航大学教授がデジタルツイン研究開発促進には、①データとモデルの標準化②データとモデルの共有化③ソフトウェアサービスの革新④異分野連携フォーラムの確立の4つが必要とNatureで発表している。特に、正確なデジタルツインを構築するには多様な専門家の緊密なチームが不可欠であること、今後すべてのモデルを実行できるデジタルツイン用のユニバーサルデザインおよび開発プラットフォームも開発する必要があることなどを述べている。

業界、データサイエンティスト、サイバーセキュリティの専門家、エンジニアリングおよびビジネスストラテジストを結び付ける異分野連携フォーラムの先駆けとして、中国では2017年、デジタルツインに関するソーシャルメディアグループがソーシャルメディアプラットフォームで設立され、英国ではキール大学のキールビジネススクールのスマートイノベーションハブが、米国ではビジネスコンサルタントのブーズ・アレン・ハミルトン運営のイノベーションハブがワシントンDCで設立されている。

(4) 課題と方向性

わが国が、今後も、ものづくり国家として他国には真似のできない信頼性の高い製品・サービスづくりの強みを継続していくためには、設計・製造基盤技術の深耕と伝承が欠かせない。しかしながら、エネルギー基盤技術(工学)俯瞰ワークショップ(2016/11/25、26開催、参考文献参照)によると、わが国の産業界では、製品、要素技術の性能品質は総じて世界有数だが、ビジネスモデルが弱く、プラットフォーム戦略による優位性の確保などができていない。ものづくりの基盤技術を担う工学系専門人材が減少しつつあり、諸外国と比較すると相対的に個々の研究力が低下していることが懸念されている。

ものづくり産業の開発・設計・製造現場で用いられる解析技術に関して、個別の要素解析技術で世界レベルにある分野がある一方、システム化や統合化は国際的に後塵を拝し、モデリング、シミュレーションの基礎となる物理・化学現象の把握・理解は十分にできていない。評価方法の開発・標準化、評価データの蓄積も立ち後れている。全体解析技術[ソフトウェア、モデル構成式、検証データ取得]でも欧米に大幅なリードを許している。このため、環境・エネルギー機器や輸送機器・サービスの開

発・製造・保守に際して、品質、納期、コスト面で産業競争力が発揮できていない。また、大学でも、基盤技術分野の研究力が中国などに比べ弱化してきているため、日本企業が、オープンイノベーションで、上記のモデリングや検証データ取得などを欧米や中国の大学に委託する例が散見され、わが国の大学の基礎研究がますます弱くなると同時に日本企業にも研究基盤が蓄積されないという悪循環となっている。

このように、ものづくり産業を支える工学系分野の研究基盤弱体化の懸念が強まっており、産業界並びにアカデミアが直面する上記問題の解決には産学官が連携して分野横断で一体となって取り組む必要がある。

ものづくりのデジタルツイン構築に必要な取り組みは多岐にわたるが、複合現象モデルの開発、検証の重要性を鑑み、以下の3つの具体方策への取り組みが期待される。

① 多様な基盤技術の統合化・分野融合による、基礎的理理解明からの複合現象モデルの開発・検証、および、モデル計算効率化技術開発、構築モデルの評価方法の開発・標準化

② ①に資する基礎科学研究からの知識基盤構築 [物理・化学現象の把握・理解、基盤となる構成方程式の確立、評価データ取得、蓄積等]

③ ②を活用した、ものづくり産業を支える工学系の基礎的理理解を深めた人材、科学的合理性に基づき、ものづくりができる人材の育成

対象とする製品分野・領域は、ものづくり技術の難易度が高く、開発期間・コストが大きい環境・エネルギー・輸送に関する機器・サービス領域を選択するのが有効である(図2-5-2)。これらは、デジタルツイン活用の効果が大きく、かつ今後の国際市場獲得や低炭素社会実現に向けた日本の貢献にも寄与すると期待される。

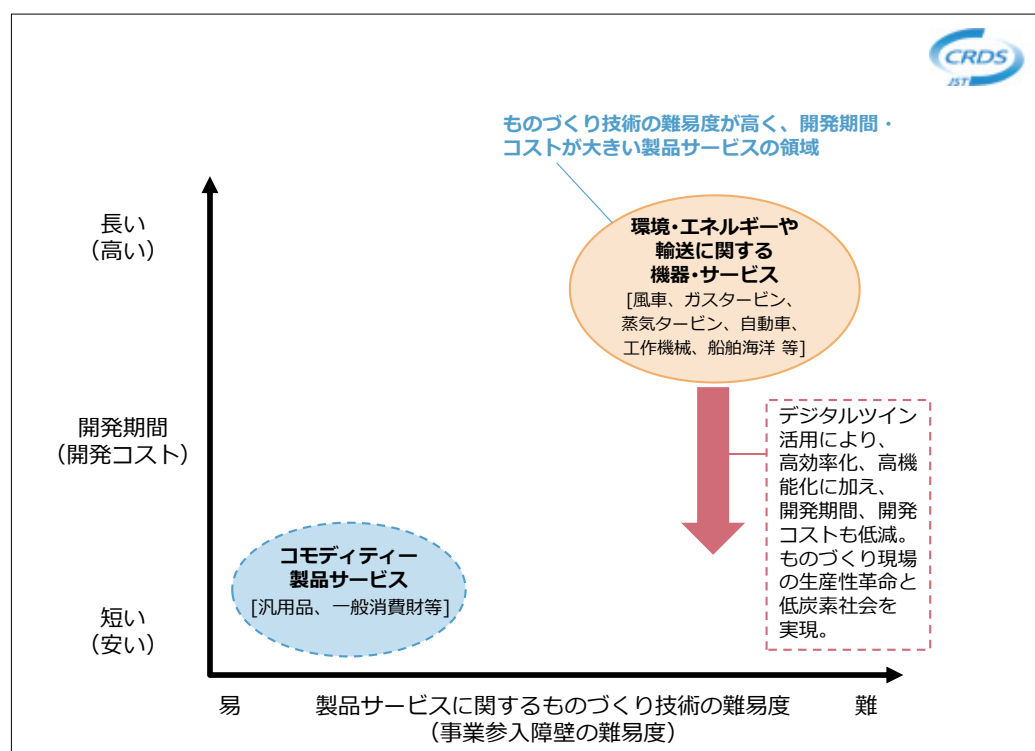


図2-5-2 デジタルツインに関する研究開発課題で対象とする製品分野・領域

上記の①から③を一体的に推進するための研究開発課題構成を表 2-5-2 に示す。これらは異なる製品間で共通的なボトルネックと考えられる研究開発課題群である。これらに係る物理・化学的メカニズムの解明、モデル構成式の確立、複合現象モデルの開発と実験的検証、モデル評価方法の開発・標準化、およびそれらに基づく力学基盤を中心とするデジタルツイン技術確立を目指す。当然、共通課題への取組みの後には個別の上記製品例へ展開するための研究開発も課題となる。製品共通課題の主要項目を以下に示す。

a 機械損傷プロセスモデル構築と検証 [複合材料も含む機械要素の損傷メカニズムや亀裂進展プロセスの解明等]、b トライボロジーモデル構築と検証 [回転機械の摺動面潤滑膜流れと表面変形などとの相互作用メカニズムの解明等]、c 流体・構造・振動連成モデル構築と検証 [回転機械や輸送機械における流体、構造、音響の相互作用メカニズムなどの解明]、d 流体・伝熱・燃焼・化学連成モデル構築と検証 [急激な過渡現象対応の燃焼素反応と乱流、伝熱・熱伝達の相互作用メカニズム解明等]、e a～d 統合化モデル(デジタルツイン) 構築と検証 [ターゲット製品例への展開と、機械要素の運動、構造変形、振動、潤滑、流体、伝熱、燃焼、化学、電磁気など複合現象支配方程式の高速連成解析モデルの開発と検証等]

表 2-5-2 具体的な研究開発課題



研究開発課題	ターゲット製品(例)					
	風車	ガスタービン	蒸気タービン	自動車	工作機械	船舶海洋
a 機械損傷プロセスモデル構築と検証	○	○	○	○	○	○
b トライボロジーモデル構築と検証	○	○	○	○	○	○
c 流体・構造・振動連成モデル構築と検証	○	○	○	○	○	○
d 流体・伝熱・燃焼・化学連成モデル構築と検証		○	○	○	○	○
e a～d統合化モデル(デジタルツイン)構築と検証	○	○	○	○	○	○

さらに、上記の①から③の取組み項目を効果的に推進し、デジタルツイン技術とそれを支える先進設計・製造基盤技術の研究開発で世界を凌駕していくための体制として、ネットワーク型の産学官連携推進体制(コンソーシアム)を中核とした体制構築を適切に進める必要がある。ここでは基礎研究から社会実装までを俯瞰した一貫通貫の研究開発が求められる。加えて、データ連携共通プラットフォーム創出、国際標準化戦略などの出口戦略や人材育成などにも着実に取り組むことが肝要である。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター 「戦略プロポーザル 革新的デジタルツイン ～ものづくりの未来を担う複合現象モデリングとその先進設計・製造基盤技術確立～」(CRDS-FY2017-SP-01) (2018 年 3 月)
- ・ 科学技術振興機構 研究開発戦略センター 「科学技術未来戦略ワークショップ報告書 革新的デジタルツイン ～デジタルツインを支える先進設計・製造基盤技術～」(CRDS-FY2017-WR-09) (2018 年 3 月)
- ・ 科学技術振興機構 研究開発戦略センター 「俯瞰ワークショップ報告書 エネルギー基盤技術(工学)」(CRDS-FY2017-WR-04) (2017 年 9 月)
- ・ 中山 力、木崎 健太郎、吉田 勝「製造業革新を加速するデジタルツイン」『日経ものづくり』2017 年 8 月
- ・ Fei Tao, Qinglin Qi, "Make more digital twins," *Nature* 573, 490-491 (2019)
- ・ 大平竜也「科学技術の潮流 デジタルツイン モノづくり技術革新」『日刊工業新聞』2019 年 9 月 13 日

3 フィールド・社会に関する研究開発とデジタルトランスフォーメーション

3.1 医療

(1) 定義・意義

今後克服すべき疾患の多くは、ゲノムと多種多様な要因が長年に亘って複雑に絡み合い、発症・重症化し、重篤な発作が起こり死に至る。そのような経緯を辿る疾患は、個々の分子メカニズムの知見の集積だけでは克服不可能であり、現実世界での疾患のリアルな記録である、健康・医療データ基盤を活用したアプローチも併せた研究開発の実践が有効である。それにより、対象群の適切な規模での個別化・層別化と各対象群における最適な予防・治療介入法、すなわち“個別予防・予見医療”技術が確立する。“個別予防・予見医療”は、人々の健康長寿に最も大きく貢献する。

近年、健康・医療データの大規模な収集・解析によって疾患・医学“知”の構造化が大きく進展し、新規医療技術の開発、診断支援システムの開発、医療行政の意思決定支援など、多方面に大きなインパクトをもたらしつつある。健康・医療データの整備・活用が、健康・医療産業の競争力や医療保障制度の最適化に直結するとの認識が、もはや世界共通のものとなっており、国内外で基盤整備が急速に進められている。

(2) 研究のトレンド

健康・医療データの収集～利活用の流れを、データ形式（定型データ／非定型データ）の観点で整理した俯瞰図を図 3-1-1 に示す。ここで述べる定型データとは、データ収集形式が規格化され、定量的な数値データや定型テキスト形式で記録されたもので、代表的なものとしてはレセプトが挙げられる。レセプトは、必要な匿名化処理を行った上で、生データ或いは若干の前処理を行うだけで大規模データ解析が可能である。非定型データとは、データ収集形式が規格化されていない、フリーテキスト、音声、動画などの記録データであり、そのままでは意味のある解析は不可能である。現状、図左側の定型データに関する取り組みは国内外で活発に成されているが、図右側の非定型データに関する取り組みは殆ど見られない。

非定型データを解析するためには、定型データへ変換せねばならず、多くの場合、人海戦術でのデータ処理作業、およびその作業を軽減するための半自動化システムが必要となる。非定型データの種類は多様であるが、その中には高品質な情報が内包されているものも多く存在する。例えば、わが国において、症例報告（年間数万～数十万件作成）、診療ガイドライン、退院時サマリ、医師などの所見、診断結果や検査依頼の根拠、問診、検査結果、添付文書、薬局文書、医学論文など、医療現場で価値のある重要なデータの多くはフリーテキストで記録されている。それらのうち、一部は SS-MIX2 拡張ストレージや学会データベースなどに電子データで保管されているものの、非定型データであるため利活用が困難で、新たな価値を生み出すに至っていない。それら未利用の貴重な非定型データ群のうち、データの質・量および解析難易度を踏まえ、特に大きな価値を生み出しうるものから優先的に取り組みを開始することが重要である。

定型データは、政策上、あるいは臨床現場での何らかの必要性に対応して規格化されたデータである。価値の高い健康・医療データ基盤を構築するためには、定型データの範囲を拡大し、現在は非定型データと位置付けられる多様なデータ群を最初から定型データとして収集することが中長期的観点から重要である。往々にして新たな医療データ収集ルールの導入は、日々の臨床の激務に追われる医療従事者に負荷と混乱をもたらしてきた。それらを極限まで低減させ、一次フィードバック（現場のメリット）を短期的に強く実感できる仕組みとすることで、入力に伴うモチベーションの向上とバイ

アスが低減し、高品質なデータ収集が加速される。その結果、臨床現場にとってもメリットとなる診療支援システムの開発・実装が進み、さらに高品質なデータ収集が行われる。このようなポジティブなサイクルを成立させ加速させることがわが国においてきわめて重要である。

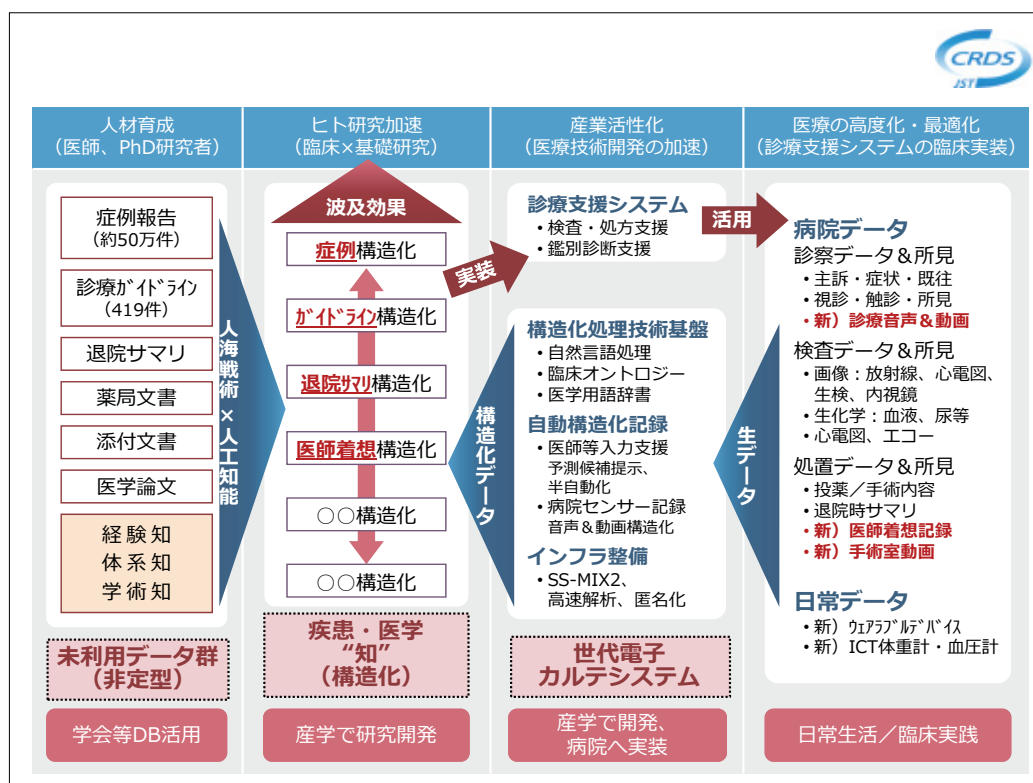


図 3-1-1 健康・医療データの俯瞰図

(3) 国内外のトピックス

海外動向

例えばエストニアは、早い時期から行政や医療の電子化を国家レベルで進め、先進的な健康・医療データ基盤を整備している。中国などの後発ながら成長著しい国家は、過去のしがらみにとらわれず優れたシステムを国内に一斉導入することが比較的容易であり、健康・医療データ基盤の整備が大きく進んでいる。

先進諸国では、センシティブな個人情報である健康・医療データを扱うことに対する、社会受容面での課題が非常に大きい。さらに、長年の伝統と共に確立した健康・医療システムを有するが故にしがらみも多く、健康・医療データ基盤の整備では遅れをとってきた。しかし、近年そのような状況は大きく改善しつつある。例えば、米国は過去 10 年間でカルテの電子化が急速に進み、現在では約 90%を超える導入率となった。電子カルテ導入で一定の成果がみられたが、当初喧伝されたほどの価値を創出するには至っていない事に対し、米国では不満が大きく広がっている。電子カルテに対する不満は、日本でも同様である。その背景として、現行の電子カルテは、臨床現場において紙媒体で記録されてきた様々な情報を電子化しただけの、いわば病院業務支援システムでしかない点にある。そのような思想で収集された定型データでは価値の創出は限定的である。より多くの価値の創出を前提とし、非定型データについても戦略的に定型データ化して収集する健康・医療データ基盤の構築が、世界的に大きな課題となっている。

なお、詳細は省略するが、ヒトゲノム基盤とこれら医療データを接続し、より価値の高い健康・医療データ基盤としようとする動きが各国で見られている。例えば米国の Precision Medicine、英国の Genomics England などが挙げられる。

わが国の政策的背景

わが国は、市場原理（米国型）でも政府主導（欧州型）でもない医療制度となっており、多様なステークホルダーが複雑に絡み合っていることから、健康・医療データ基盤の整備には大きな困難が伴う。また、わが国の医療行政の一環で収集される健康・医療データは、諸外国と比して質・量ともに優れているものの、諸外国に見られるような“健康・医療データは国家の財産であり積極的に活用せねばならない”との認識が弱く、健康・医療データに基づく新たな価値創出には及び腰な状況が続いてきた。しかし、ここ 3-4 年の間に、わが国においても健康・医療データ分野の活性化に向けた様々な取り組みがついに開始された。

平成 28 年 10 月、厚生労働省は、わが国の健康・医療データの収集・利活用を大きく加速させるため、“ICT を活用した「次世代型保健医療システム」の構築に向けて”を公開し、「つくる：集まるデータ→生み出すデータ」「つなげる：分散したデータ→データの統合」「ひらく：たこつば化→安全かつ開かれた利用」の 3 つのコンセプトと、それを実現するための施策群が示された。その一環で、例えば「次世代医療基盤法の施行（2018.5）」「全国保健医療情報ネットワーク（2020 ～）」「医療等分野における識別子（ID）（2020 ～）」の準備など、関連する法整備・環境整備が着々と進められている。また、健康・医療データ関連の研究開発プロジェクトもここ数年で次々と開始されている。ライフサイエンス・臨床医学分野全体を見渡しても、健康・医療データ領域は最も注目を集めており、産官学の取り組みは加速度を増している。

(4) 課題と方向性

現在、国内外で活発に解析されている健康・医療データは、レセプトなどの解析しやすい定型データ群である。しかし、中長期的に新たな価値を創出し続けるためには、現状の定型データ群だけでは不足すると考えられる。現時点では解析困難とされる非定型データも解析対象とするチャレンジが、今後特に重要となる。

そのためには、2 つの方向性が考えられる。1 つ目は、ある程度の人手と半自動処理技術を組合せ、非定型データを解析可能な定型データへと変換する。特に注目すべき点として、わが国には、極めて高品質な医療情報を含むものの、非定型データであるが故に死蔵されてきた健康・医療データ（症例報告など）が存在する。それらの解析・構造化は、世界に類を見ない新たな疾患・医学“知”の源泉となり、わが国の強みとなる。

2 つ目は、現時点では非定型データとして臨床現場で生成・収集されているものを、可能な限り定型データとして収集・蓄積するシステムの構築である。データ入力に伴う医師等の負担の最小化（医師の着想記録などの新たなタイプの情報収集も含む）とメリットの最大化（診療支援システムの提供など）を両立させる、「次世代電子カルテシステム」とも呼べる新たなシステムの開発と、臨床現場への実装が重要となる。

これらを戦略的に推進することで、わが国の健康・医療データ基盤が中長期的に大きく強化され、基礎医学、医療技術開発、医療提供、健康医療行政などの全てにおいて、計り知れないインパクトをもたらす。

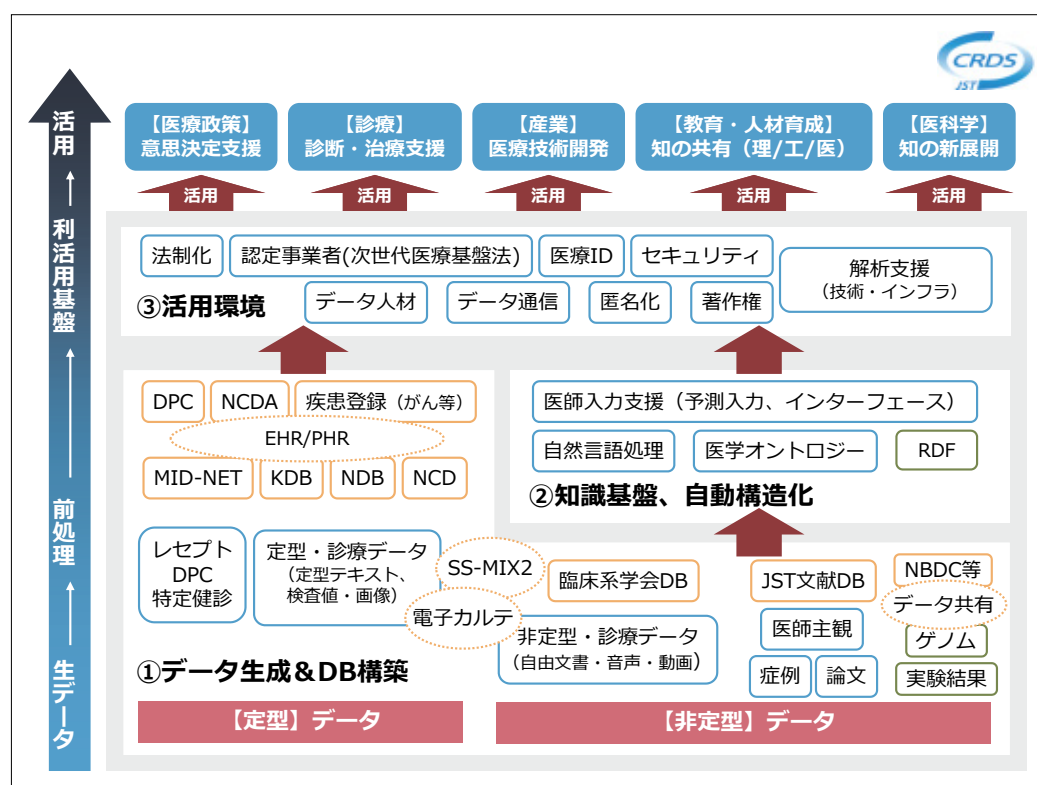


図 3-1-2 テーマ概要図

健康・医療データから人々のQOLの向上、産業活性化、サイエンスの新展開などの新たな価値を創出し続けるため、わが国において推進すべき課題は次の5つである。

- ① 『非定型データの自動／半自動構造化技術開発』・・・現状、解析が困難とされる非定型データを対象として、「非定型データ群の自動構造化技術開発（自然言語処理など）」「疾患用語集、疾患概念地図の作成」の取り組みを通じ、解析可能な範囲を拡大させる。
- ② 『未利用の非定型データ群の構造化推進、疾患・医学“知”の知識基盤の整備』・・・既にわが国に存在する、高品質な情報を含むものの非定型データであるが故に死蔵され続けてきたデータ群（例えば症例報告など）について、「既存の症例報告データの構造化」「既存の診療ガイドライン、退院時サマリ、医師などの所見、診断結果や検査依頼の根拠などの構造化」「継続的に構造化データ基盤がアップデートされる仕組みの整備」を実施することで、新たな価値を生み出し続けるデータ基盤とする。
- ③ 『新しいタイプの非定型データなどの取得、構造化の推進』・・・最先端の機器を活用し「病院“内”センサー記録の収集・構造化」「病院“外”センサー記録の収集・構造化」など、従来では収集対象として考えられてこなかったデータを収集し構造化する。
- ④ 『次世代電子カルテシステムの開発』・・・現行の電子カルテが新たな知を生み出す形となっていないため、「医師の記録時の負担の大幅軽減」「医師の診療支援」などのシステムを実装し、「従来は記録されてこなかった高品質な生データの収集」を行ない、得られたデータをもとに「付加価値の高い構造化データの自動・半自動生成、知識基盤の自動アップデート」を実施する。
- ⑤ 『次世代電子カルテシステムの社会実装、応用展開』・・・上記④のシステムはより多くの医療機関で実装されることで新たな価値創出が加速するため、「主要な大病院へのシステム実

装（既存の電子カルテシステムを拡張）」「“ヒト疾患研究統合イノベーション拠点（仮称）”へのシステム実装」を実施する。また、健康・医療データが国家繁栄の礎となる財産であるとの認識の上で、「健康・医療情報活用基本法（仮称）」の議論、整備を行なう。

以上の課題を実践し、わが国において疾患・医学“知”の構造化を行なった先に、“個別予防・予見医療”が実現すると考えられる。そして、関連する技術を制する国・企業がこれからの健康・医療産業を制すると言え、産業競争力強化の点からも大きな意味を持つ。

(5) 参考資料

- Vindell Washington et. al., “The HITECH Era and the Path Forward” *The New England Journal of Medicine*, September 7, 2017; 377:904-906, DOI: 10.1056/NEJMp1703370
- 厚生労働省「保健医療分野における ICT 活用推進懇談会 提言書」（2016 年 10 月）
- 科学技術振興機構 研究開発戦略センター「戦略プロポーザル データ統合・ヒト生命医科学の推進戦略（IoBMT）」（CRDS-FY2018-SP-06）（2019 年 3 月）

3.2 ヘルスケア

(1) 定義・意義

高齢化の進捗そして医療費の増大に伴い日本の医療が大きな転換期を迎えており、従来の治療中心の医療から予防医療・予見医療へのシフトが喫緊の課題である。その中で、新たな手法を用いた早期診断や日常生活（医療機関外）におけるモニタリングの重要性が高まってきており、ゲノム診断、リキッド・ガスバイオプシー、ウェアラブルデバイスといった診断・モニタリング手法、さらに AI による進化が期待される手法として画像診断が注目を集めている。

これらの手法が普及することにより、消費者や社会に以下のような波及効果をもたらしうると考えられる。

- ・ 遠隔、医療機関外での診断の選択肢が広がる。
- ・ 自身の健康状態をより正確に把握でき、健康に対する意識の向上につながる。
- ・ 診断の効率化・質の均一化が図れ、医者や医療費が限られる中で診断の質を維持できる。
- ・ さまざまな計測モダリティで得られたデータを統合することで、新たな指標を用いた超早期診断や各個人に合わせた個別化診断が可能となる。

本稿では主に、ヘルスケア IoT のキーとなるウェアラブルデバイスと AI 技術の適用が進む画像診断に焦点を当てる。

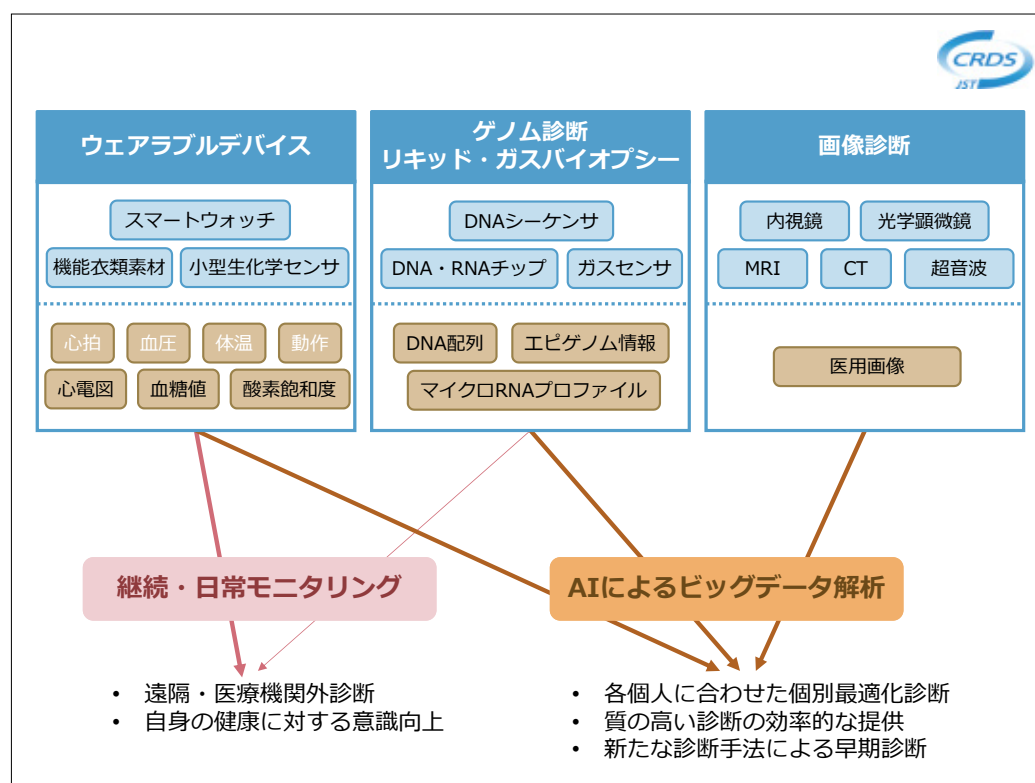


図 3-2 ヘルスケアにおける、新しい診断・モニタリング手法および AI 活用による効果

(2) 研究のトレンド

①ウェアラブルデバイス

スマートフォンの次世代デバイスとして、Apple Watchのような民生のウェアラブルデバイスが普及しつつある。その中で、心拍や心電、呼吸等のバイタルデータをモニタリングするデバイスも出てきており、Fitbitを始めとしたスポーツトレーニングやフィットネスなどの領域で活用されてきている。ヘルスケア・医療分野での活用がさらに進むには、小型化・長時間駆動といったデバイスに対する一般的な技術要求に加え、医学的な有用性やそれに伴う経済的・社会的便益が創出されることが不可欠である。ウェアラブルデバイスに医学的な有用性を持たせる方向性としては、大きく以下の2つがある。

1) 医学的に有用性の高いパラメータを測定できるセンサーを搭載したデバイスの開発

ウェアラブルデバイスに搭載されるセンサーは、何を感じ取るかによって、機械的センサー（加速度計、ジャイロセンサー、圧センサーなど）、生理学的センサー（心拍、血圧、体温モニターなど）、生化学的センサー（持続血糖モニター、アルコールセンサーなど）に分けられる。心拍数、動作（歩数、活動量モニター）、呼吸数、体温など、日常生活でのバイタルを評価するセンサーは比較的実装が容易である一方、あるパラメータ単独から医学的に意味のある結果を見出すことは難しい。より医学的な有用性の高いパラメータとしてECG、酸素飽和度、生体電位などの応用も始まっているが、今後はバイオ/生化学情報のリアルタイム計測への展開が期待されている。例えば、Abbott社の血糖値モニタリング用バイオデバイスである「Free Style リブレ」は、生化学測定である血糖値評価が継続的に行えるバイオデバイスとして2017年にFDAの承認を得ており、従来の自己測定器より簡単で痛みも少ないということで糖尿病患者の間で急速に普及しつつある。

2) ビッグデータ解析による、有用性の高い情報の抽出

IoTをヘルスケア分野に適用した概念として、通信機能を内蔵したデバイスと医療情報ICTシステムをオンラインのコンピューターネットワークを通じて繋ぐというIoMT（Internet of Medical Things）があるが、通信機能を内蔵したウェアラブル機器はネットワークへデータを送り出すセンサー端末として機能する。個々のデータとしては医学的な有用性が低いパラメータであっても、多数のデバイスから継続的な測定により得られる大量のデータを解析することで、各個人に最適化したモニタリング・治療や特定の疾患・発作に特有の前兆の感知に活用可能であると考えられる。例えば前出の「Free Style リブレ」では、継続的測定に基づいて患者ごとの血糖プロフィールを解析し、低血糖に陥るタイミングの把握やインスリン投入方法の決定に活用している。同じくFDAの承認を得ているEmpatica社のスマートウォッチEmbraceは、高度な機械学習を用いててんかんの大発作や全般発作である強直間代発作を感知し、介護者に即座に警告を発信する機能を備えている。

②画像診断

光学顕微鏡やMRI、CTなどの画像診断機器によって取得した画像をベースにした画像診断において、医療機器の分解能やスループットは改良を重ねてきた。一方で、診断については長らく病理医・放射線科医に委ねられていたが、画像に含まれる豊富な情報を学習させることでAIに診断させる研究が、近年数多く出てきている。より質の高い診断の実現や診断の効率化による医師不足の解消といった効果だけでなく、人間の眼ではとらえるのが難しい手掛かりを検出することでより早期の診断や認識できていなかった病態の検出が行なえる可能性がある。

最も盛んに行われているのは、医師の判断結果を教師データとして学習させ医師よりも高い感度や正解率を目指す研究であり、例えば国立がん研究センターのグループは、ディープラーニングを活用した大腸がん及び前がん病変発見のためのリアルタイム内視鏡診断サポートシステムの開発に取り組む、隆起型の病変に関して、感度98%・特異度99%・正診率98.8%の精度を示すシステムの開発に

成功した。光学顕微鏡による病理診断画像解析や、CT、X線などによる放射画像解析、さらに MRI、超音波など他の様々なモダリティでも AI 技術の活用が積極的に検討されている。AI の診断への活用については、いくつかのステップを踏んで発展させていく必要がある。第 1 ステップは異常部位（病変）の検出であり、本件に関しては、世界的に精度は高くなってきていると考えられ、今後はいかに臨床応用（社会実装）していくかという点が大きな課題である。第 2 ステップは質的診断であり、特にゲノム医療時代においては、遺伝子変異を医用画像からの確に判断できると治療方針の決定・予後予測など臨床的にも大変意義があることから、世界中で研究が進められている。例えば、病理組織切片から非小細胞肺癌の組織分類及び遺伝子変異を高精度に予測できたという論文が米国で発表されている。今後は、このように AI によってマルチモーダルに解析・診断を試みる例が増えると考えられる。

③ゲノム診断

近年の DNA シーケンサーの高速化・低コスト化により、ゲノム解析を臨床の現場に利活用する例が出てきており、医療機関では主に疾患治療を目的として、がんの原因遺伝子の推定や先天性希少疾患の確定診断などが実用化されつつある。一般消費者向けにも、米国 23andMe 社を始めとしたゲノム解析サービスが展開されており、ヘルスケアを含んだ情報提供を行なっている。シーケンシング技術の進展により、DNA 配列の情報だけでなく DNA 修飾などエピゲノム情報も解読できるようになったため、それらの情報の活用も進むと考えられる。将来的には、AI 技術との組み合わせることでゲノム情報から健康状態やさまざまな疾患を予測できることが期待されるが、AI を適用するにはゲノム解析結果と他の医療・ヘルスケア情報との紐づけたデータを蓄積する必要がある。そのため、イギリスを皮切りに米国、中国など世界各国で 10-100 万人規模の遺伝子解析と合わせて検体や健康・医療データ収集を進めるプロジェクトが次々と立ち上がっている。

④リキッド・ガスバイオプシー

従来は液体や気体の生化学分析はベンチトップの装置を使わなければならなかったが、センサーの高感度化やシーケンシング技術の高度化を背景に、液体や気体の検体の生化学検出を行なう小型デバイスが研究されている。検体としては、呼気ガスや皮膚ガス、少量の血液や尿といった比較的取得が容易なものを対象としている。特に、DNA 片や RNA 片を検出する手法は得られる情報量が非常に多いため、解析に AI 技術が活用されている。現在実用化されている診断デバイスは主にがんを対象としたものであるが、それらデバイスにより集まる検査結果を解析することにより、他の疾患への適用拡大が期待される。

(3) 国内外のトピックス

【日本】国内メーカーからもウェアラブルデバイスがいくつか発売されているが、ヘルスケア分野としては、前述の Abbott 社の FreeStyle リブレが 2017 年に保険適用された以外で目立った保険適用例はない。研究・開発としては、もともとセンサーや材料は日本の強みであることからアカデミア・企業問わず積極的な研究が進められているが、本分野を主導するような国家プロジェクトは現状ない。日本の強みを生かした例としては、東レと NTT が共同開発した「hitoe」という計測機能性の衣類素材があげられる。応用研究として、リハビリ患者の心拍数や運動量の常時モニタリングを行なうことができる服を開発し、定量的なリハビリ支援に生かすといった実証実験がなされている。

イメージングがもともと日本の得意な分野ということもあり、AI を画像診断に応用する研究は内視鏡、顕微鏡、MRI などさまざまな診断方式において行われている。その中で、医療機器・支援ツールとして認可を受けるものも出てきている。日本では、2018 年に昭和大学・AMED の成果として、機械学習により内視鏡画像中の腫瘍と非腫瘍を区分するソフトウェア、2019 年にはディープラーニングにより脳の MRI 画像から脳動脈瘤を自動検出するソフトウェアが薬事承認を取得した。前者の

大腸内視鏡診断支援ソフト“EndoBRAIN”は医療機器として、2019年国内で承認された（オリンパス社）。超拡大内視鏡“Endocytoscopy”で撮影した画像をAIで解析し、腫瘍か否かの識別に要する時間は1秒を切り、正診率98%は専門医に匹敵する。

リキッドバイオプシーでは、大型研究プロジェクトの成果をベースに実用化を目指す例がある。名古屋大学発ベンチャーIcariaは、JST さきがけの研究成果をもとに、尿から検出したマイクロRNAにより早期がんを診断するデバイスの開発を行なっている。具体的には、ナノワイヤを用いて約1300種類のマイクロRNAを検出してプロファイル化し、機械学習により解析して診断を行なう。また、がん研・東芝・東レなどの共同のNEDOプロジェクトにおいて、血液中のマイクロRNAを検出して血液1滴で13がん種を同時診断する技術を開発し、東芝はマイクロRNAチップと検出装置の実用化を目指している。

【海外】米国を中心に医療機器として認証されるウェアラブルデバイスやAIを用いた画像診断が次々と出てきている。特にFDAは、2017年7月にデジタル技術を用いたイノベーションを適切に評価する体制を整えるために「Digital Health Innovation Action Plan」を始動させ、規制機関としてウェアラブルデバイスやAIを含むデジタルヘルス分野の安全で有益な発展の促進に積極的である。

このような状況も相まって、ここ2-3年で医療器としてFDAに認可されたウェアラブルデバイスが出てきている。例えば、血糖モニター用として前述のAbbott社に加えてMedtronic社やSensonic社、心臓モニター用としてZoll社のデバイスが認可された。

リキッドバイオプシーとしては、血液中のctDNA(循環腫瘍DNA)から肺がん患者のEGFR遺伝子変異の有無を検出する手法として、ロッシュ社のキットが2016年にFDAに承認され、日本でも保険適用されている。

またAIを使った医療支援としては、2018年に、脳卒中の可能性があることを医療提供者に知らせることができるコンピューター断層撮影(CT)のAI解析ソフトウェア、および網膜画像から医師の判断なしで糖尿病網膜症を自動診断するソフトウェア(IDx社)がFDAにより認証された。

AI技術を最大限に生かすには、質の高い大量のデータをいかに蓄積できるかがカギになるが、その中で近年、ヘルスケア分野でもGAFA、特にAppleとGoogleの存在感が下記のように増しつつある。

Apple :

- ・スタンフォード大学とのApple Watch 心臓健康調査のために、40万人を募集し医療情報を共有
- ・2018年9月、AppleWatch4の「ECG App(心電図アプリ)」FDA認可

Google :

- ・2019年11月、ヘルスケア向けウェアラブルデバイスのFitbitを買収
- ・2018年、Google社の画像認識AIによって「99%の精度で転移性の乳がんを検知できる」、「主要な2種類の肺がんを97%の精度で判別できる」など発表

(4) 課題と方向性

ウェアラブルデバイス、AIを用いた画像診断、共に今後益々普及すると期待されているが、そのためにそれぞれ固有の課題と共通した課題がある。

ウェアラブルデバイスのデバイスとしての課題は、既存の医療・ヘルスケアと比較して明確な医学的有用性を示すことができているデバイスが、今のところ限られている点である。今後、デバイスが有用性を得るための研究の方向性としては、データの収集、解析による診断に留まらず、それらの結果を基にした医学的介入、治療までの一連の操作を、単一のプラットフォーム上で統合的に行なえるデバイスの開発である。例えば、血糖値に応じて適切な量・タイミングでインスリンを注入するデバイスなどの研究が行われている。また、別な方向としては、マルチモーダルな測定によるより高度な診断であり、汗からさまざまな種類の分子やイオンを検出できるデバイスといったデバイスが開発中

である。実用的な観点では、デバイスの購入コストの抑制、小型化、バッテリーの長寿命化、非侵襲の低い測定方法の開発、セキュリティー、クラウドからエッジへのデータ解析の移行、といったハード・ソフト両面で課題がある。

AI、特に機械学習に不可欠な、質の高い医用画像の収集については、画像診断機器が他の先進国と比べても充実しているにもかかわらず、日本は他国と比較して遅れている。その要因としては、個人情報としての医療データの利用に法的制限があること、および、医療機関により情報が分散していることである。その対策として、匿名加工した医療情報の利用に関して定めた次世代医療基盤法が2018年に施行され、2019年末に匿名情報を扱う事業者が初めて認定された。データ均一性についての定めはないなど課題は残るものの、医療画像の利活用が促進されることが期待される。個人情報としての医療情報の保護と活用をいかに両立させるかは、法律家や生命倫理の専門家など様々な分野の専門家を巻き込んで今後も議論すべき課題である。

AIのさらなる適用については医療・ヘルスケア部門特有の課題がある。まず、日進月歩である領域の法規制、特に継続的に学習・変化し続ける Adaptive AI を医療機器としてどのように規制・認可するかについては、従来の規制とは異なる新しいアプローチが不可欠である。それに対し、FDAは Adaptive AI を含んだ医療機器を把握・規制するための新たな枠組みの検討を2019年4月に開始し、日本でも2019年11月に Adaptive AI を使った医療機器の審査を簡素化するための条項が改正薬機法に盛り込まれた。また、日本では今のところ「AIは診断支援ツールであり判断主体は医師である」という見解であるが、今後AIがより高度化した場合、AIによる診断の位置づけを再考する必要が出てくると考えられる。

最後に、デジタルヘルス分野に限らず AI 開発において無視できないのが、大量データ収集のプラットフォームやデバイスを持つ GAFA である。また、中国では、政府方針により医療系データについてはテンセント社に独占的に集約させ、AIによるソリューションの実用化を加速させようとしている。AIを扱える技術者が全世界的に不足している中、GAFAや中国系IT大手が囲い込もうとしている状況もあり、この分野でいかに日本の存在感を示していくかは大きな課題である。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 ライフサイエンス・臨床医学分野（2019年）」（CRDS-FY2018-FR-04）（2019年3月）
- ・ 科学技術振興機構「出資型新事業創出支援プログラム（SUCCESS）における Icaria 株式会社への出資について」
<https://www.jst.go.jp/entre/release/190213.pdf>
- ・ 株式会社東芝「血液1滴から13種類のがんを99%の精度で検出する技術を開発 ー独自のマイクロRNA検出技術を使った健康診断などの血液検査により、生存率の高いステージ0からがん罹患の有無を識別ー」株式会社東芝 研究開発ライブラリ
https://www.toshiba.co.jp/rdc/detail/1911_06.htm
- ・ U.S. Food and Drug Administration, “Statement from FDA Commissioner Scott Gottlieb, M.D. on steps toward a new, tailored review framework for artificial intelligence-based medical devices”,
<https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-steps-toward-new-tailored-review-framework-artificial>

3.3 農業

(1) 定義・意義

現在の農業が直面する最大の課題は、気候変動と環境負荷低減であり、これは世界共通の難題である。気候変動に適切に対応しながら、環境負荷を低減させつつ収量増加を目指すため、無駄なく最大の収量を確保するためのスマート農業への取り組みが世界的に進められている。一方、日本の農業の現場では、農業従事者の高齢化による労働力不足が深刻であり、農業生産性・効率性の向上、農作業による負荷軽減の観点からスマート農業の実現と普及に期待が集まっている。

基礎研究では、作物の成長や収量と圃場環境について精密な計測で数値化してモデル化し、シミュレーションすることで、適切な栽培介入法を求めようとする努力が続けられている。応用～実装研究では、各種センサーを駆使して圃場環境を正確に把握し、商業規模の栽培で過剰な施肥を防ぐ可変施肥や、農薬のピンポイント散布、農家の農作業負荷を軽減するロボティクスの研究が進められている。

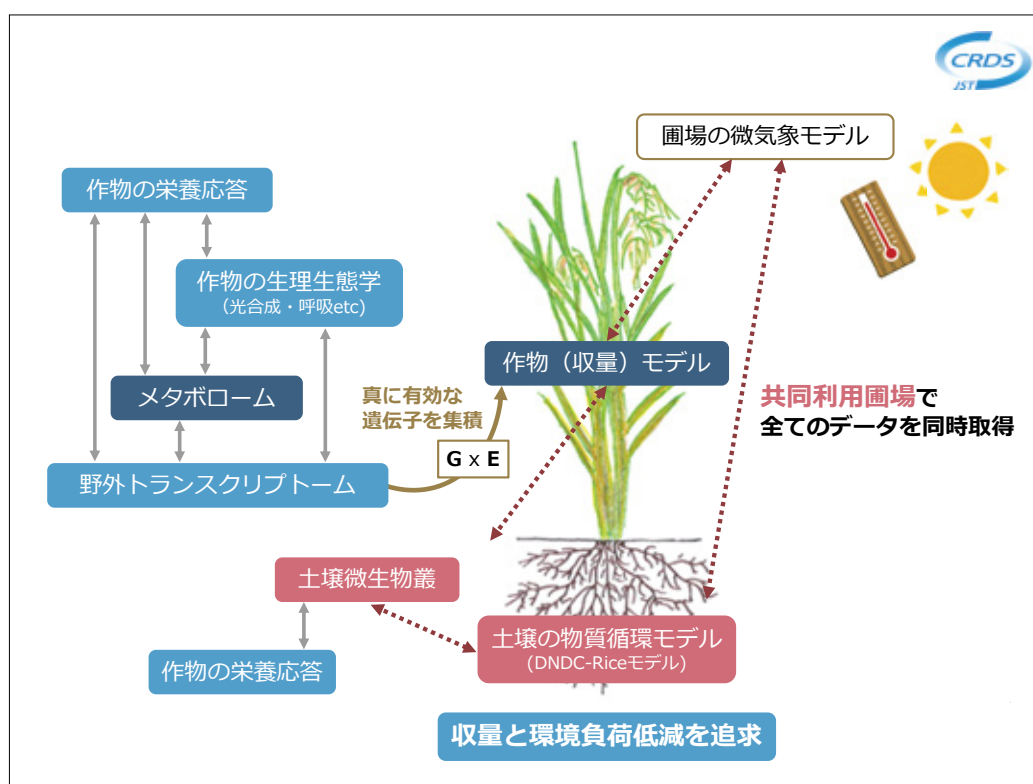


図 3-3 気候変動に対応し、環境負荷低減を目指す農業のための融合的アプローチ

(2) 研究のトレンド

作物の成長と収量は環境条件に大きく影響されるので、作物にとってより適切な環境を整備するための栽培介入が行われる。近年、作物の成長は土壌や圃場の気象条件と互いに相互作用し、どれか一つを改良してもベストな状態にはならないことが分かってきた。

圃場の作物は、土壌から水分や窒素、リン、その他のミネラル成分などを吸収し、光合成によって大気中の（二酸化）炭素を体内に固定し、人間がその固定された炭素化合物を、食糧や燃料源、繊維や飼料など、様々な形で利用する。農業では人間が利用しなかった農業残渣（例えば稲藁など）を土

壤に戻ることが多いため、光合成によって固定された炭素化合物や、土壌から吸収した窒素、リンなどは、ある程度土壌中に戻ることになる。農業残渣などの有機物は土壌中の微生物層によって分解され、有機物として地中に残ったり、二酸化炭素やメタンとして大気中に放出されたりする。また、土壌中の微生物の中には、植物に窒素やリンを供給するものがあることが知られている。このとき、植物は微生物に見返りとして炭素化合物を与える。このように、圃場では、窒素、リン、炭素化合物、水分などの様々な物質が、土壌、作物、大気中といった圃場環境を循環しているのである。植物の生育を含む、こうした物質循環は、温度や湿度といった気象条件に大きく左右されるが、植物群落そのものが圃場の気象にも影響を与えることが知られている。例えば、植物の蒸散活動によって葉面温度が大きく変わり、また蒸散活動は土壌からの吸水速度にも影響を与える。このように植物自身の活動が圃場の微気象に影響を与え、収穫量だけでなく、圃場の物質循環全体に影響を及ぼす。この圃場の物質循環を理解することは、より効率よく高収量を得るためだけでなく、農業の環境負荷を低減し、持続可能な農業へと転換するためのカギとなる。

圃場の物質循環の理解のためには、以下の3モデルを核としたプロセスモデルをリンクさせ、作物が圃場環境と相互作用しながら収量形成を行う過程を数値で辿れるようになることが不可欠である。即ち、土壌の物質循環モデル、作物（の成長と収量形成）モデル、圃場の微気象モデルである。現在、土壌の物質循環については、DNDCモデルが構築され、精密な計測に基づいた、土壌微生物-土壌-作物の物質循環をプロセスモデルとして計算・予測できるようになっている。作物の成長と収量形成については精密な実測データを積み上げた精度の高いプロセスモデルが稼働しており、環境条件を入力すれば、開花期の予測や最終収量予測が可能である。圃場の微気象については、植物の蒸散量と草体の表面積に依存してパラメータが変化するため、遺伝子型を取り入れた圃場の微気象モデルの構築が進められている。これら三つのモデルでは、作物の品種によって異なるパラメータを計測によって求めて使っているが、どんな遺伝子がどんな環境条件でどれだけ収量形成に貢献するかを数値評価として知ることができれば、地域ごとの土壌特性と気象特性に最も適した作物をデザインし、環境負荷を低減させることも可能になる。そのための試みとして、生理生態学の知見を活かした新品種育成のための研究が進められている。ここでは、光合成効率やストレス耐性などの生理生態学的特質をターゲットにした量的形質遺伝子座（QTL）解析や、ゲノムワイド関連解析（GWAS）などが行われている。また、実験室で解析された分子メカニズムやオミクス情報と圃場での表現型とを融合させる試みの一つとして、大規模野外トランスクリプトーム解析が行われた。ここでは、トランスクリプトームデータと気象データを照合し、環境応答・体内時計・日齢を取り入れた統計モデリングを行うことで、トランスクリプトーム変動の多くの部分を予測可能なモデルが得られた。この研究は、野外環境のような複雑な環境においても、現実的な学習データ量からトランスクリプトーム変動の予測を行うモデルが作成可能であることを示した点で画期的であった。

農作業の効率化・負荷軽減のため、センシングやロボット技術を活用した取り組みが進められている。ドローンによる農薬散布や、自動走行トラクタ、収穫ロボットなど、大規模化と省力化を可能にする技術が開発されている。各種センサーやハイパースペクトルカメラ、衛星からの観測データを用いた生育状況の把握・管理に向けた研究開発も進んでいる。また、これまで熟練農業者の作業ノウハウ（暗黙知）として蓄積されていたプロセス条件を、センシングデータのビッグデータ解析、AIの活用等により形式知化する取り組みもおこなわれている。

(3) 国内外のトピックス

欧州では、Modelling European Agriculture with Climate Change for Food Security (MACSUR) がFACCE-JPI (Joint Programming Initiative for Agriculture, Climate Change, and Food Security) と呼ばれる共同研究の中で立ち上がり、18か国70機関のグループが、モデルと各国のデータベースを活用

したモデル評価が行われた。対象は、作物だけではなく、家畜や草地も含まれ、これらを横断して評価するような試みも行われた。MACSUR と AgMIP 間では MOU が締結され、データやモデルの有効活用が図られている。

圃場における光合成関連形質を網羅的に評価するための計測機器の開発が急ピッチで進んでいる。David Kramer のグループではクロロフィル蛍光パラメータなどを簡易に評価可能できる低価格光合成測定システム PhotosynQ を開発し、世界的に普及している。またオーストラリア CSIRO の John Evans のグループではハイパースペクトルカメラを用いて葉の光合成に関わる生化学的性質を予測するモデルを開発した。日本では葉のガス交換速度を 15 秒程度で測定可能である迅速光合成測定装置 MIC-100 が 2018 年より市販されている。NDVI や群落温度を測定するスペクトルカメラや熱画像カメラ、その解析プログラム等は広く普及しており、トラクタやドローンなどに搭載するなど様々な方法が試みられている。

日本では 2018 年より農林水産省戦略的プロジェクト研究推進事業「民間事業者等の種苗開発を支える「スマート育種システム」の開発」が開始された。1. 育種ビッグデータの整備および情報解析技術を活用した高度育種システムの開発、2. 民間事業者、地方公設試等の種苗開発を支える育種基盤技術の開発、の 2 課題を柱とし、ゲノム・作物形質のビッグデータデータベースの構築および育種の加速化・効率化技術の開発を目的とする。また、スマート農業に関する府省横断的な取り組みとして、総合科学技術・イノベーション会議が推進する戦略的イノベーション創造プログラム (SIP) の枠組みの下、「次世代農林水産業創造技術」(SIP 第 1 期：2014 年～2018 年) および「スマートバイオ産業・農業基盤技術」(SIP 第 2 期：2018 年～2022 年) が実施されている。

豪州では作物光合成改良を目的とするプロジェクト "ARC center of Excellence for Translational Photosynthesis" が行われ、Australian Research Council から 22M 豪ドル /7 年の資金を受けている。(i) CO₂ 同化効率の向上、(ii) 光受容効率の向上、(iii). 自然変異の利用、(iv) 圃場生産性の検証をテーマに掲げ、遺伝子操作による植物改変や、ハイスループット計測技術の開発などを進めている。

欧州では植物のフェノタイピングに関わる国際コンソーシアム "European Plant Phenotyping Network" が組織されている。全 22 の研究機関が参加し、圃場環境や植物形質のハイスループットフェノタイピング技術の開発やその活用方法について研究が行われている。同様のフェノタイピングプロジェクトは加、米、中、豪でも実施されている。

米国では New York University の Michael D. Purugganan らのグループが、フィリピンの IRRI (International Rice Research Institute) と共同でイネの野外トランスクリプトーム研究を行っている³⁷。サンプル数はそれほど多くないものの、熱ストレスや乾燥ストレスをかけたサンプルのデータと合わせた解析などに特徴がある。

モンサント傘下の Climate Corporation は高分解能の気象観測とシミュレーションをもとに農家向けに栽培管理支援から保険商品の提供まで統一的に行っている。

(4) 課題と方向性

作物の成長や収量と圃場環境についての統合モデルに、肥料分吸収に関わる分子遺伝学的な知見、大規模野外トランスクリプトームなどの結果を合わせて導入することで、土壌特性や気象特性に合わせた最適な遺伝子型を持ち、気候変動に対応しながら環境負荷を低減する作物のデザインが可能になるだろう。

土壌、作物、そして群落の微気象モデルを融合して、環境と耕地管理との交互作用を定量的に評価する手法を開発し、圃場の物質循環をコントロールすることは、温暖化への適応、および環境負荷低減農業の実現のために不可欠である。しかし複合的な気象条件を取り扱い、土壌や作物の生態生理学的な特性を反映した包括的なモデルを構築して運用し、モデルの問題点を洗い出して、改良を進めて

いくには、十分なデータがない。

環境負荷を下げるために養水分投入は抑制されねばならず、水利用効率・肥料利用効率の向上の重要性が今後一層高まる。新たな遺伝要素の解明に加え、これまで特定された収量性や環境ストレス耐性に関わる遺伝子や育成系統を用い、遺伝子型 (G) × 環境 (E) との相互作用の解析研究が加速される必要がある。あわせて、地球温暖化による圃場環境の変化の予測と、それに適応した作物の育成研究の強化が求められる。

転流や登熟など収量に直結するが解析が難しい生理学的性質について、その評価を容易にする手法の開発が求められる。また作物のシンク・ソース・転流のバランスの最適化が今後の課題となる。

以上のような知見を気候変動の研究に取り入れるには、組織的あるいは協調的な仕組みが必要である。

ようやく開放系大気 CO₂ 増加実験 (FACE) などの環境操作実験を利用して、作物モデル精度比較と改良を行う試みが始まったが、これまでの環境操作実験では、CO₂ 濃度、温度、水といった単独の環境要因を扱うことが多かった。現実の気候変動下では様々な環境要因が複合的に変化するので、それらを複合的に取り扱うことができる新規の環境操作実験プラットフォームを構築し、そこから精度の高いデータを取得して、モデル研究で問題となった新たな課題（仮説）に取り組む必要がある。

まずは、農耕地における作物生理や物質循環における環境と相互作用を研究する拠点を構築し、農学、生態学の研究者で連携しながらモデル化研究を進めることで、モデル開発に有益なデータを得ると同時に、物理的・数学的センスで複雑現象の解明に取り組む人材も育成することが重要である。

農作業の効率化・農業生産性の増大に向けて期待されるロボットや自動農機の実現には、高精度の測位技術とその低価格化が求められている。LIDAR、カメラ、測位システムなどのセンサーフュージョンにより、実現する必要がある。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 ライフサイエンス・臨床医学分野 (2019 年)」(CRDS-FY2018-FR-04) (2019 年 3 月)
- ・ 科学技術振興機構 研究開発戦略センター 「戦略プロポーザル 次世代育種・生物生産基盤の創成 (第 3 部) 気候変動下での環境負荷低減農業を実現する基盤の創出～圃場における微生物、作物、気象を統合的に扱うモデルの開発に向けて～」(2020 年 3 月刊行予定)
- ・ 永野惇、工藤洋「野外の環境における生物の環境応答の理解にむけて：トランスクリプトームデータと気象データの統合」『領域融合レビュー』2014;3(e009):1.
DOI: 10.7875/leading.author.3.e009
- ・ 長谷川利弘「作物の環境応答における生理生態・遺伝・微気象モデルの融合的アプローチ」『学術の動向』 2016 2 月号 :77.
- ・ 科学技術振興機構 研究開発戦略センター「科学技術未来戦略ワークショップ報告書 多様なデータの取得・統合処理を可能にするセンサー融合基盤技術～センシング情報の高付加価値化に向けて～」(2020 年 3 月)

3.4 生物多様性

(1) 定義・意義

生物多様性や生態系は社会の持続的発展に多様な価値をもたらす源泉であり、持続可能な開発目標（SDGs）の中でも陸域と海域それぞれの持続可能な管理・活用に向けたゴールが掲げられている。関連する研究開発には生物多様性や生態系そのものの理解（維持機構解明など）、時間的・空間的な分布の変化等の把握や予測、あるいは生物多様性や生態系の管理・活用のための手法の開発などが含まれる。またこうした分野の研究開発では、様々な生物種の時間的・空間的な分布やその動態を従来通りの方法で追跡するばかりでなく、新しい技術や手法との融合が積極的に行われている。昨今では情報通信技術、センシング技術、遺伝情報の解析技術、機械学習等を用いた解析や各種の統計モデル、生物や物理プロセスを網羅したモデル等、様々な基盤技術の進展が本分野の進展を支えている。本節ではそうした各種技術や手法の進展・普及により利用可能なデータの種類や量が増大している状況、ならびにそれら膨大なデータの活用を生物多様性や生態系の仕組みの解明等につなげるための解析技術が進展しつつある状況について紹介する。

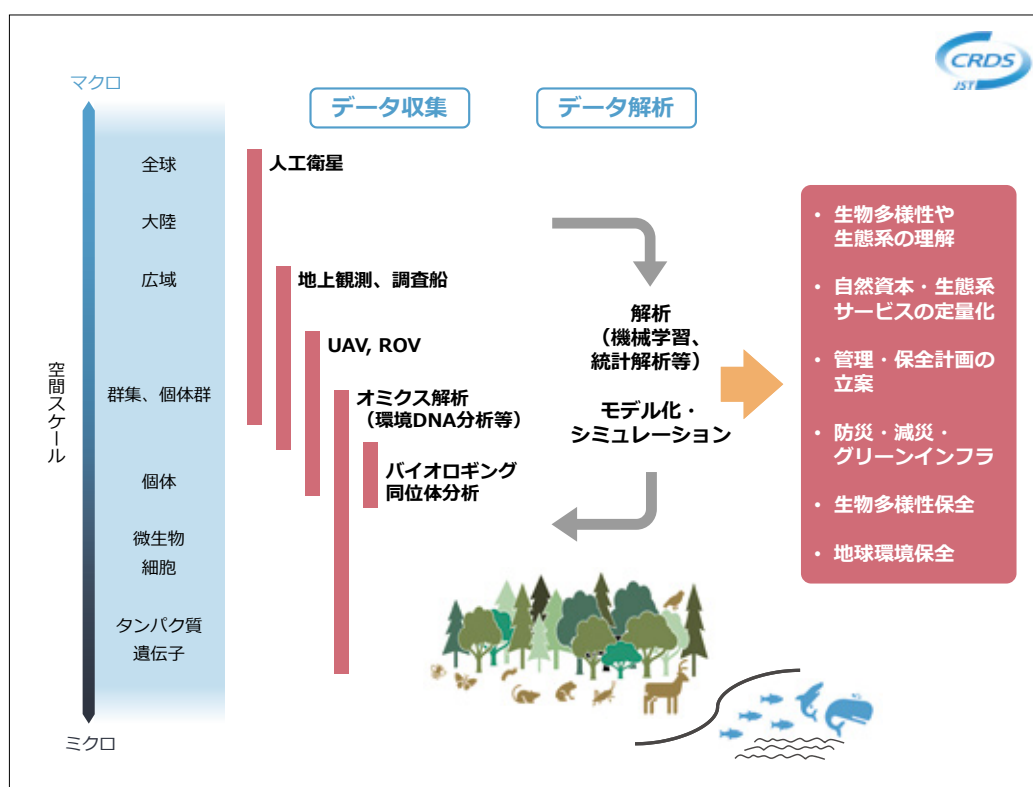


図 3-4 生物多様性・生態系分野におけるデジタルトランスフォーメーションの概観

(2) 研究のトレンド

従来、生物多様性や生態系の研究では多大な労力と時間をかけて個体群や群集あるいは種レベルの分布調査やモニタリングが行われてきた。これに加えて最近では近年のゲノム科学の進展を背景に遺伝子情報の蓄積・活用が進んでいる。特に水、土壌、大気などの環境中に存在する生物由来の DNA を採取し分析する「環境 DNA 分析」が活発に行われるようになってきている。この背景には第 3 世代

PCR（ポリメラーゼ連鎖反応）とも言われるデジタル PCR を利用した遺伝子解析技術の進展がある。

環境 DNA 分析は従来の調査やモニタリングに完全に置き換わるものではないが、生物種の生息状況の把握を以前よりも効率的に行うことができ、かつ遺伝情報など新たな情報を付加できる可能性も秘めていることから注目されている。例えば特定の種の在・不在を把握するだけでなく、環境中に存在する生物種を DNA 塩基配列から系統学的手法によって特定する「DNA バーコーディング」、あるいはそれを用いた網羅的な特定手法である「メタバーコーディング」やメタゲノミクス解析を行うことで生物多様性の迅速な把握につながると期待されている。DNA バーコーディングに関しては「International Barcode of Life Project、iBOL」という国際的なライブラリー構築プロジェクトが2010年に開始された。国内でも微生物に関して多種多様な情報を遺伝子・系統・環境の3つの軸に沿って整理統合したデータベース「MicrobeDB.jp」が既に開発・運用されており、微生物群集の定量化が一層容易になっている。また昨今は DNA のみではなく RNA、代謝物、脂質など解析対象の拡大も見られ、一連のエコ・オミクス研究の今後の進展が注目されている。

動物の食性や物質循環に用いられてきた安定同位体分析についても、近年では炭素・窒素以外にも様々な元素を測定する多元素分析によってより精緻な解析ができるようになってきている。また分子を対象にした同位体分析も行われており、食物網構造のような生態系の仕組みの理解に基づいた解析もできるようになってきている。更に環境傾度を表す同位体地図（Isoscape）を作成することによって物質の移動や生物の移動を解析する手法も出来つつあり今後の発展が期待されている。

生態系レベルの広範な時空間スケールを横断する観測にはリモートセンシングが有用な技術である。現在は地球環境変動の衛星観測ミッションが展開されており、高頻度、高解像度、多バンドの衛星リモートセンシングが計画されている。日本では2017年に気候変動観測衛星「しきさい」(GCOM-C)が打ち上げられて海洋・陸域環境の観測が強化された。温室効果ガス観測技術衛星「いぶき」(GOSAT)は大気中の二酸化炭素とメタンの観測を継続している。近年はこれら衛星からのデータを活用してフィールド下の生物多様性や生態系の研究とリモートセンシング観測技術の融合を進める取り組みが行われている。例えば地球観測に関する政府間会合（GEO）の「生物多様性観測ネットワーク」(GEO BON)には国際的な地域活動として「アジア太平洋生物多様性観測ネットワーク (AP-BON)」があるが、その AP-BON では生物多様性に関する生物学的調査とリモートセンシング観測の融合がいち早く提案されている。

衛星観測データの無償公開とその活用も近年急速に進んでいる。ランドサット衛星の画像データが2013年から無償公開されるようになったため、森林変化等の全球レベルでの定量的評価がそのデータを使って進められている。その他にも米 NASA の地球観測衛星に搭載された可視・赤外域の放射計 (MODIS または Moderate Resolution Imaging Spectroradiometer) の取得データが無償公開されている。

フィールドにおけるセンシング技術も多様化している。データロガー、マイクロコンピューター、カメラ、レコーダー、測位・情報通信技術等の普及により様々な生物と環境に関する局所スケールでのトラッキングやデータロギングが進展している。今後、従来困難であった小型動物や海洋生物についての行動データ、植物や陸上植生の季節性（フェノロジー）の年変動や地理的分布に関する画像データ、生物・非生物を問わず長期の連続観測データ、移動や分布データの収集が進むと期待されている。

加えて局所リモートセンシング技術ではドローンなどの無人機 (Unmanned Aerial Vehicle、UAV) に代表されるように従来の衛星や航空機よりも小型かつ自律的なプラットフォームが急速に普及している。海洋分野でも技術的には UAV 相当である自律型無人探査機 (Autonomous Underwater Vehicle、AUV) や洋上中継器 (Autonomous Surface Vehicle、ASV) といった遠隔操作無人探査機 (Remotely Operated Vehicle、ROV) が存在し、海洋保護区でのモニタリングに活用するなどの検討が進められている。2017年頃から中国をはじめとするドローン系の企業による安価な ROV が水中ドローンとして販売され始めたという動きもある。水中での観測は音響技術の蓄積があり、音響デー

タ合成開口技術、地層データの自動合成などが試みられている。

データ取得方法の多様化により巨大なデータ基盤が構築されつつある中、その膨大なデータの処理や解析技術にも進展が見られる。例えば画像解析をはじめとしてクラウド上での解析ツールが充実してきており、ドローンなどにより得られた画像データの SfM (Structure from Motion) 技術による結合をオンライン上で行うツールが普及してきている。環境閾値 (臨界点) の予測に関しても、従来は理論研究や微生物を用いた単純な系の実験だったが、リモートセンシングにより得られた画像を用いた大陸スケールでの解析等も進みつつある。

衛星画像の活用についても Google Earth Engine 上にランドサット衛星データ等がアーカイブされており、高速な処理が全球で一括して実施できるようになっている。生物の分類にはまだ使えないものの、画像認識についても既に Google や Microsoft のクラウドでラベル化された情報を抽出できるようになっている。そのため例えば植生状況を把握するための代表的な指標である正規化植生指数 (Normalized Difference Vegetation Index, NDVI) を季節ごとに全球一括で毎年計算することも可能となっている。これらの技術は地域スケールから全球スケールまでの生態系の変化の把握に貢献しつつある。ただ、その反面、こうした民間のクラウド上での解析は利便性に優れるものの、提供企業に偏った情報集積を招く可能性があり、その点への懸念が指摘されている。

高度な解析手法の開発も進んでいる。状態空間モデルを含む階層ベイズモデル等のベイズ統計、非線形時系列解析法 (Empirical Dynamic Modelling, EDM)、あるいは機械学習を取り入れた因果関係推定法やその応用に係る研究が注目されている。特に機械学習を生物多様性や生態系の研究に取り込む事例は近年急激に増加している。例えばディープラーニングを取り込んだ事例は個体、集団、景観、生態系、地球規模まで様々なスケールの研究で見られ、その目的も同定・分類、疾病検知、数や密度の推定、多様性の推定、保全・再生・資源管理への応用等、多岐に亘っている。

データの多様化・膨大化を背景に、生物多様性や生態系の評価のためのモデルの開発・高度化も進んでいる。特に生物分布情報と環境データを用いて各生物種の地理的分布の予測を行うモデリング手法である生態ニッチモデリングの発達が著しい。生態ニッチモデリングの中にも多数のモデリング手法があるが、主要なものとしては Maxent、Biomod2、Random Forest、複数モデルの平均などが知られている。海洋では生態系モデルも多数存在し、Ecopath や Atlantis 等が使用されている。また生態系サービスを評価するためのモデルも開発が進んでいる。気温や降水量などの環境情報と植生図などの生物情報を初期条件として入力することにより生態系サービスの潜在的な供給量の数値化・地図化が可能なモデルが多数提案されており、GUI 操作が可能なソフトウェアも整備されている (InVEST、TESSA、ARIES、LUCI 等)。さらにこれらを使って生態系サービスのホットスポットの特定、生物多様性保全優先区との重ね合わせ、生態系サービス間のトレードオフの検証等を行う事例も世界各地から報告されている。

(3) 国内外のトピックス

国内における対応や状況としては、まず生物多様性や生態系の観測に関して、複数の国際プログラムの日本ノードを各々に設立し、国際的な連携の下で取り組んでいる。衛星による観測では、JAXA が国際宇宙ステーションの日本実験棟「きぼう」に搭載する地球観測用 LiDAR (レーザを照射することで地面の標高や森林の樹冠高を高精度に計測できるセンサー) の開発に 2021 年頃の打ち上げを目指して取り組んでいる (MOLI ミッション、Multi-footprint Observation Lidar and Imager)。

近年は遺伝子情報の蓄積・活用の動きが活発である。2018 年には環境 DNA の普及に応じて「環境 DNA 学会」が発足している。遺伝子情報のデータベースの整理も進みつつあり、例えば微生物に関しては前述のとおり「MicrobeDB.jp」が既に開発・運用されている。

生物多様性や生態系に関するデータの収集・蓄積には多くの時間と労力がかかるが従来はそれだけ

では学術論文となりづらく、研究成果としての評価が十分でなかった。またオープンデータ化の国際的な流れもある。こうしたことを背景に、日本生態学会の国際学術誌 *Ecological Research* では Data Paper というデータの提供と公開に特化した学術論文のセクションが追加された。

国際的な動きとしては、長期かつ広域の観測に基づく生物多様性の評価・予測に関連した注目事例を以下に挙げる。

- ・「PREDICTS」・・・英の大学等研究機関が中心となって世界中から陸域の生物多様性の分布情報を集約し、そのビッグデータを活用した統合的な研究に取り組んでいる。2018年7月現在では38,000種以上を含む250万件以上の生物多様性情報が記載されている。これまでの主要成果としては土地改変が陸域の生物多様性に与える影響の評価や、それを基にした臨界点（閾値）の評価研究等がある。
- ・「BIOTIME」・・・英セントアンドリュース大学のグループが牽引する国際プロジェクトで、生物多様性の時系列変化の実データを収集している。主要な成果としては、世界中の陸域・海域の生物群集を対象に、種数と種構成の時間変化を定量化した研究や、熱帯の淡水群集を対象にした種構成の時間変化の定量化等がある。
- ・「Biodiversity Exploratories」・・・独の大学等研究機関が中心となって土地改変が生物多様性と生態系サービスに与える影響評価に取り組んでいる。独国内の3地域を対象に、1,000以上の調査区を設け、動植物や微生物についてのデータを収集している。300名以上の研究者やスタッフが関わっており、個別に50以上のプロジェクトが進行している。近年は土地改変が生物相の均質化をもたらすことや栄養段階の異なる生物群集の多様性が生態系の多機能性を支えること等を示した研究成果をあげている。
- ・「Isobank」・・・米ニューメキシコ大の研究者が中心となってNSFの支援により2018年に開始された研究プロジェクト。安定同位体分析のデータを一元的に蓄積・管理し、その活用を促進するためのデータプラットフォーム作りを目指している。
- ・ワークショップ・・・米National Center for Ecological Analysis and Synthesis (NCEAS) や独German Centre for Integrative Biodiversity Research (iDiv) は世界中で行われた実証研究や実験研究のデータを統合したり、理論研究等を統合したりすることを目的としたワークショップを頻繁に開催している。世界中から来る参加者の旅費を開催者側が負担する代わりにそこでの成果を統合研究としてNatureやScience等へ発表するといった仕組みが構築されている。
- ・「iNaturalist」・・・カリフォルニア科学アカデミーによって運営されている市民科学のプロジェクト。世界中の自然愛好家と研究者が集うソーシャルネットワーキングサービス（SNS）で、地球上の生物多様性に関する観察記録を共有し、種同定を助け合い、地図上に残すこと等に取り組んでいる。2015年10月の時点で既に約40万種に関する190万件の生物確認情報や個体写真が集積されていた。

(4) 課題と方向性

DNA バーコーディング解析は歴史が浅いため、情報の蓄積量は既知の種を網羅するには未だ遠い状況と言われている。比較的進んでいる土壌圏の生物でも情報が不足していると言われている。今後はDNA バーコーディング可能な分類群を広げるため、プライマーの開発や種同定のための情報の追加が課題である。また在・不在を確認するだけでなく、存在する種の個体数やバイオマス量なども把握可能な技術へと発展させることも課題である。

遺伝子情報の抽出や海洋観測技術をはじめとして技術的には容易に多様な大量の情報が得られるようになってきたが、第一線の研究に必要な量と質の情報を収集し利用可能な状態にするための金銭的・労力的コストは従前と変わらず、研究実施にあたっての障壁となっている。例えば大規模な観測イン

フラの場合にはその維持・管理やデータ取得そのものに費やされる時間と労力が非常に大きい。膨大なデータを解析するための技術や計算インフラの整備も必要になる。また異なるプロジェクト等で個別に構築された複数のデータベースを整理、統合して横断的に利用するための技術・手法の普及が必ずしも十分に進んでいない。

加えて日本では長期的視野をもったモニタリングやデータベースの継続的運用の多くは研究者個人による資金調達に依存しており、プロジェクト終了のたびに財政的不安を抱えざるを得ない状況がある。研究者・技術者等の人材不足も深刻化している。日本では情報科学の専門家が生物多様性や生態系研究の分野に関わってくる例がまだ非常に限られてもいる。

また海外では前述の事例のように世界中に散在するデータを収集・集約して統合的に解析するプラットフォーム作りが様々な形で進められている。しかしそうした取り組みは現在のところ日本では見られない。世界と伍する研究を実施していくためには世界中を巻き込むような大胆な研究構想力も求められる。

以上から、今後はデータの収集、集約、解析・活用において更なる技術開発や手法開発が必要であり、また同時に新たな研究者の参入を促す仕組みや研究構想などの立案も重要となる。

(5) 参考資料

- ・ 大澤剛士、神保宇嗣「ビッグデータ時代の環境科学ー生物多様性分野におけるデータベース統合、横断利用の現状と課題ー」『統計数理』第 61 巻 第 2 号 (2013) : 217-231
- ・ Sylvain Christin, Eric Hervet & Nicolas Lecomte, “Applications for deep learning in ecology” . *Methods in Ecology and Evolution* vol. 10 (2019)、1632-1644
- ・ 科学技術振興機構 研究開発戦略センター 「研究開発の俯瞰報告書 環境・エネルギー分野 (2019)」 (CRDS-FY2018-FR-01) (2019 年 3 月)

3.5 気象・気候

(1) 定義・意義

地球温暖化、オゾン層破壊、砂漠化、海洋汚染、酸性雨など、地球規模もしくは広域規模での環境問題が深刻化する中で、地球表層・海洋・大気についての必要十分な観測と、それに基づく地球規模の中長期的な気候変動予測技術の高精度化が強く期待されている。さらには、IPCC（気候変動に関する政府間パネル）第4次報告書で、20世紀後半以降の温暖化が人間活動によるものとほぼ断定されて以降、気候モデルを高解像度化し、今後避けられない温暖化に社会が適応するための政策（適応策）立案に資するよう、地域ごとに詳細な情報を発信する動きが盛んである。

また、気候モデルによる予測は、適応策だけでなく、温暖化の進行そのものを止める緩和策の立案にも貢献している。気候モデルに生態系の働き等を組み入れた地球システムモデル（ESM）が用いられることが多く、さらには、温暖化と人間社会との相互作用の整合的・包括的推定には、社会経済分野で開発されている統合評価モデル（IAM）と、ESMとの結合モデルを開発することが望まれている。

気象・気候の研究分野は、計算科学（シミュレーション）、そしてスーパーコンピュータの発達に伴っていち早く発展してきている。逆に言えば、計算機の性能向上がこの分野の研究を牽引してきている。そして、環境観測技術の高度化・多様化により環境観測データのビッグデータ化やオープン化が進み、容易にデータが入手できる環境が整ってきた。近年では、それらの膨大なデータをAI・機械学習等と統合する研究への取り組みも広がりつつある。さらにこれらのビッグデータとしての気象データは、農業・小売業・運輸業といった幅広い産業へ活用され、新たな気象ビジネスによる社会的価値の創出という局面を迎えている。

(2) 研究のトレンド

観測データの取得とそのプラットフォーム、気候変動モデルによる予測・評価、そして局地気象予測の3つに分けて研究のトレンドを紹介する。

まず、観測データに関して、現在、人工衛星によって全球の地表面および近傍の表層を周期的・継続的に計測することが可能となり、国際的枠組みのもと、各国の観測衛星データが情報共有化されている。また、大気中の温室効果ガスや微粒子（エアロゾル、雲）、短寿命気候汚染物質（SLCPs）等の濃度やその変化、気候変動に大きな影響を与える海洋や極地、森林、土地利用変化等についても衛星観測や地上／船上／航空機観測によって、多種多量の観測情報・データを取得することが可能となった。さらに、無人航空機（UAV）・遠隔操作型潜水機（ROV）・自律型無人機（AUV）等の新しい機器を利用した観測が普及しつつあり、それに伴って関連する各種センサーの開発も進められている。

このように膨大な観測データの取得が可能となり、それらを蓄積し、相互に利用するために、国内外でクラウドベースあるいは大型計算機システムを用いたプラットフォームの整備が進んでいる。日本では国主導のDIAS（データ統合・解析システム、MEXT）や民間企業によるプラットフォームやクラウドサービスが提供されている。

2009年の米国地質調査所（USGS）Landsatデータの無償化を皮切りとして、欧州 Copernicus 計画における地球観測データ利用システム「Copernicus Data and Information Access Services（Copernicus DIAS）」や米国海洋大気庁（National Oceanic and Atmospheric Administration; NOAA）における「NOAA Big Data Project（BDP）」等、衛星データに関するプラットフォームの状況は大きな転換期を迎えている。すなわち、今まで限られた専門家等のみが取り扱うことができた情報やデータが、個人や民間企業でも手軽に利用できるように開放されたということだ。この流れを受け、日本では2019年2月に Tellus（政府衛星データのオープン＆フリー化及びデータ利用環境整備事業、METI）が一般公

開され、政府衛星データを利用した新たなビジネスマーケットプレイス創出を目的としたプラットフォームとして注目されている。

次に気候モデルによる予測・評価に関して、地球温暖化の顕在化という認識が一般社会でも広まりつつある。その影響の深刻化を避けるためにこれらの予測・評価がますます重要となっている。温暖化が不可避で、ますます深刻化する将来に向けては「緩和」努力に加え、温暖化を前提として「適応」していくことが欠かせない。「適応策」はあらゆる分野に関わり、その内容・優先順位は地域ごとに異なっている。そのため、気候変動の将来予測、その地域レベルへのダウンスケール、その結果を踏まえた地域への影響予測、地域特有の脆弱性とリスク評価、優先順位の特定、対策技術、社会実装化など、多岐に渡る研究が実施されている。

世界的に見ると、気候変動によって熱ストレスによる死亡者数、下痢性疾患、洪水による死亡者数の増加や、感染症を媒介する生物の生息可能域の拡大などを通じたマラリアやデング熱による死亡の増加がもたらされると予測されている。わが国のような先進国では、熱ストレスによる影響と洪水や台風などによる災害などの極端な気象による影響も大きい。2018年の猛暑に相当するような異常高温が発生したことや、特にここ数年連続で集中豪雨が発生し、各地で甚大な被害が起きた。

そのため局地気象予測では、より高精度の気象予測や信頼性の高い防災気象情報の提供が求められている。現在では気象観測データが充実し、スパコンを用いた数値予報の精度もかなり向上している。この数値予報は天気予報・防災気象情報として翻訳され（予報をガイドすることから予報ガイダンスと呼ばれている）、広く一般市民に届けられている。ちなみに気象庁では1995年より予報ガイダンスに、逐次学習型の手法であるニューラルネットやカルマンフィルターなどの機械学習を取り入れている。しかし、難しいのは短時間で大きく変化するメソスケール現象（数百キロ以下の規模を持つ気象システム）の予測である。メソスケール気象の特性が機械学習でのパターン認識を困難にしていると推測されており、データ駆動型研究の更なる深化が期待されている分野でもある。

(3) 国内外のトピックス

・国際Argo計画による海洋観測

海洋観測の分野では、漂流しながら潜航・浮上を繰り返すことで水温や塩分の鉛直構造を自動で観測できるArgoフロートが開発され、それを用いた国際Argo計画の全海洋規模展開によって、海洋の物理的変動の観測が飛躍的に発展した。Argoフロートは現在、およそ3900台が展開されており、10日に1回ほどの頻度で、水温・塩分の鉛直分布を水深2000mまで測定し、準リアルタイムでデータが陸上に送信されている。送信されたデータは、取得後24時間以内に全球気象通信網（GTS）により、世界中の気象機関などに配布されるほか、6ヶ月以内に高度な品質管理を施されて、インターネット上で提供される。なお、観測データの高品質化に機械学習等を用いる試みが行われている。

・観測データのプラットフォーム化、大規模データアーカイブ化と機械学習・ビッグデータ解析

膨大な衛星観測データを蓄積し、相互に利用するためにクラウドベースあるいは大型計算機システムを用いたプラットフォームが国内外で整備されている。また地球観測衛星データ、地上データなどのビッグデータから社会情報を抽出する紐付ける機械学習、ディープラーニングなどのIT技術開発については、産業技術総合研究所や民間などにより急速に進みつつある。

国内におけるプラットフォームのひとつとして前述のTellusがある。Tellusには政府の提供する衛星データだけでなく、民間が提供する地球・宇宙のデータも多数保持されている。これらはクラウド・GPU・ストレージなどのコンピューティングリソースを使い、オープンかつフリーでの利用が可能である。さらに機械学習やディープラーニング等により、防災・農業・水産業・インフラ管理等の様々な分野での気象情報のサービス化と新たな価値提供が進んでいる。

GAFA 等の取り組みとして、Google Earth Engine の事例が挙げられる。Google Earth Engine では、様々な衛星データ、気象データが収集されており、大量データの解析が行える仕組みを提供しており、海外では米国を中心として大学院生や若手研究者を中心に既に多くのユーザーが利用している。

- ・ FLUXNET2015 と温室効果ガスの推定における機械学習の導入

大気中温室効果ガス濃度を大きく左右する陸域の炭素循環については、渦相関法によるタワー観測ネットワークが展開されており、全球ネットワークとして FLUXNET、そのサブ組織としてアジア域では AsiaFlux が展開されている。これらの観測により大気―陸域間の熱・水・二酸化炭素等のフラックスが連続観測され、データ公開が進められている。FLUXNET については近年大幅なデータのアップデートがあり FLUXNET2015 としてユーザーが使いやすい形でデータが提供されている。さらに、これら地上観測ネットワークデータと衛星観測データなどに対して機械学習を行うことにより、地上観測ネットワークデータを広域化した全球プロダクトも構築され、コミュニティが利用できる形を提供している。

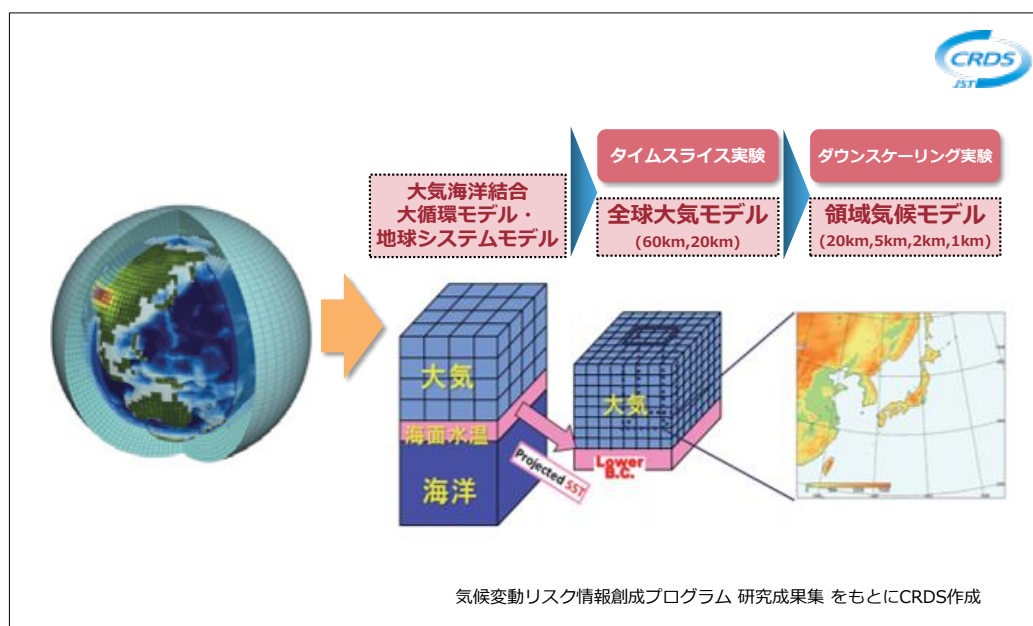
- ・ 十年規模変動予測

現在から先の 1 年以上、最大 10 年程度の時間スケールを対象とした予測研究が盛んになってきている。十年規模予測では、100 年程度の時間スケールを対象とした予測と異なり、自然変動による揺らぎの位相を正確に取り込む必要があるため、データ同化と呼ばれる手法を用いて特定の時点に対応する観測データになるべく近い初期値を作成して予測を始めることになる。データ同化は従来、天気予報のような比較的短期間の短い予測のために発達してきた手法であるが、温暖化予測に対する要請の高度化に伴いより長い時間スケールや、生物・化学過程などへの応用が進みつつある。例えばフランスの研究チームは、「パーフェクトモデルアプローチ」（モデル結果を観測に見立てて行う予測可能性検証実験）に基づいた研究により、地球規模の炭素循環に最大 6 年間の予測可能性があることを示している。国内では海洋研究開発機構が、データ同化により赤道太平洋域からの二酸化炭素フラックス変動を予測する研究などに取り組んでいる。

- ・ 温暖化ダウンスケーリング

気候変動が都市生活や住民の健康に及ぼす影響を予測・評価するためには、温暖化を適切に解析・予測できるシミュレーションモデルの開発が不可欠である。通常、全球気候モデルの空間解像度は 100 ～数 100km 程度であり、そのままでは都市生活や住民の健康への影響予測・評価に用いることができず、「温暖化ダウンスケーリング」と呼ばれる技術が用いられる。全球気候モデルの出力結果をより高解像度のシミュレーションモデルを用いて空間詳細化を施す力学的ダウンスケーリングという手法がある（他に、統計的なモデルを用いて高解像度化を行う統計的ダウンスケーリングという手法もある）。ダウンスケーリングには計算機資源の面から高解像度化に限界があり、計算機能力の向上が望まれているが、AI（ディープラーニング）を利用することで、膨大な計算機を使用せずに複数の予測が可能となる、新たな手法にも期待が高まっている。

なお、国内における地球温暖化予測モデルおよび力学的ダウンスケーリングに用いる領域モデルの開発は、主に、文部科学省「統合的気候モデル高度化研究プログラム」（2017-2022）において行われている。本プログラムの成果物としての予測データは、プログラム内での防災研究に用いられるほか、同省「気候変動適応技術社会実装プログラム」（SI-CAT）や、環境省「気候変動の緩和策と適応策の統合的戦略研究」（S-14）においても活用され、気候変動適応法に基づいた適応策立案に役立てられる。



・ 気象観測・予測における AI 技術の導入

気象庁では、AI 技術導入によって気象観測・予測の精度を大きく向上させることを目指し、理化学研究所革新知能統合研究センターと共同研究を行っている。AI 技術の活用による、気象観測データの品質管理手法や、全球モデル・メソモデル・局地モデル等の複数の数値予報結果や観測データを最適に組み合わせた統合型ガイダンスの開発の大きく 2 つに取り組んでいる。

また、甚大な被害をもたらす台風がここ数年、日本に年複数回上陸しており、台風予報のさらなる高度化が求められている。台風予報においては、強度変化のメカニズムが十分に解明されておらず、強度予測の改善が課題となっているが、気象研究所において米国の研究者の協力を得て、機械学習を導入した台風強度予測手法の開発が行われており、現在、試験的に台風予報に導入されている。

・ スーパーコンピュータ「富岳」開発と気象予測

2014 年から開始された文部科学省スーパーコンピュータ「富岳」開発事業のもと、「京」の後継機である「富岳」の開発・整備、および富岳を用いて重点的に取り組むべき社会的・科学的課題（重点課題）に向けたアプリケーション開発が開始されている。重点課題のひとつとして「観測ビッグデータを活用した気象と地球環境の予測の高度化」が設定されており、富岳では京と少し視点を変え、観測データの取り組みに重点を置いた研究開発が実施される。研究テーマのひとつとして、データ同化技術の高度化により、集中豪雨の発生をいち早く予測することを目指している。

(4) 課題と方向性

多種・多様な環境観測データが前提となる気候・気象の分野で、わが国として今、大きな課題と言えるのは、整備すべき観測データ項目の認識とデータ入手先についての戦略明確化、データ処理機能の強化である。これまで、地球観測衛星に関する物理観測パラメータは国連 World Meteorological Organization (WMO) 下に置かれた Global Climate Observation System (GCOS) によりリスト化された Essential Climate Variables (ECVs) を参考にしつつ、その充足を目指して計画立案や各々のミッション推進を実施してきた。しかしながら、わが国として気候変動観測につ

いてどのような観測データベースを整備すべきかについて、議論が活発には行われておらず、気候変動観測衛星についての戦略立案がここ数年停止している状態である。また、他国との協力によるデータ取得も含め、観測データベース整備について、オープンな議論の下、国としての方針を定めていく必要がある。データを扱うためのプラットフォームの整備、運用とともに、データサイエンティストの観測データへの習熟やアプリケーションの研究開発などが喫緊の課題であり、またアウトプットとしての社会情報の位置づけの検討についても課題である。

また、地球システムモデル開発の今後の大きな方向性として、高解像度化、アンサンブル数増大、多くのプロセスを取り込む複雑化の3つが挙げられる。高解像度化は、細かな空間スケールでの大気の大気対流活動や地形の変化が重要な役割を果たす降雨過程などについて、モデルの再現性を向上させるために必須である。また、100年や1000年に一度と言った、非常にまれな集中豪雨、熱波、渇水の発生確率の変化を検出するためには、従来高々100程度が普通であったアンサンブル数を1桁は増やす必要がある。さらに前述のSLCPsの取り扱いには大気中の化学反応などについて、これまでモデルに含まれていなかったプロセスも導入する必要がある。これらの方向性はいずれも計算負荷の増大を伴うことになるため、今後大型計算機の性能が順調に向上することがこの分野の発展にとって決定的に重要であると同時に、AIや機械学習といった手法をより積極的に活用する必要があると考えられる。

さらに、地球観測衛星によって毎日膨大な観測データが生産され、アーカイブ化されているが、観測データアーカイブから社会利益をもたらす情報化手法の研究の高度化にAI技術が果たす期待は大きい。目標とする事象が自然現象ではなく社会現象などの場合には、従来の論理的な物理的解析アルゴリズムの研究ではデータ量が多く対応できない。社会現象との相関そのものが社会的利益に直結するので、得られた情報と目標とする事象の相関解析では、ニューラルネットを用いる機械学習手法の実用化確立が課題である。

最後に、この分野における研究開発人材については、世代交代が必ずしも適切にできておらず急速な科学技術力低下が発生しつつあるとも言える。これまでに得られた知見を体系化し、若い世代を教育・啓蒙し、科学技術力を維持する実施体制や拠点の整備も大きな課題である。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 環境エネルギー分野（2019年）」（CRDS-FY2018-FR-01）（2019年3月）
- ・ IPCC, "Climate Change 2007: The Physical Science Basis," Working Group I contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2007.
- ・ 気象庁「IPCCのRCP8.5シナリオを用いた非静力学地域気候モデルによる日本の気候変化予測」『地球温暖化予測情報』9巻, 79, 2017.
- ・ Collins W. D., et al., "The Integrated Earth System Model Version 1: Formulation and Functionality," *Geoscientific Model Development* 8: 2203-2219, 2015.
- ・ 三村信男 監修、太田俊二、武若聡、亀井雅敏 編『気候変動適応策のデザイン』（2015）.
- ・ S. Hales et al. ed., *Quantitative risk assessment of the effects of climate change on selected causes of death, 2030s and 2050s*, (Geneva: World Health Organization, 2014).
- ・ 東京大学生産技術研究所 吉兼隆生「機械学習による局地気象予報の試み」（2018）The 32nd Annual Conference of the Japanese Society for Artificial Intelligence
<https://pdfs.semanticscholar.org/ec03/0673991bbee8b0d197e7619beb91c161ef15.pdf>

- FLUXNET Network, "Fluxdata,"
<http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/> (2020 年 3 月 1 日アクセス) .
- Boer George J., et al., "The Decadal Climate Prediction Project (DCPP) Contribution to CMIP6,"
Geoscientific Model Development, 9 : 3751-3777, 2016.
- Roland Séférian, et al., "Assessing the Decadal Predictability of Land and Ocean Carbon Uptake,"
Geophysical Research Letters, 45: 2455-2466. 2018.
- 国土交通省 気象庁、平成 31 年度報道発表資料「気象庁と理化学研究所（革新知能統合研究センター）との共同研究について」（プレスリリース、2019 年 1 月 23 日）
https://www.jma.go.jp/jma/press/1901/23a/20190123_ai_besshi.pdf
- 理化学研究所計算科学研究センター「気象予測とスパコン」
<https://www.r-ccs.riken.jp/r-ccssite/wp-content/uploads/2018/03/postk-kishou.pdf> (2020 年 3 月 1 日アクセス)

3.6 エネルギー

(1) 定義・意義

将来の大幅な CO₂ 削減に向けて、再生可能エネルギーの拡大、および利用面での電化拡大が潮流になっている。加えて IoT の潮流により、需要家の電力消費量、電力システム運用に関わる電流・電圧や設備情報などは、デジタル化が進むと想定される。複雑化する電力ネットワーク運用のためには、これらデジタルデータを用いて最適化するマネジメントシステム（GEMS：Grid Energy Management System）が必要となるが、個別の需要家においても太陽光発電（PV）、蓄電池（EV 含む）、エアコンなどの電力消費機器などの分散型エネルギー資源（DER: Distributed Energy Resource）を最適に運用するためのマネジメントシステム（BEMS/HEMS: Building/Home Energy Management System）が必要になってくる。このように分散して存在する EMS は個々に運用が最適化されるが、EMS 間でデータを共有し、全体として活用する仕組みがあればさらに大域的なエネルギーシステム全体での効率向上、温暖化ガス排出削減、設備のスリム化・稼働率向上等の効果が期待される。

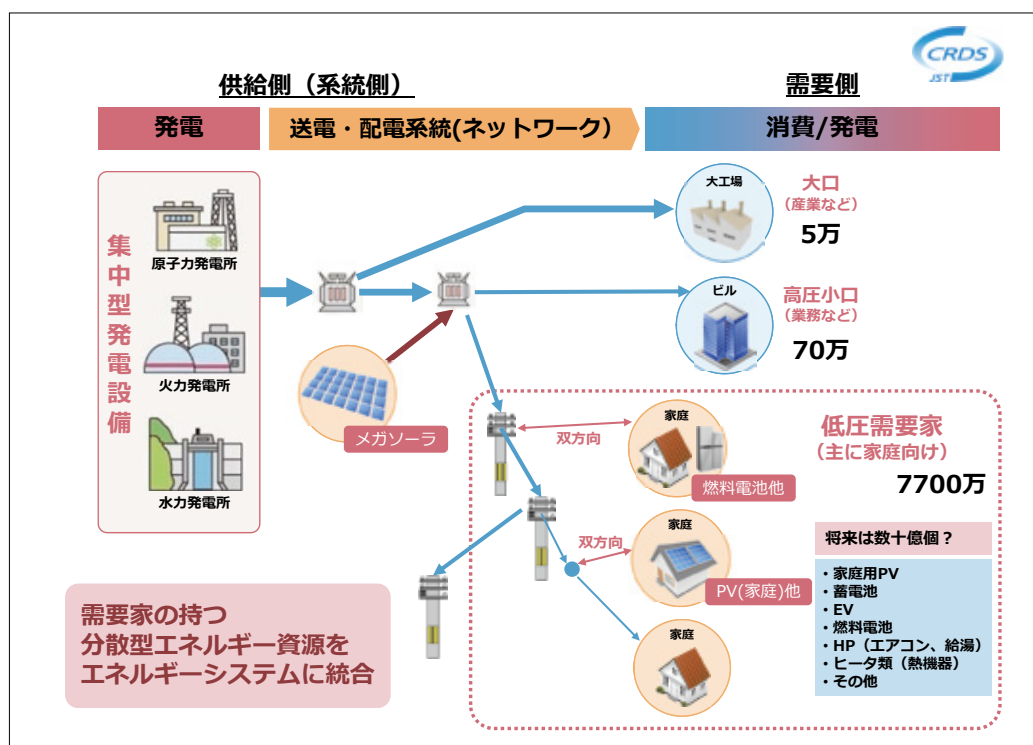


図 3-6-1 将来の電力ネットワーク

上記は日米欧等の先進国に共通した潮流であり、World Economic Forum(世界経済フォーラム)では、将来の三つのトレンド（①電化の進展、②配電系における DER の拡大、③デジタル化として Internet of Things (IoT) の出現）に対応するためにシステムレベルでの提案を行っている。これは生産者、送配電者、顧客といった従来の境界がぼやけ、システムガバナンスの複雑さが増加している中で、「グリッドエッジ」である一般家庭の顧客がビジネスに参加できるようなプラットフォームを構築し、グリッドエッジを対象とした技術を統合化することで分散型エネルギー資源の価値最大化を可能にすることが必要とされている。

(2) 研究のトレンド

歴史的に電力システムを振り返ると総括原価型電気料金が世界共通的に採用され、電力供給と料金の安定化に寄与した。一方、欧米では 1990 年代から自由化が進められ、市場取引への移行による競争原理の導入、効率化による料金低減が追及されてきた。現在はそれに加え、経済性の観点のみでは導入が進まない再生可能エネルギーをはじめとする低炭素技術や省エネルギー推進について、政策による導入誘導を図っている。例えばドイツでは 2000 年から、日本では 2012 年からの固定価格買取制度（FIT）の導入が開始され、コスト低下が著しい風力や PV といった変動性再生可能発電（VRG：Variable Renewable Generation）の導入が急速に進んでいる。特に 2010 年以降は PV のコスト低下もあり、急速に増加し、それに伴い、電力需給を「同時同量」でバランスさせる必要のある電力システムへの対策として、前述の需要側を巻き込んだ統合システム化が重要になってきている。なお、電力システムの安定化のためには広域化が有利とされているが、わが国は島国で国際連系線が無いという点で困難さが増す。また設備投資の不確実性が高まる電力市場の自由化が VRG の増加とほぼ同時であることが困難さをより一層増している。

電力ネットワークにおける技術開発や取り組みは PV や風力発電などの VRG や需要家が所有する DER をデータを活用した ICT 技術により、どのようにして電力システムの計画・運用に取り込み、変動の緩和・制御に貢献させ、市場に取り込んでいくかにある。また DER の主要機器である蓄電池の急速な価格低下が進行中であり、今後の拡大を考慮することも重要になる。このような状況を踏まえた上で、以下の研究開発のトレンドがある。

- ・ 住宅での PV 発電電力の自家消費促進： 例えば、バッテリーの設置、PV・電力需要との協調制御、このための HEMS、クラウドコントロールシステム等
- ・ 分散型エネルギー資源（DER/DR）の活用・統合制御による Flexibility の創出と系統運用サポート、および市場への統合： 例えば蓄電池によるアンシラリーサービス（需給バランスの監視、系統運用、電圧・周波数の調整など）、EV 充放電・給湯機による周波数制御・需要シフト（電力ピークマネジメント）
- ・ DER 統合の技術的検討： 需要家側にある DER/DR の統合制御のため高度な制御手法と情報通信技術・標準化の検討など
- ・ VRG の発電量の予測、気象データ活用、精度評価と計画・運用への活用
- ・ その他
 - ◇ 社会的に進展する IoT、Big Data と解析、電力エネルギー用 AI の開発・活用
 - ◇ 需要家の需要制御や保有する DER 制御に対する理解促進

(3) 国内外のトピックス

電力ネットワークにおけるさまざまなデジタルデータやネットワークとつながるリソースの活用に向けた検討が見られる。

- ・ 需要家の電力消費データの活用： 需要家の電力消費データは今後の電力システム運用や新たなサービスビジネスの展開において有用性が高い。世界的にスマートメータ²の導入が進められており、わが国でも 2024 年に全世帯の電力メータがスマートメータに置き換わる予定である。これにより全需要家の 30 分間隔での電力データが蓄積され、活用されることになる。またスマートメータの電流センサーデータ等を活用したデータ解析により需要家内の機器毎の電力消費パターンに分解する技術（Non-Intrusive Load Monitoring：NILM または Disaggregation と呼ばれる）の開発が進んでおり、需要内の個々の機器の使われ方を把握し、機器故障検知、

2 スマートメータ：電力使用量をデジタルで計測する、通信機能を搭載した電力量計

不在検知などのサービスへの活用が指向されている。

- 国内において大規模な VPP (Vartural Powoer Plant³) の実証事業が行われており、リソースの拡大とともに近未来に制度化される需給調整市場の各商品メニューに対応した VPP の制御特性、確実性等の検討が進められている。電気自動車 (EV) を新たな VPP リソースとして実証に取り込むことも実施。
- 電力取引へのブロックチェーン技術の活用
ブロックチェーン技術の特徴の一つであるデータ改ざん耐性を用いることで、電力供給と需要の個別情報をつなぎ、再エネ電力の認証技術につなげることなどが行われている。また将来期待されている PV 余剰電力などの近隣家庭間の直接電力取引 (P2P 取引) の検討も行われており、ブロックチェーン技術の一つであるスマートコントラクト (プログラムによる契約の自動化) を用いた検討も行われている。以下に P2P 取引事例を示す。
 - ▶ 米国：2017 年に米国ニューヨーク市のブルックリン地区で、ブロックチェーン技術に基づき P2P で再生可能エネルギーの電力取引を地域内で行う実証 (Brooklyn Microgrid プロジェクト) が始められた。LO3 Energy 社によって開発された電力取引システムとスマートメータを用いて実施。
 - ▶ 日本における実証試験
 - ◇ 浦和美園地区における実証事業 (環境省事業 H29-31：デジタルグリッド)：リアルタイム電力取引・制御システムであり、ブロックチェーン技術 (イーサリアム) におけるスマートコントラクトを用いることで電流制御を実施。
 - ◇ 余剰電力の P2P 取引実証試験例 (東大 / 日本ユニシス / 関西電力 / 三菱 UFJ との共同研究 (関電・異実験センター内)：太陽光発電設備が設置されたプロシューマー宅で発生した余剰電力を、電力の消費者とプロシューマーの希望価格から、各種方式 (オークション、ザラバ、ダイナミックプライシング) により取引価格を決定し、ブロックチェーンを用いて模擬的に取引を行い、複数の電力消費者宅へ送電。
- 米国「Transactive Energy」：Transactive Energy は、電力システムの特徴である集中型の制御により行われていた需給バランス、系統の安定性、信頼性の維持を、市場経済ベースの制御メカニズムを通じて行う新しいアプローチ。2004 年に DOE が検討組織として「GRIDWISE Architecture Council」を設立し、実証試験も含めて導入の検討。各ノードにおける電力価格 (+ 託送料) の情報を元に、自動的な電子商取引を行うことで、電力系統全体の安定化・最適化を図る。デモンストレーションのための実験も行われた。

3 VPP (仮想発電所)：工場、ビル、家庭の DER を IoT による高度なエネルギー管理技術で束ね (アグリゲーション)、遠隔・統合制御することで、電力の需給バランス調整に活用する仕組み。

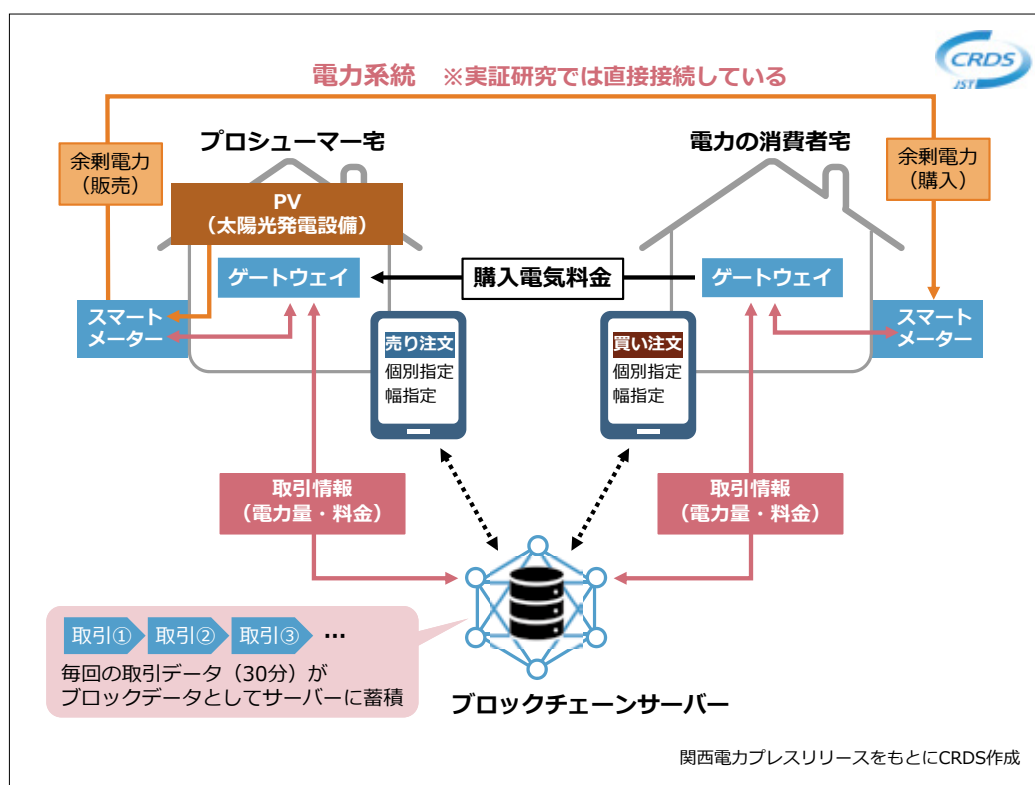


図 3-6-2 余剰電力の P2P 取引実証試験例（関電・実実験センター内）

(4) 課題と方向性

- VRG/DER 拡大に伴う電力系統との連携・協調、諸課題への対応：VGR/DER が大量に導入された電力システムで起こる諸現象の解明。特に慣性力低下、パワエレ機器の高度化、系統故障時に起こる過渡現象において安定供給に必要なとなる対策の開発、システムのモデリングの精緻化などが必要。
- 階層型監視制御アーキテクチャ、分散システム間協調：デジタル化の潮流において、各事業者・需要家はそれぞれの領域を最適化するシステムを保有する必要がある。このような階層型の監視制御の機能分担、必要なデータ種・粒度・交換タイミング、分散するシステム間での協調と目的とする最適化の実現に関する研究開発ニーズが高まっている。
- 小規模電力の需要家間取引（P2P 取引）：DER/DR の電力システムへの統合の視点が主軸になっている一方で、市場を介さない電力売買の仕組みの研究（P2P 取引、系統制約調整問題等）、取引データの管理とセキュリティの確保（Blockchain の活用、決済との連携等）が関心を集めている。
- 消費者行動分析、行動経済学的分析：エネルギー消費や自動車走行などのビッグデータを収集・活用した、消費者行動分析や行動経済学的分析を、プライバシー保護可能なオープンデータベースの開発も含めて実施する。また、消費と供給電力の動的な挙動を考慮した電力価格、インセンティブ設計手法などの行動経済学的研究も必要となる。
- IoT/ ビッグデータ / 人工知能の応用：時間粒度・空間粒度が異なる膨大なエネルギーデータ（Energy Granular Data）に対して、通信ネットワークの負荷低減を目的としたデータスリム化・スクリーニング・特徴抽出を実施する方法論の研究開発が必要。また、マルチスケールエネルギー

ギーリソースのアグリゲーションによる高速・高精度需給調整力の創出方法論に関する研究も有用になる。こうした課題に、IoT/ ビッグデータ / 人工知能という急速に進化している科学技術を取り込むことも今後の重要な方向性である。

- ・ オープンデータ整備：新たなデータ処理方法（data driven analytics）や制御手法の有効性検証に実データの利用が不可欠であるが、現在の国内外の実証試験は、既存技術の統合が中心で、経験的に性能向上を図っている段階である。ベンチマーキングに使えるオープンデータが整備されていないため、大量データ処理、リアルタイム情報通信制御などでの根本的な障壁があり、制御効果の確実性を確保する需要家行動など未解明である。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター 「研究開発の俯瞰報告書 環境・エネルギー分野 (2019 年)」 (CRDS-FY2018-FR-01) (2019 年 3 月)
- ・ 科学技術振興機構 研究開発戦略センター 「戦略プロポーザル 未来エネルギーネットワークの基盤技術とエネルギー需要科学 ～ 2050 年超の一般家庭でのエネルギー需給構造変化に向けて～」 (CRDS-FY2016-SP-02) (2017 年 3 月)
- ・ The Future of Electricity New Technologies Transforming the Grid Edge
<https://www.weforum.org/reports/the-future-of-electricity-new-technologies-transforming-the-grid-edge>
- ・ 関西電力プレスリリース「電力売買価格の決定を含むブロックチェーン技術を活用した電力直接取引の実証研究の開始について」 2018 年 10 月 15 日
https://www.kepc.co.jp/corporate/pr/2018/1015_1j.html

3.7 都市

(1) 定義・意義

1950年には25億人であった世界の人口は、その後急速に増大し、2000年には60億人を突破し、2020年には80億人に迫ると予想されている。一方、2018年には全人口の55%が都市に住み、2030年までには60%を超えると予想される。都市には産業が集積し、強い経済力がある。その結果、さらなる人口の集中が加速される。この集積が都市化のメリットである。しかし、人口や産業の過度の集中には、大気汚染や交通渋滞の発生、治安の悪化、住居のスラム化などさまざまな負の側面もある。

スマートシティの研究開発においては、これらのデメリットを補い、持続可能な経済社会システムである都市を実現しようとしている。IoT(Internet of Things:モノのインターネット)などの技術を活用して、エネルギーや交通などのライフラインの効率化、住環境の改善、災害への対応力向上などを目指している。

(2) 研究のトレンド

日本では、2010年頃に再生可能エネルギーを効率よく使い環境負荷を抑える次世代環境都市が、エネルギー環境問題を解決するソリューションとしてさまざまな実証実験で取り上げられ、スマートメータの導入などがなされた。関西文化学術研究都市や柏の葉スマートシティなどのモデル都市が街づくりを進めているが、活発な実証実験が終了した後は、スマートシティという言葉を目にする機会は減った。一方、米国、欧州ではスマートシティを中心とした動きがIoTなどの技術の進展、環境への意識の向上などにより再活発化している。社会基盤実現に向けた研究開発が必要なSDGsとの関連もあり、わが国においても、再び重要な取り組みになってきている。スマートシティに関連したプロジェクトをシーズ志向かニーズ志向か、利用者主導か行政主導か、という軸でまとめたものが図3-7-1である。

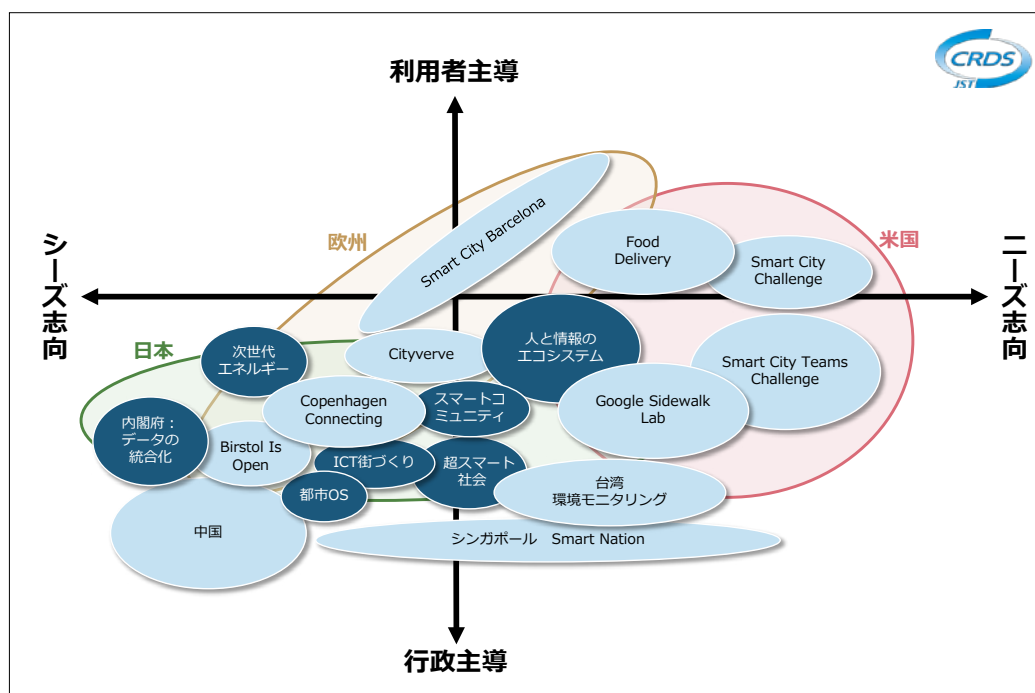


図 3-7-1 各国のスマートシティプロジェクトの現状

欧州においては技術導入を主目的とした政府によるトップダウンからスタートしたが、日本ではスマートシティに限らず、実証フェーズの施策を複数実施し、市民の課題解決を主目的とした市民によるボトムアップへとその活動がシフトしている。先端技術の導入だけを目的としたスマートシティの実現は実証実験レベルで終了することが多い。一方、継続してスマートシティに取り組む例では、例えばバルセロナで見られるように、市民のニーズを汲み取り、市民の課題解決を目的としたスマートシティ構築が目立っている。わが国においても、平成30年8月に国土交通省都市局が策定した「スマートシティの実現に向けて【中間とりまとめ】」に、今後目指すべきスマートシティのコンセプトとして、①技術オリエンテッドから課題オリエンテッドへ、②個別最適から全体最適へ、③公共主体から公民連携へとといった、バルセロナと同様な項目を挙げている。

関連する研究項目を列挙する。

- ・ 行動経済学における「ナッジ」による行動変容：ナッジはリチャード・セイラー（2017年ノーベル経済学賞）により提唱された。経済学では人の行動が完全に合理的で効用最大化を求める原理に基づくことを前提としているが、実際の人の行動は環境によって首尾一貫しない（限定合理性）ため、実際の経済が理論通りにならない。その一方でそれらの首尾一貫しない行動の中に一定の法則性のクセ（バイアス）があることが、行動経済学や心理学によりわかってきた。このクセを用いて人の行動を変容する方法として検討されている。わが国では環境省「低炭素型の行動変容を促す情報発信（ナッジ）による家庭等の自発的対策推進事業」においてナッジによる節電とその効果検証を実施中である。
- ・ 情報技術におけるセンサー・情報機器による行動認識と行動変容への適用：センサーが小型化し、スマートフォンやスマートウォッチのように常に持ち歩くことが可能になることで、動きだけではなく、生理的な状態（心拍や呼吸）なども取れるようになってきている。このためストレスや Quality of Life などの内面も計測可能になってきている。また行動変容においては、報酬等として、例えばポイント付与を利用した外発的動機付けや位置ゲーム（例：ポケモンGO）

のような好奇心などを用いた内発的動機付けが検討されている。

- ・ 計算社会科学：ビッグデータやシミュレーションを用いて、社会経済現象を研究する、社会科学、データサイエンス、計算科学の文理融合領域である。ビッグデータを利用することで、社会現象を定量的に分析したり、あるいはビッグデータに基づき導いた社会モデルを用いてシミュレーションすることで社会を分析することなどを可能にする。2015年に国際計算社会学会（IC2S2）が開催され、日本でも2016年に「計算社会科学研究会」、2017年には「計算社会科学研究センター」が立ち上がっている。なおビッグデータを用いた機械学習などの過去の行動をベースとする帰納的手法は環境変化に適応しながら変化していく人の行動を予測するには十分とは言えず、人の行動の裏にある動機等を含めた演繹的なモデル化との統合が必要になる。
- ・ 「実験社会科学」の展開：社会科学におけるヒトの特性について、実験（生物学、神経科学、工学などの融合）を通じて、エビデンスベースの実験社会科学の基盤と融合研究ネットワークを創出しようという分野が立ち上がっている。自然科学はヒトをモノとして扱い、社会科学はヒトを人格として扱うが、ヒトのモノ的な基礎を認識することが社会科学にとって重要との認識に立つものである。

(3) 国内外のトピックス

- ・ NSF Smart & Connected Communities：米国 NSF(National Science Foundation) が2016年から実施しているプログラム。コミュニティは、急速に変化するインテリジェント技術によって居住者とその周辺環境が相互に接続され、新しい変革の時代に入りつつある。この変革は、福祉と繁栄を大いに改善すると期待されるが、一方で技術と社会の関りにおいて課題も提起する。NSF スマートアンドコネクティッドコミュニティ（S & CC）プログラムでは、スマートで接続されたコミュニティが新しいレベルの経済的機会と成長、安全とセキュリティ、健康とウェルネスおよび全体的な生活の質をもたらすことを可能にする科学および工学的基盤を生み出すことを目的としている。
- ・ Side Walk プロジェクト：2015年6月に、Alphabet 社は新会社「Sidewalk Labs」の設立を発表した。Sidewalk Labs は、人々を中心とした都市デザインと先端技術を組み合わせることで、持続性があり、手頃な価格で暮らせ、機動性があり、経済的機会が均等な都市を実現することをミッションとしている。2017年にSidewalk Labsは、カナダが進めるトロント東部ウォーターフロント地区に新たなコミュニティを作るプロジェクトに参画することを発表した。都市の抱えるエネルギー、ゴミ処理、健康管理、住居不足、交通といった課題を技術で解決して、住民、労働者、観光客の Quality of Life を向上させる計画である。そのために、都市プラットフォームを構築して、都市で生み出されるさまざまなデータを利用するとしている。
- ・ スマートメータからのデータ活用：スマートメータの情報の利活用の動きは日欧米で検討されているが、プライバシー保護や情報管理の問題が絡むため、まずはその法整備や仕組みづくりの検討がされている。わが国ではスマートメータを含むデータ利活用の仕組みとして、公的機関の認定を得た団体として「情報銀行」⁴でデータを扱うことを条件にその利用が認められる見通しになっており、今後の取り組みが一気に進むことが予想されている。既にデータの活用を検討する組織も立ち上がっており、防災計画、商圈分析、新サービスなどへの活用検討が行わ

4 情報銀行：個人とのデータ活用に関する契約等に基づき、個人情報管理・蓄積するシステム（PDS）等を活用してデータを管理するとともに、個人の指示又は予め指定した条件に基づき、個人に代わり妥当性を判断の上、データを第三者(他の事業者)に提供する事業

れている。またさまざまな産業でも電力データの活用による人の行動予測を基にした新たなサービスや付加価値創造が期待されている。(図 3-7-2)

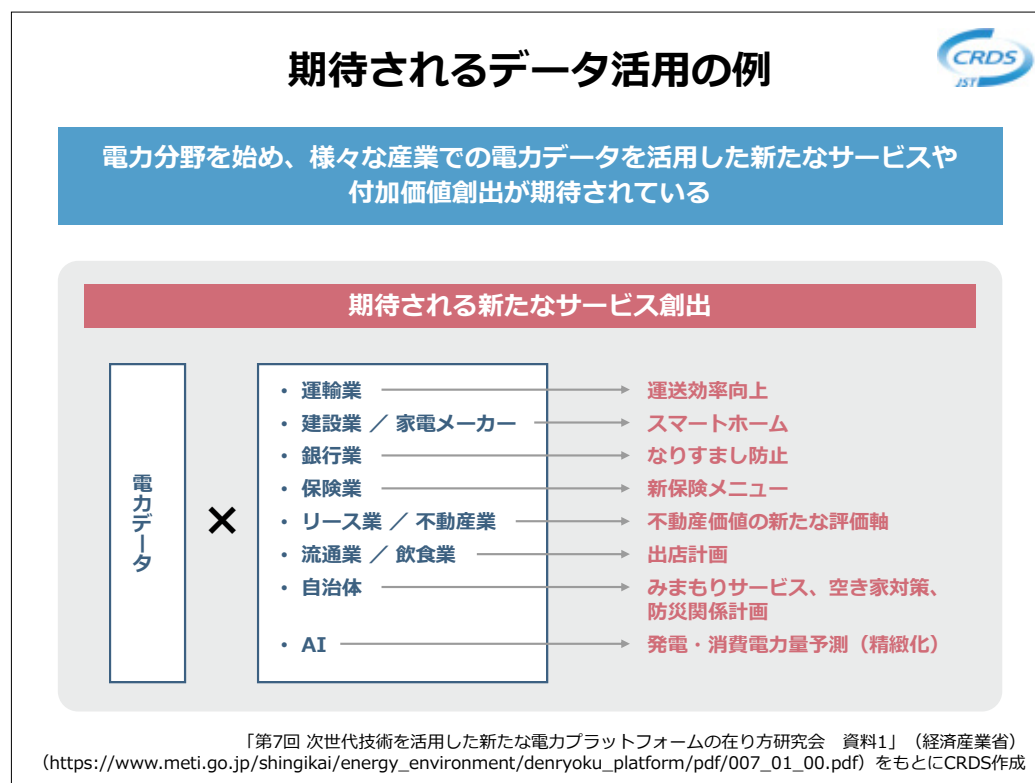


図 3-7-2 電力データの期待される活用例

- ・ 分野横断的なデータ流通を可能にするデータ管理基盤 FIWARE: これまではスマートシティサービスを実現しようとする、交通サービスの高度化には車やバス、電車の運行データを使い、また、防災のためには防災カメラや河川の水位センサーのデータを使い、電力の制御には電力データを使うなど、分野毎に個別のデータを活用し、分野毎に個別にシステムを構築していた。その結果、分野やサービスを横断したデータの活用や連携が困難であり、都市としての統合されたサービスが提供されにくかった。FIWAREは分野横断的なデータ流通を実現するデータ管理基盤であり、IoT プラットフォームとして欧州委員会の官民連携プログラムで開発された。これにより、交通データと防災データを組み合わせ、高度な災害対応サービスが実現されたり、気象データと電力データを統合的に解析することによって、正確な電力需要予測が可能になる。欧州を中心に都市のさまざまなデータの連携・利活用を推進し、行政の効率化、新たなビジネスの創出、市民の生活向上などに向けた取り組みが行われている。
- ・ サイバーフィジカル&ヒューマンシステム (CPHS): 国際自動制御連盟 (IFAC) および IEEE 制御システム学会(CSS)が2014年に共催ワークショップを開始した。これまでのサイバーフィジカルシステム(CPS)に人間をシステムの一部として取り込む CPHS の考え方が始まっている。わが国でも計測自動制御学会において CPHS 調査研究会が発足している。
- ・ FuturICT2.0 プロジェクト: EU のフラッグシップ概念実証プロジェクト (FLAG-ERA JTC2016) として2017年に開始。ICTからのデータを利用して、社会経済システムのモデルを開発し、web ベースの実験とシミュレーションにより社会科学の領域での大規模なシミュレーションの

基礎を提供するとともに Finance 4.0 と呼ばれる経済モデルのデモを提供することを目指している。

(4) 課題と方向性

このようにデータ活用が社会システム構築・改善に大きな役割を果たすために必要な研究開発課題としては、空間的・時間的に異なるさまざまなデータを構造化していくデータサイエンスの研究開発に加え、社会システムの一部として関与する人の行動に係る研究領域が挙げられる。すなわち後者については、「人の行動のモデル化」が必要であり、行動予測のみならず行動変容の方策に関する行動科学の研究も必要である。

これはインフラ系設備（通信網、道路網、送配電網など）を用いるシステムでは、通信量、交通量、電力容量などに物理的上限を持つことから、交通であれば、渋滞緩和のために経路を分散させること、電力であれば最大需要を下げるために電力需要を平滑化することなど、個人の満足と社会システムの最適化を両立することが求められ、このためには人の行動変容を促す方策（情報や金銭的インセンティブなど）が必要となるためである。このような領域は自然科学のみならず、ヒトに関する研究の歴史がある人文・社会科学、例えば行動経済学、社会心理学などとの学際的な分野融合研究が必須となる。

- ・ 自然科学と人文・社会科学の融合：人の行動を理解し、行動変容を促すことはエネルギー関連のみならず、人が関与するすべての社会システムにおいて重要である。今後このような人の行動に係る分野を推進するためには、数理モデルやデータを定量的に扱う行動経済学が、情報科学と接合していく形で融合することが親和性の点で良いと考えられる。ただし行動（Behavior）そのものに着目する自然科学に対して、人文・社会科学では行為（Action：意味、価値、理由、主体性、人間性など）に着目しており、研究の目的の違いを理解した上での融合研究のあり方を検討する必要がある。
- ・ オープンデータ整備：新たなデータ処理方法（data driven analytics）や制御手法の有効性検証に実データの利用が不可欠であるが、現在の国内外の実証試験は、既存技術の統合が中心で、経験的に性能向上を図っている段階である。ベンチマーキングに使えるオープンデータが整備されていないため、大量データ処理、リアルタイム情報通信制御などでの根本的な障壁があり、制御効果の確実性を確保する需要家行動など未解明である。

(5) 参考資料

- ・ 国際連合経済社会局 <https://population.un.org/wpp/>
- ・ United Nations Department of Economic and Social Affairs, “The World Cities in 2018”, https://www.un.org/en/events/citiesday/assets/pdf/the_worlds_cities_in_2018_data_booklet.pdf
- ・ Ajuntament de Barcelona “Barcelona Digital City” <https://ajuntament.barcelona.cat/digital/en>
- ・ 荒川豊、総務省勉強会「行動認識と行動介入による情報駆動型社会システムの実現に向けて」
- ・ 科学技術振興機構 研究開発戦略センター 「戦略プロポーザル 進化的社会システムデザイン ～自然科学と社会科学の連携協調による持続可能な社会の実現～」(CRDS-FY2019-SP-01) (2019年7月)
- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 システム・情報科学技術分野(2019年)」(CRDS-FY2018-FR-02) (2019年3月)

- National Science Foundation, “Smart and Connected Communities”
https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505364
- Sidewalk Infrastructure Partners, “Sidewalk Labs is reimagining cities to improve quality of life,”
<https://www.sidewalklabs.com/>
- 「家庭の電力使用データ、暮らしの分析・災害対応に活用」『日本経済新聞』2019年11月19日
- 亀井亜希子「2019年度、情報銀行が本格開業へ」『大和総研レポート』2019年6月25日
- 経済産業省 第7回 次世代技術を活用した新たな電力プラットフォームの在り方研究会 資料2 『グリッドデータバンク・ラボのご紹介 - スマートメータのデータ活用とその可能性』
- 経済産業省 第7回次世代技術を活用した新たな電力プラットフォームの在り方研究会 資料1 『電力データの活用』
- European Commission, “The Future Internet platform FIWARE” ,
<https://ec.europa.eu/digital-single-market/en/future-internet-public-private-partnership>
- 計測自動制御学会制御部門「Cyber-Physical & Human Systems 調査研究会」
<http://is.eei.eng.osaka-u.ac.jp/hatanaka/CPHS/index.php>

3.8 政策立案

(1) 定義・意義

一般に、「政策」とは社会がより良くなることを目指して、さまざまな課題を解決するための国の方針・方策・構想・計画などの総称⁵である。デジタルトランスフォーメーションの流れの中で、政策そのものの研究がどのように変化しているかについて説明する。ここでは特に科学技術、イノベーションに係わる政策を対象にして、各種の大量データから社会動向を分析し、合理的な方法で政策立案や政策評価をおこなうための手法やシステムについて、その動向を示す。

(2) 研究のトレンド

図 3-8 に政策形成から実施のプロセスの一例を掲げる。まず社会や自然から課題を抽出し、それを分析する。次にその課題を解決する方法をさまざまな観点から検討し、複数の政策案（政策オプション）を作る。その中から最も良い評価を得た政策オプションを実行に移し、その結果、社会や自然に作用した成果を測る。これを繰り返すことにより、国の発展を進める。

このようなプロセスにおいて、さまざまな測定、分析、評価の方法が必要であるが、確固としたものはまだ存在しない。しかし、今やリアルタイムにビッグデータを収集、分析することによって、政策が対象とする複雑な社会や自然の動きの実態を総体として把握することにも可能性が見え始めた。

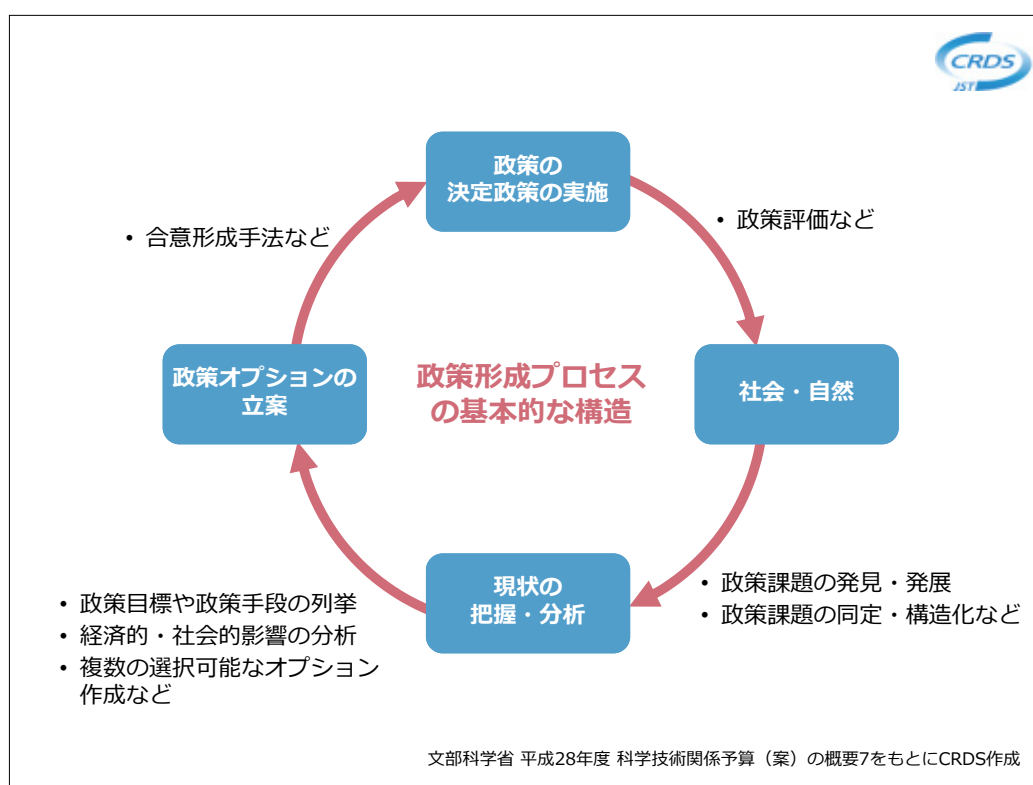


図 3-8 政策形成プロセス例

5 政策の具体的手段として通常用いられるのは法規制、徴税、資金供与、国民へのサービス等である。

上に挙げた政策のプロセスに沿って、それぞれ研究要素を説明する。

①データの収集

社会、自然界の現象をなるべくリアルタイムに知ることが必要である。科学技術に係る範囲では、研究への投資の状況、研究成果である研究論文や特許の公表数、研究成果を利用した製品やサービスの売上額等が対象といえる。研究機関や研究者を対象とした統計調査データ、研究論文の書誌データベース、特許データベース等がそれらのデータ・ソースとなる。

この収集段階で問題となるのは、収集データの種類と質である。政策のために必要十分なデータが常に揃うわけではなく、別々に存在するデータベースを組み合わせる、あるいは新規に調査収集する等の作業が必要になる。さらにすでに収集されたデータが正しいか否かを確認する必要も生じる。このような収集にかかるコストや技術的難しさを考慮することが求められる。

②データの分析

政策上の課題を浮かび上がらせるために、どのようにデータを組み合わせ、論理を展開するかを設計する必要がある。ここでは、用いるデータが目的にあっていないか、正しく解釈されているか、が検討点である。場合によっては、同じデータからいくつもの異なる解釈が生じる可能性がある。そのような多義性を解消するために新たなデータを追加するか、または複数の解釈をそのまま後段の政策形成プロセスに持ち込むか、を考える必要がある。

③政策オプションの作成

抽出された政策上の課題を具体的な政策の形にしなければならない。政策を立案する場合は複数の実施案を検討し、その中から最適と考えられるものを選択することになる。

政策は将来に向けて実施策を立てるものであるため、手元にある過去データにもとづいて、政策の実施案によって生じる効果を「予測」する必要がある。そこでシミュレーション技術が求められる。複数の実施案に対応して、シミュレーションに与えるさまざまなパラメータの組み合わせを変えて結果を予測し、比較することになる。これら複数の実施案と予測結果の組が政策オプションである。

最終的に、政策立案の責任者は政策オプションをさまざまな角度から比較し、日本の将来にとって望ましく、かつ現実に実行できる案を採択することになる。

④政策の評価

いわゆる PDCA サイクルの Check に該当するもので、政策を実施した結果、どのような成果が生まれたか、目標にどれほど近づいたか、何か新たな問題が生じたか等を調べる。

政策は社会を対象とするために、施策と結果の間の因果関係を正確に把握することが本質的に難しい⁶ことから、これまで客観的な評価はほとんど行われてこなかった。しかし最近では政策効果についても「エビデンス」による説明が重視されるようになり、政策評価も指標中心になりつつある⁷。

6 アクション・リサーチと同様に、政策が実施されると、社会のシステムの系に影響を与えてしまうため、当初の想定モデルから外れるという問題がある。他方、自然を対象とする施策の場合は科学的検証がある程度可能であることが多い（例：建設・土木分野における安全基準等）。

7 たとえば第5期科学技術基本計画（2016～20年度）では、期初に若手の大学教員数の増加、トップ10%論文数の増加等の計8つの目標を設定し、それを測る主要指標21個、さらに政策分野ごとに状況を把握するための細かい第2レイヤー指標群を設定して、科学技術に係る約1,200事業の全体効果をさまざまな角度から見ようとしている。

(3) 国内外のトピックス

・データベース・ネットワーク

科学技術政策に関して総合的な分析ができるような仕組み作りについて、下記のような海外事例を挙げておく。これらはもっぱらデータの収集から分析まで（政策プロセスの①②）の範囲をカバーするものといえる。

STAR METRICS (Science and Technology for America's Reinvestment: Measuring the Effects of Research on Innovation, Competitiveness and Science) と UMETRICS (Universities: Measuring the Impacts of Research on Innovation, Competitiveness, and Science)

米国で2007年から「科学イノベーション政策の科学」(Science of Science Innovation Policy:SciSIP)事業が開始された。これは政策担当者をサポートするために必要なデータ、ツール、方法論を生み出す実践コミュニティの構築を目指したものである。

SciSIPの研究の基礎となるデータ・情報基盤としてSTAR METRICSが整備された。これは、大学に所属する研究者ごとの公的研究資金、研究成果、支出に関わるデータをSTAR METRICSに参加する大学間(米国中西部の大学連合(CIC)等、約100校)で統合し、さらに省庁が保有するデータとリンクさせるものである。これにより、公的研究開発投資の大学研究への影響評価が可能になり、SciSIPの採択研究に広がりをもたらされることになった。

2012年から開始されたSTAR METRICS Level2では、広範な研究開発投資の効果を測るべく、より多くの成果(特許、論文引用数、研究上の成果等)のデータと研究開発投資のデータを結び付ける試みが行われている。しかし、ツールとしての需要(行政上必要な検索、広報用の視覚的資料作成等)と、研究のための分析データのニーズとの乖離が大きくなったため、2015年に前者はNIH主導のレベルII/Federal RePORTERに、後者はUMETRICSに分離した。UMETRICSは省庁の支援を受けず、民間財団によって維持されている。

RISIS (Research infrastructure for science and innovation policy studies)

欧州における科学イノベーション政策研究のための研究インフラ RISIS は、既存データ基盤のネットワーク化や共同研究の推進により新しい視点を与えられたデータ基盤の構築を目的としている。ファンディング、人材、論文・特許等に関する12のデータベースと、それらを整合・連携させるシステムから構築されている。

第1期RISISは欧州連合(EU)のFramework Programme 7 (FP7)の枠組みの中で、25個の活動プログラムを実施した(2014～17年)。参加団体は9カ国13大学・研究所。現在、第2期RISIS(RISIS2)に移行し、Horizon2020の中で10個の活動プログラムを実行している(2019～22年)。参加18機関。行政担当者との対話や研修を頻繁に行っていて、研究者と行政とのコミュニケーションを促進している点が特徴である。

日本のデータベース・ネットワーク

「(5) 国内外の政策動向」で述べたように、日本の国や自治体を持つ公的データはかなりオープン化が進んでおり、政策分析の一次ソースとして利用可能である。しかしいわゆる5つ星モデルから見ると、機械判読可能なLODに到達しているものはまだ少ない。

日本においては、データ提供機関が国に集中していて(DATA.GO.JPやRESAS等)、大学や民間の機関同士がネットワークを形成していない点が上記の海外事例と比較して遅れているといえる。しかし民間や地域を持つビッグデータと、公的データを組み合わせようとする試みは徐々に進んでいる。

また、2018 年から内閣府は「エビデンスに基づく政策立案」を推進するために、国立大学法人・研究開発法人等の経営データを統合するための標準化に乗り出した。2019 年度内にこれらの法人からデータ収集を始め、2020 年度以降は研究成果データ（論文・特許等）ともリンクして、大学の研究力を総合的に把握する計画である。これは政策プロセス④の評価を支援するための一つのデータ基盤になるといえる。

・個別技術

政策プロセスの①～④に係る個別の技術をいくつか挙げておく。これらはまだ調査段階ないし実験段階で、小規模のものであるが、政策形成の今後のキー技術に発展する可能性がある。

「JUDGIT！」（行政事業レビューシート検索システム）

政策プロセス①②を支援する。各府省庁の毎年度の行政事業レビューシートをデータベース化して、キーワード検索可能としたもの。2014 年度以降、約 5,000 件 / 年ずつ集積されている。たとえば“災害に対する補助”という言葉で事業を横断的に探することができる。

「議事録分析」（NISTEP）、「国会答弁自動作成システム」（経産省）

政策プロセス①～③まで支援する。省庁において蓄積された各種の委員会議事録、国会質問主意書・答弁書は膨大な量があり、かつ年々増加している。過去の記録の中から、類似した案件を参照し、政策形成の過程を分析したり、過去事例をもとに適切な文書を効率的に用意することを狙っている。類似検索や文書作成支援のために AI 技術を活用する。

「経済シミュレーション」

政策プロセス③を支援する。政策が国の経済の将来に及ぼす影響をシミュレーションする。これまでに、下表のようなさまざまな手法が提案されている。

表 3-8 経済シミュレーションモデルの例

＜一般的なモデル＞	経済財政モデル(内閣府)、Q-JEM(日銀)、短期日本経済マクロ計量モデル(内閣府)、MULTIMOD Mark III (IMF)、FRB Global(FRB)
＜科学技術の政策や進歩を考慮したもの＞	NISTEP 従来モデル(NISTEP)、NEMESIS(EU)、MULTIMOD Mark III 拡張版(IMF)、R&D 動学一般均衡 (DGE)モデル(一橋大)、多部門経済一般均衡的相互依存モデル(GRIPS)、MaeSTIP(NISTEP)

「社会シミュレーション」（日立京大ラボ）

政策プロセス③を支援する。あらかじめ有識者が考察した社会的要因とそれらの因果関係モデルをもとに、確率的シミュレーションと AI 技術によって、長期にわたる社会変化の予測をおこなう。大きな分岐点に至る前に必要となる諸政策の気づきが得られる。

（4）課題と方向性

これまで述べたように、政策の分野でデータを駆使して本格的に研究を進めている例はまだない。海外事例を含めて、ようやくデータベースの基盤が作られ始めている段階である。

現時点で今後の発展のために必要と考えられるポイントをまとめる。

(ア) データ基盤の整備

政策の立案には多種多様なデータが求められ、それらが即時に利用できる環境が望ましい。日本では法律で定められた基幹統計含めて多くの統計データが存在し、定期的な更新を経て、DATA.GO.JPを通して公開されている。しかしその公開レベルはまだ低く、またデータベース間を連携する仕組みも存在しない。データベースを再利用可能な形にすることが急がれる。

(イ) 研究コミュニティの拡大

STARMETRICS や RISIS の例に見られるように、海外では大学や研究機関がそれぞれ保有している独自のデータベースを積極的に公開し、相互利用をはかるコミュニティが存在し、政策研究のためのデータ基盤を維持している。日本においても国のデータ整備事業に頼るだけでなく、研究者が自ら動いてデータベースを中心としたコミュニティを作ることが必要ではないか。

(ウ) 政策実務者と研究者の協働

政策のプロセスを実行するのは主として府省庁や自治体の実務者である。その政策プロセスを調査・分析して、よりよいプロセスのあり方を提案するのは研究者である。(さらにステークホルダーである国民がいる。) 実務者と研究者(とさらに国民)が相互の立場から意見交換し、データをよりどころとして政策立案の質を高めていく必要があるが、現実には両者の距離はまだ遠い。その間を埋める試み⁸も一部で始まっているが、さらに広範に協働の輪を作っていくことが望ましい。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター 『調査報告書 米国「科学イノベーション政策のための科学」の動向と分析」(CRDS-FY2015-RR-04) 2015 年 11 月
- ・ 科学技術振興機構 研究開発戦略センター 「調査報告書 欧州における科学技術イノベーションネットワーク形成とデータ連結・拡張の動向」(CRDS-FY2016-RR-04) (2017 年 3 月)
- ・ 科学技術振興機構 研究開発戦略センター 『調査報告書 科学技術イノベーション政策の科学における政策オプションの作成 ～ ICT 分野の政策オプション作成プロセス ～」(CRDS-FY2015-RR-07) (2016 年 3 月)
- ・ 曾根原登「地域ソーシャル・ビッグデータ連携基盤とデータ駆動政策決定支援システム事例」、総務省情報通信審議会・先端技術 WG 第 5 回資料、2016 年 4 月 22 日
- ・ 内閣府「エビデンスに基づく政策立案」
<https://www8.cao.go.jp/cstp/evidence/> (2020 年 3 月 1 日アクセス)
- ・ 小柴 等、森川 想「議事録を用いた我が国における議会・行政の関係性分析手法」、人工知能学会論文誌、34 巻 5 号 p.E-J47_1-10、2019 年 9 月
- ・ 経済産業省：「国会会議録等の分析高度化への人工知能利活用の可能性検証に関する調査研究」調査報告書(平成 29 年度電子経済産業省構築事業)、2018 年 3 月
- ・ 株式会社日立製作所、ニュースリリース「AI の活用により、持続可能な日本の未来に向けた政策を提言」2017 年 9 月 5 日
<http://www.hitachi.co.jp/New/cnews/month/2017/09/0905.html> (2020 年 3 月 1 日アクセス)
- ・ 黒木 優太郎、横尾 淑子、科学技術・学術政策研究所、STI Horizon「シリーズ ―未来を創る― 日立京大ラボの描く未来」

8 文科省「科学技術イノベーション政策における『政策のための科学』推進事業」では、2019年度から「共進化実現プロジェクト」を開始している。これは行政側の担当部署と大学研究者が直接対話して課題を設定し、政策上での解決策を探索するものである。現在、人材育成やイノベーションに関する9課題を研究中。<https://scirex.grips.ac.jp/project/> (2019年11月1日閲覧)

<https://www.nistep.go.jp/wp/wp-content/uploads/NISTEP-STIH5-2-00174.pdf> (2020 年 3 月 1 日アクセス)

【参考】日本のオープンデータ

以下に日本国内の代表的なオープンデータをいくつか掲げる。これらはすべて国が公開しているものである。

DATA.GO.JP

[URL] <http://www.data.go.jp/> (2020 年 3 月 1 日閲覧)

[概略] 二次利用が可能な公共データの案内・横断的検索を目的としたオープンデータの「データカタログサイト」。政府を代表するオープンデータサイトであり、内容は充実している。今後、データ登録の負担軽減やオープンデータの一元的な提供という観点から、DATA.GO.JP と行政情報の総合的なポータルサイトである e-Gov との統合を 2022 年に実現できるよう計画調整が進められている。

[企画・立案] 内閣官房情報通信技術 (IT) 総合戦略室

[運用] 総務省行政管理局

[公開ポリシー] 基本的には CC-BY で運用されている。

[データ規模] 2019 年 10 月末時点で、データセット数：25,001、提供省庁数：22。

政府統計の総合窓口 e-Stat

[URL] <http://www.e-stat.go.jp/> (2020 年 3 月 1 日閲覧)

[概略] 政府統計ポータルサイト

[企画] 総務省統計局

[運用] 独立行政法人統計センター

[データ規模]

- 2019 年 10 月末時点で 609 種の統計調査データが掲載されている (データセット 1,554,946 件)。最も基本的な統計である基幹統計 (55 種) の他にも、各省庁の独自統計が含まれている。
- データ形式の多くは XLS や PDF 形式である。
- 主な統計データについては XML、JSON、CSV で出力できる API を用意している他、重要な統計データ 14 種についてはオープンデータの 5 レベル (最高) である LOD (Linked Open Data) でも利用できるようにしている。

NISTEP「政策のための科学」データ・情報基盤

[URL] <http://www.nistep.go.jp/research/scisip/data-and-information-infrastructure> (2020 年 3 月 1 日閲覧)

[概略] エビデンスに基づく科学技術イノベーション政策の推進のための体系的なデータ・情報基盤。この中の「データ・情報基盤リンク集」には国内外の科学技術イノベーション関連のリンクがおかれている。(http://www.nistep.go.jp/research/scisip/data-and-information-infrastructure/datalink_index)

[企画・運用] 文部科学省科学技術・学術政策研究所 (NISTEP)

地域経済分析システム RESAS

[URL] <https://resas.go.jp/> (2020 年 3 月 1 日閲覧)

[概略] オープンデータのわかりやすい応用例の一つ。産業データや人口データなど地域に係す

るデータを地図情報と組み合わせて、視覚的でわかりやすく提供するサービス。RESAS-API を通して加工可能なデータを公開している。利用者コミュニティや出前授業、ハッカソンなどを通じて普及や応用拡大をはかっている。

〔企画・運用〕内閣官房 まち・ひと・しごと創生本部

・日本の公的オープンデータの課題

政府の IT 統合戦略室では毎年、全府省庁が保有している統計データの棚卸を実施して、オープンデータ化がどの程度進められているかを公表している。上記のように政策は積極的に推進されているが、現実にはまだ道半ばという状況である。

たとえば最新の棚卸（2018 年 3 月時点）では、公開データのうち、5 つ星モデルの第 2 段階に相当する構造化 PDF 形式（約 29%）と Excel 形式（約 19%）で半数近くを占める。第 4 段階以上の高度な段階のものは 10% にとどまっている。今後は非公開の理由を具体的に分析することによって、公開に向けた検討を継続する必要がある。また公開データのオープン化レベルをさらに高めて、機械判読可能としていくことも求められる。

4 デジタルトランスフォーメーションの基盤技術

4.1 データ基盤フレームワーク・システム

(1) 定義・意義

ビッグデータとAI技術の発展・普及によって、「データに基づく問題解決」の枠組みが幅広く活用されるようになった。この枠組みは基本的に、①データ収集・蓄積ステップ、②データ分析ステップ、③アクション実行ステップ、で構成される(図4-1)。これをここでは「問題解決パイプライン」と呼び、このフレームワークについて述べる。

問題解決パイプラインの技術的発展は、図4-1に示すように、(a)より深い分析自動化へ、(b)一気通貫の自動化へ、(c)サイバー空間から実世界へ拡張、という3方向へ発展してきた。また、この問題解決パイプラインを支えるデータ基盤の面で、データ処理基盤技術、データ保護技術、オープンデータ技術等の研究開発も進められている。

このようなデータに基づく問題解決の発展によって、実世界で起きる現象・活動の状況の精緻かつリアルタイムの把握・予測、膨大な選択肢の網羅的検証、大規模複雑タスクの自動実行等が可能になり、問題解決手段の質的变化、科学の方法論の革新、産業構造・社会システムの転換等がもたらされ得る。

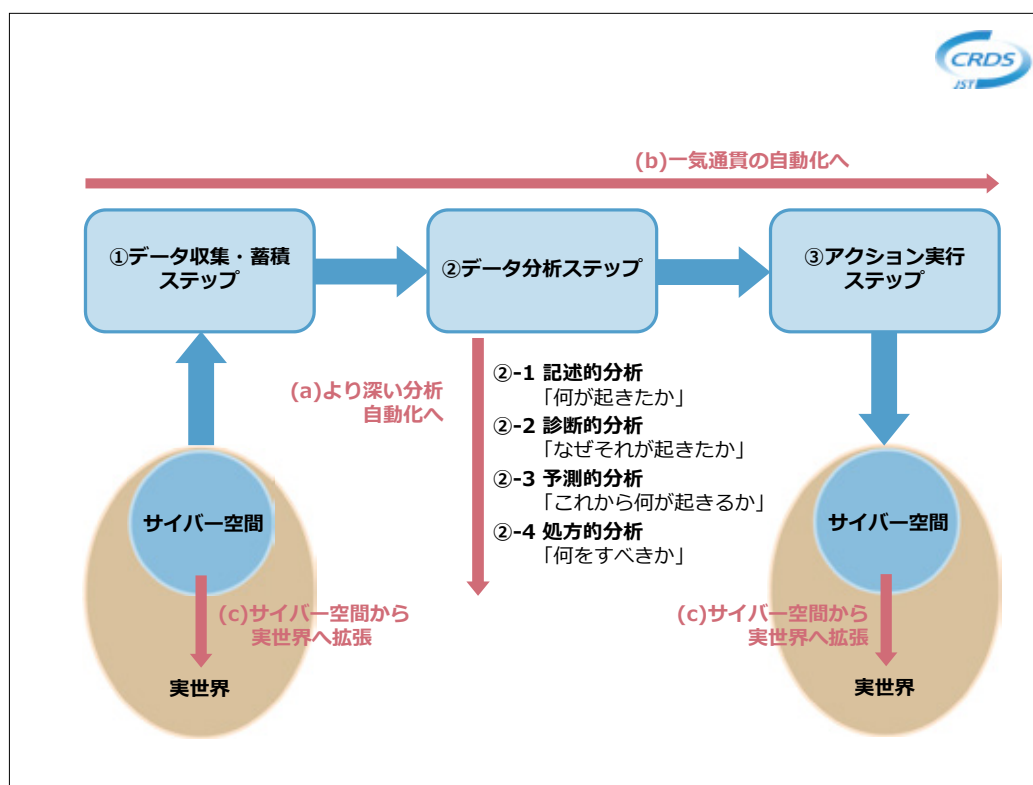


図4-1 問題解決パイプラインとその技術的発展の方向

(2) 研究のトレンド

上で述べたように、問題解決パイプラインの技術的発展は、(a)より深い分析自動化へ、(b)一気通貫の自動化へ、(c)サイバー空間から実世界へ拡張、という3方向に進んでいる。

- (a) より深い分析自動化へ：②データ分析ステップは、さらに、データ分析の深さによって段階がある。米国のGartner社は、データ分析の段階を、②-1) 記述的分析(Descriptive：何が起きたか)、②-2) 診断的分析(Diagnostic：なぜそれが起きたか)、②-3) 予測的分析(Predictive：これから何が起きるか)、②-4) 処方的分析(Prescriptive：何をすべきか)という4段階で深まるとしている。②-4)によってアクションが計画され、③アクション実行が可能になる。このような技術的発展は、主に機械学習技術の発展によって実現されてきた。特にディープラーニングによるパターン認識や判別・分類・予測の高精度化は、②-1)、②-2)、②-3)に大きく貢献した。②-4)に対しては、さらに深層強化学習技術、予測型意思決定最適化技術、知識ベース推論技術、シミュレーション技術等が組み合わせて用いられる。
- (b) 一気通貫の自動化へ：問題解決パイプラインで、①→②-1→②-2→②-3→②-4→③とステップを深めるほど、問題の解決に近づき、価値が高まる。つまり、たとえば①と②-1しか自動化されなければ、②-2以降は人間が行うことになるが、①から③まで一気通貫で自動化されれば、人間は実行状況をモニタリングしていればよいことになる。電力マネジメントの例で具体的に説明するならば、前者のケースは、電力消費状況の計測・可視化までが自動化され、その状況に基づいて人間が今後の必要量を判断し、アクションを考えることになる。後者のケースは、電力消費状況を自動計測し、今後の必要量を自動予測し、最適な状況になるように自動制御も行われる(人間はその様子を見ていけばよい)。このような①から③まで一気通貫での自動化を可能にする方向で、技術開発が進められている。特に注目される技術として機械学習プロセスの自動化技術(AutoML)があるが、これについては(3)のトピックスにて述べる。
- (c) サイバー空間から実世界へ拡張：前述したように、問題解決パイプラインは、当初、インターネット上のサービスに集まるデータを収集・解析し、そのサービスを改良・強化するために使われた。つまり、サイバー空間に閉じたパイプラインであった。しかし、現在は実世界(フィジカル空間)からデータを収集し、その解析結果に基づいて、実世界のシステムにフィードバックをかけるような応用へも広がっている。つまり、サイバーフィジカルシステム(Cyber Physical Systems：CPS)としての問題解決パイプラインへと拡張されている。この拡張は、問題解決パイプラインにおける、①データ収集・蓄積ステップと③アクション実行ステップが、サイバー空間から実世界に広がったということである。そのために、センサーやアクチュエーターを含むさまざまなIoTデバイス、あるいは、ロボットが①や③に導入されるようになった。軽量化、省エネ化、高解像度化、スマート化等の技術改良が進められている。

さらに、問題解決パイプラインを支えるデータ基盤技術として、データ処理基盤技術、データ保護技術、オープンデータ技術等の技術開発も進められている。

- データ処理基盤技術は、大規模なデータを高速に処理するための技術群である。ますます大規模化するデータを、より高速に処理するという要求が高まり、分散並列処理技術、圧縮データ処理技術、ストリームデータ処理技術等が発展している。
- データ保護技術は、分析対象となるデータの保護のための技術群で、暗号化等のセキュリティー技術に加えて、分析対象データが個人属性や行動履歴のようなパーソナルデータである場合に、そのプライバシーを保護するための技術、さらには、データの分析と保護を両立させるプライバシー保護型データ分析技術が取り組まれている。データ匿名化、分散プライバシー、秘密計算等の技術がある。
- オープンデータ(Open Data)は、最小限の制約のみで誰でも自由に利用、加工、再配布ができるデータのことである。さまざまな問題解決にデータ利用が促進され、また、他のデータと組み合わせた新しい価値創出・サービス創出が活性化される。そのために、共通的なデータ形

式や付加的な情報（メタデータ等）の記述形式等がデザインされている。特に、セマンティックウェブ分野で開発・標準化された技術を用いたリンクト・オープンデータ（Linked Open Data：LOD）がよく知られている。さらに、分野・組織をまたいだデータの連携を容易にするため、共通語彙の設定も含むデータ連携基盤の構築が推進されている。米国では NIEM（National Information Exchange Model）、欧州では SEMIC（Semantic Interoperability Community）、日本では共通語彙基盤 IMI（Infrastructure for Multilayer Interoperability）の構築が進められており、IMI、NIEM、SEMIC の間の国際的な相互運用性も検討されている。

(3) 国内外のトピックス

データに基づく問題解決の共通の枠組みを強化する技術として、注目されるのは機械学習プロセスの自動化技術 AutoML（Automated Machine Learning）である。機械学習を用いたデータ分析作業において、従来、分析のための前処理や構造設計（特徴エンジニアリング、モデル選択、ハイパーパラメータチューニング等）のステップは、人手で試行錯誤しなくてはならなかった。しかし、AutoML は、それら各ステップにおいて考えられるさまざまな組み合わせの中から最適なものを効率よく探索することを自動化する。代表的なシステムとしては、DataRobot、dotData、Google Cloud AutoML、Optuna 等が知られている。AutoML 技術によって、機械学習・データ分析のプロセスに要する期間が大幅に短縮されることを加えて、必ずしも高い専門知識を持たない人でも簡単に実施可能になり、機械学習を用いたデータ分析の適用場面が広がる。

一方、本節に示したような枠組みが生む大きなインパクトとして、AI やディープラーニングが科学の考え方や枠組みそのものに変化を与えることが論じられている。AI やディープラーニングは、人間の認知限界・思考限界を超えた発見を科学にもたらし得る。特に以下の二人からの示唆が注目されている。

- ・ ノーベルチューリングチャレンジ（AI による科学的発見というグランドチャレンジ）：北野宏明氏が 2016 年に AI による新たなグランドチャレンジを立ち上げた。そのチャレンジは「2050 年までに、ノーベル賞かそれ以上の科学的発見を行う人工知能を開発する」ことを目標とするものであり、彼自身は特に生命科学分野をターゲットとして取り組んでいる。「科学的発見は、仮説空間の網羅的探索と検証によってなされる」という仮説を置き、このプロセスにおける人間の限界を AI 技術によって超えようとする試みである。ノーベル賞級の科学的発見を行うために道具として AI 技術を使うことに主眼があるのではなく、科学的発見の本質を解明し、人間の知識の飛躍的拡大をサポートする AI 技術の実現によって、人類が直面する問題の解決につなげることを目指している。
- ・ 高次元科学：丸山宏氏が 2019 年に自身のブログにて提唱し、同年の人工知能学会全国大会での招待講演においても言及した新しい見方である。科学はこれまで、現象を正確に説明できる、よりシンプルなモデル・原理（少数のパラメータによる記述）を求めてきたが、それは人間が認知できる範囲に制限されていることを意味するのかもしれない。ディープラーニングは人間には理解することが困難なほどの多数パラメータで現象を記述し、人間の認知限界を超えた現象のモデル化を可能にしたと見ることができる。このような新たな科学の可能性を「高次元科学」と呼んでいる。

(4) 課題と方向性

本節に示したようなデータに基づく問題解決の発展によって、(1) でも述べたように、実世界で起きている現象・活動の状況の精緻かつリアルタイムの把握・予測、膨大な選択肢の網羅的検証、大規模複雑タスクの自動実行等が可能になり、問題解決手段の質的变化、科学の方法論の革新、産業構造・社

会システムの転換等がもたらされ得る。ただし、そのような革新・転換が進むかどうかは、その問題に関わる実世界ビッグデータを取得できるか、また、その分析結果を実世界へフィードバックして状況改善に結び付ける制御手段が整っているかといった点に大きく依存する。取り組みが強く求められる分野、解くべき問題に関して、そのような仕組みを整備していくことが求められる。

一方、そのような革新・転換が起こることに対して、人材育成や社会制度の面の対応も課題である。人材面では、革新・転換の結果、そこで人々に求められる能力・スキルも変化する。しかも、その変化が早いいため、人間の能力・スキル獲得のスピードが追いつかないことが問題になる。人材の再教育システムが重要になる。また、革新・転換の効果を高めたり、社会や人々の受容性を考慮したりする上で、制度設計・規制緩和の施策も考えていく必要がある。日本の社会に合わせる面だけでなく、グローバルな調和と競争環境も意識して設計していくことが求められる。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 システム・情報科学技術分野(2019年)」第2.1.6節「データに基づく問題解決」(CRDS-FY2018-FR-02)(2019年3月)
- ・ 北野宏明、「人工知能がノーベル賞を獲る日、そして人類の未来―究極のグランドチャレンジがもたらすもの」『人工知能』31巻2号(2016年3月)、pp. 275-286
- ・ 丸山宏、「高次元科学への誘い」、CNET Japan ブログ、2019年5月1日
https://japan.cnet.com/blog/maruyama/2019/05/01/entry_30022958/

4.2 データ処理技術：AI・機械学習

(1) 定義・意義

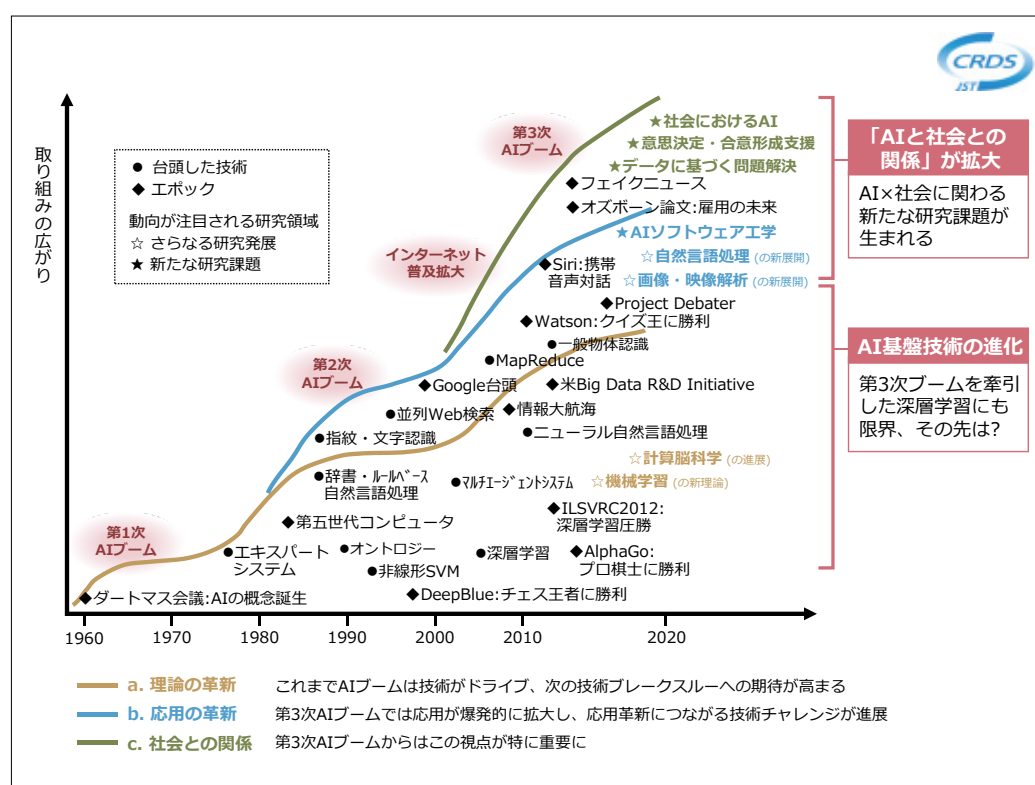
ビッグデータとAI技術の発展・普及によって、「データに基づく問題解決」の枠組みが幅広く活用されるようになった。この枠組みは基本的に、①データ収集・蓄積ステップ、②データ分析ステップ、③アクション実行ステップ、で構成される（図4-1）。

ここでは、②の根幹となるデータ処理技術について取り上げる。人工知能（Artificial Intelligence：AI）技術は、人間の知的活動（認識、判断、計画、学習等）をコンピューターで実現するための技術群である。

(2) 研究のトレンド

昨今、第3次AIブームと言われるが、各ブームでAIが次の世代に移ったとみなすならば、1950年代後半から1960年代にブームとなった第1世代は、AIに関わる基礎的な概念が提案され、AIが新しい学問分野として立ち上がったが、実用性という面ではまだ“トイシステム”であった。1980年代の第2世代は、人手で辞書・ルールを構築・活用するアプローチが主流で、エキスパートシステム、指紋・文字認識、辞書・ルールベース自然言語処理等（カナ漢字変換等）の実用化にも結び付いた。2000年代以降の第3世代は、インターネットやコンピューティングパワーの拡大を背景として、ビッグデータ化と機械学習の進化によって、高精度な分類・識別、予測、異常検知、生成・変換等が可能になり、さまざまな応用システムが実用化され、社会で広く使われるようになった（図4-2）。

第3世代AIの中核はディープラーニング（深層学習）技術であり、さまざまなタスクで大幅精度向上を達成してきたが、いくつかの限界・問題も指摘されつつある。具体的には、①学習に大量の教師データや計算資源が必要、②理論的に未解明で、学習結果の信頼性が確保できていない、③実世界環境・文脈・常識等の理解や適応ができていないといった問題である。このような問題を克服することが、次の世代のAI（第4世代AI）につながる。そのための取り組みとして、機械学習に基づく帰納型の処理に、知識・記号推論のような演繹型の処理を融合させたアーキテクチャの発展・最適化や、人間の脳情報処理に学ぶアプローチ等が検討されている。



まずディープラーニングを中心とした第3世代AIの発展が著しい。さまざまな特定用途で次々と、人間を上回る精度・性能を達成してきた。第3次ブームの初期の頃、画像認識分野のタスクである一般物体認識のコンペティションILSVRC(ImageNet large scale visual recognition challenge)では、2012年にディープラーニングが初めて導入されて、前年までのエラー率を一気に10%近く改善して大きな注目を集めた。その後も精度改善が進み、2015年には人間の認識精度を上回ったと言われる。また、コンピューターゲーム分野では、ディープラーニング・深層強化学習を用いたGoogle DeepMind社のAlphaGo（アルファ碁）が2016年～2017年に世界トップランクプロに圧勝して大きな話題になった。囲碁は完全情報ゲームであるが、不完全情報ゲームにおいても、2018年にはリアルタイムストラテジーゲームStarCraftIIでGoogle DeepMindのAlphaStarがプロプレイヤーに圧勝し、2019年には6人制ポーカーでFacebookとCMUのPluribusがトッププロに大勝した。さらに、質問文に対して関連文書を読解して回答を抽出する文章読解問題についても、2018年にコンピューターが人間を上回る精度を達成し、その後、BERT、XLNetといったシステムが次々に最高精度を更新している。

このような成果の拡大は、ディープラーニング技術そのものが発展し続けていることに伴うものである。ディープラーニングが強化学習と結合した深層強化学習によって、繰り返し試行を通して、最適なアクションを決定することが可能になり、ゲームやロボット制御等への応用が広がった。自然言語処理の分野では、文脈類似性に基づく意味の分散表現とその演算が可能になり、ルールベース言語処理からニューラルネットワーク言語処理へと処理系が再編され、画像・映像・音声とのマルチモーダルな処理系も構築しやすくなった。さらに、敵対的生成ネットワークGAN(Generative Adversarial Networks)等、ディープラーニングに基づく生成モデルが開発されたことによって、新しい生成系アプリケーションが次々に生み出されている（たとえば、あたかも本物のような実在しない画像・映

像の生成、他者の動作のコピー、線画の自動着色・写実画変換、見えない面まで推測した 3D モデル生成等)。

しかし、その一方で、ディープラーニングの限界・問題（前述の問題①②③）が指摘されるようになってきた。さらに、AI 技術がさまざまな形で社会に広がるにつれて、AI と社会との関係から生じる新たな問題も生まれてきた（図 4-2）。たとえば、以下のような問題があげられる。

- ・ ブラックボックス問題：精度が高くとも人間に理解可能な理由説明を出してくれない、動作保証ができない、事故発生時に原因解明や責任判断ができないといった問題であり、説明責任に関わる。
- ・ バイアス問題：機械学習の判定結果は学習データの傾向を反映するため、学習データが偏見を含んでいれば、判定結果にも偏見が反映され、差別を増長するという公平性に関わる問題である。
- ・ 脆弱性問題：機械学習は学習データと比べて想定外のデータに対してどう振る舞うかは不明であることや、誤認識を誘発する攻撃（Adversarial Examples）が可能であること等の脆弱性が指摘されている。
- ・ 品質保証問題：機械学習は仕様（正動作）が定義されないためテストの成否が定まらないことや、精度 100%は無理で、間違いは不可避であること等、製造物責任にも関わる問題である。
- ・ フェイク問題：フェイクニュースが拡散しやすいことに加えて、GAN 技術の進化によって、本物と見紛うフェイク動画等の生成が可能になり、SNS を介した世論誘導・政治干渉の危険性が増大している。写真や動画が証拠として信じられるのかという問題も生じてくる。

このような社会との関係から生まれた AI の新たな問題に対処するため、機械学習の精度・性能を高めるのとは異なる研究開発も取り組まれるようになった。具体的には、説明可能 AI（XAI）、公平性配慮型機械学習、Adversarial Examples 対策、AI ソフトウェア工学（機械学習応用システムのテスト手法・開発方法論等）、フェイク検出手法等が取り組まれている。

また、ディープラーニングの限界を克服する新たな AI モデルに向けて、人間の脳情報処理から学ぶアプローチが注目されている。AlphaGo をはじめ革新的な AI 技術を開発している DeepMind は「知能の解明」を企業ビジョンとして掲げ、Neuroscience-Inspired AI のアプローチを推進している。AI 研究におけるディープラーニング、強化学習、アテンション、エピソード記憶、転移学習等は、脳情報処理の知見にインスパイアされたと言われる。

（3）国内外のトピックス

日本政府は 2019 年 3 月に「人間中心の AI 社会原則」を決定した。ここでは、「AI-Ready な社会」をビジョンとし、人間の尊厳が尊重される社会（Dignity）、多様な背景を持つ人々が多様な幸せを追求できる社会（Diversity & Inclusion）、持続性ある社会（Sustainability）を基本理念として掲げている。そして、AI 社会原則として、①人間中心の原則、②教育・リテラシーの原則、③プライバシー確保の原則、④セキュリティ確保の原則、⑤公正競争確保の原則、⑥公平性、説明責任及び透明性の原則、⑦イノベーションの原則を示した。2019 年には他にも、欧州委員会が「信頼できる AI のための倫理ガイドライン」、IEEE が「倫理的に配慮されたデザイン」、OECD が「人工知能に関する OECD 原則」、中国が「次世代 AI のためのガバナンス原則」を発表する等、AI 社会原則・倫理ガイドラインが国・世界レベルで議論・策定されるようになった。

このような AI 原則・倫理ガイドラインの内容はいずれも似たような方向性を掲げているものの、国の AI 強化の戦略は各国の特徴が表れている（表 4-2）。米国は、GAFA がビジネスと基礎研究の両面で圧倒的な優位性を持っており、GAFA やスタートアップが活発な民間技術開発を推進している一方、DARPA が国としての中長期的な研究投資をシャープに打ち出している。中国は、AI リード企業

5 社（Tencent、Baidu、Alibaba、iFlyteck、SenseTime）を選定し、政府が AI 産業を後押ししていることに加えて、AI を活用した監視・管理社会（社会信用システム、天網、金盾）の構築を進めており、他国と大きく異なる AI 応用技術開発が見られる。欧州は GDPR（一般データ保護規則）のようなデータ保護戦略によって GAFA に対抗しようとしている。日本は前述の「人間中心の AI 社会原則」「AI 戦略 2019」を策定したが、研究開発面では「Trusted Quality AI」を打ち出し、AI 品質を競争力に近づけようという点に特徴がある。

表 4-2 各国の AI 関連戦略・政策



米国	- GAFAがビジネスと基礎研究の両面で圧倒的に優位 - GAFAやスタートアップによる民間の活発な技術開発の一方、DARPAが国としての中長期的な研究投資をシャープに打ち出している：DARPA AI Nextに20億ドル投資 - 経済・国家安全保障のためのAIを強化、「AIにおける米国のリーダーシップ維持」を大統領令で宣言
中国	「次世代人工知能発展計画(AI2030)」を掲げ、AIリード企業5社(Tencent、Baidu、Alibaba、iFlyteck、SenseTime)を選定、政府がAI産業を後押し - 同時にAIを活用した監視・管理社会(社会信用システム、天網、金盾)の構築推進、他国と大きく異なるAI応用技術開発を推進
欧州	「信頼できるAIのための倫理ガイドライン」を策定 - データ保護戦略(GDPR)を含めて、AIに関わる国際ルール作りを通じて米中やGAFAに対抗 - 各国のAI戦略に加えて、Horizon 2020、Horizon Europeによる国横断のAI研究推進
日本	「人間中心のAI社会原則」「AI戦略2019」を策定、Trusted Quality AIを打ち出し、実世界AIに加えて、AI品質も競争力に - 理研AIP、産総研AIRC、NICTが中核国研として国のAI研究を牽引 - 米中2強への対抗として欧州(特に独仏)と協調

第4世代AIに向けた研究開発として、(2)でも触れたように、機械学習に基づく帰納型の処理に、知識・記号推論のような演繹型の処理を融合させたアーキテクチャの発展・最適化や、人間の脳情報処理に学ぶアプローチ等が検討されている。これに関わる国内トピックとしては、東京大学の主催による公開シンポジウム「深層学習の先にあるもの—記号推論との融合を目指して」が2018年1月と2019年3月に開催され、AI分野のリーダー的研究者からこの方向に向けた取り組みが示されたことや、新学術領域研究として「人工知能と脳科学の対照と融合」（2016年6月～2021年3月）が立ち上げられ、AIと脳科学の主要研究者による連携・融合研究が推進されていることがあげられる。

(4) 課題と方向性

脳科学分野では、蛍光イメージング法、脳波測定法、PET、fMRI等、脳の活動を計測する技術が発展し、低侵襲・非侵襲化、分解能向上が図られてきた。これにビッグデータ解析・機械学習技術を組み合わせて、ブレインデコーディング、モデル化解析、ニューロフィードバック・DecNef法等、脳情報処理をより深く理解する手段も生み出されてきた。このような技術発展に日本の研究者は大きく貢献してきた。これをAI研究のブレークスルーにつなげるためには、個々のニューロンや脳内各部の神経

回路といったミクロなレベルから、脳全体の活動をとらえるマクロなレベルまで、さまざまな階層の研究成果を統合し、脳情報処理を総合的に解明していく取り組みが今後いっそう重要になっていく。そのために、多階層でビッグデータを蓄積していくことや全脳シミュレーションのためのコンピューティング基盤の研究開発も重要である。

また、AI 研究の発展には、脳科学や人間理解と AI、社会課題やソーシャルメディアと AI、倫理や法制度と AI、あるいは、AI の中でも帰納型 AI（機械学習等）と演繹型 AI（論理推論等）といった異なる研究分野間の横断的な研究推進が不可欠になってきている。そのような人材を育成するためには、各分野の研究者間の連携・交流だけでなく、大学教育の段階から多分野にまたがる学習が促進されるような仕組みも考えていくべきと考える。さらに、AI 研究人材、脳研究人材、多分野横断的人材等は、グローバルに不足しており、その獲得競争やリテンションにおいて、現状、日本の制度・環境に優位性があるとは言いがたい点にも課題がある。

AI 分野は、米国・中国が圧倒的な投資で技術開発を加速している状況において、日本は勝ち筋を描いて、そこに投資をフォーカスしていくことが必要である。従来、GAFA がサイバー世界で圧倒的優位であるため、日本が狙うのは実世界データに基づく AI だという考えが示されてきた。しかし、GAFA や中国企業も当然ながら、サイバー世界だけでなく実世界も攻めてきている。この状況において、「AI 戦略 2019」で新たに打ち出した Trusted Quality AI（CRDS からこれまで提言してきた AI の信頼性・安全性確保）は、日本が強みを出し得る方向性である。また、脳科学分野も日本に技術蓄積がある分野であり、脳科学と AI 研究を融合させるアプローチも日本の強みを構築し得る方向性と考えられる。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター「研究開発の俯瞰報告書 システム・情報科学技術分野（2019 年）」（CRDS-FY2018-FR-02）（2019 年 3 月）
- ・ 科学技術振興機構 研究開発戦略センター「戦略プロポーザル AI 応用システムの安全性・信頼性を確保する新世代ソフトウェア工学の確立」（CRDS-FY2018-SP-03）（2018 年 12 月）
- ・ 科学技術振興機構 研究開発戦略センター「戦略プロポーザル 複雑社会における意思決定・合意形成を支える情報科学技術」（CRDS-FY2017-SP-03）（2018 年 3 月）
- ・ 内閣府（統合イノベーション戦略推進会議決定）「人間中心の AI 社会原則」2019 年 3 月 29 日
- ・ 内閣府（統合イノベーション戦略推進会議決定）「AI 戦略 2019」2019 年 6 月 11 日

4.3 データ取得技術：物理・化学的（量）計測

(1) 定義・意義

IoT 社会に必要なコア技術としての計測技術は、放射光、中性子等、量子ビームの大型研究施設から電子顕微鏡、NMR 等の汎用計測機器に至るまで、物質・材料やライフサイエンス・医療など様々な研究開発分野に浸透し、産業界、学术界において有効活用されている。計測データから有意な情報を読み解く際に、現状では研究者の経験と勘に頼る部分もまだまだ多いが、最近の計測機器のコンピューター制御技術とデータストレージの進歩に伴い、計測にデータ科学を融合させた“計測インフォマティクス”と呼ばれる新たなパラダイムが誕生している。

計測技術の発展における ICT の貢献には、①計測の高性能化、②計測の高次化、③計測の多面化、の 3 タイプがある（図 4-3）が、ここでは“計測インフォマティクス”が及ぼす科学技術上の効果として、自然科学における計測、すなわち物理量計測を基礎とする①と②のタイプに限定して記載する。計測技術の中心である物理量計測がより高精度、高効率で可能になったり、AI 技術による分析結果（解釈）を含む意味的計測、そのフィードバックまでを行う自律的計測といった、より高次の計測が可能になると期待される研究開発領域である。

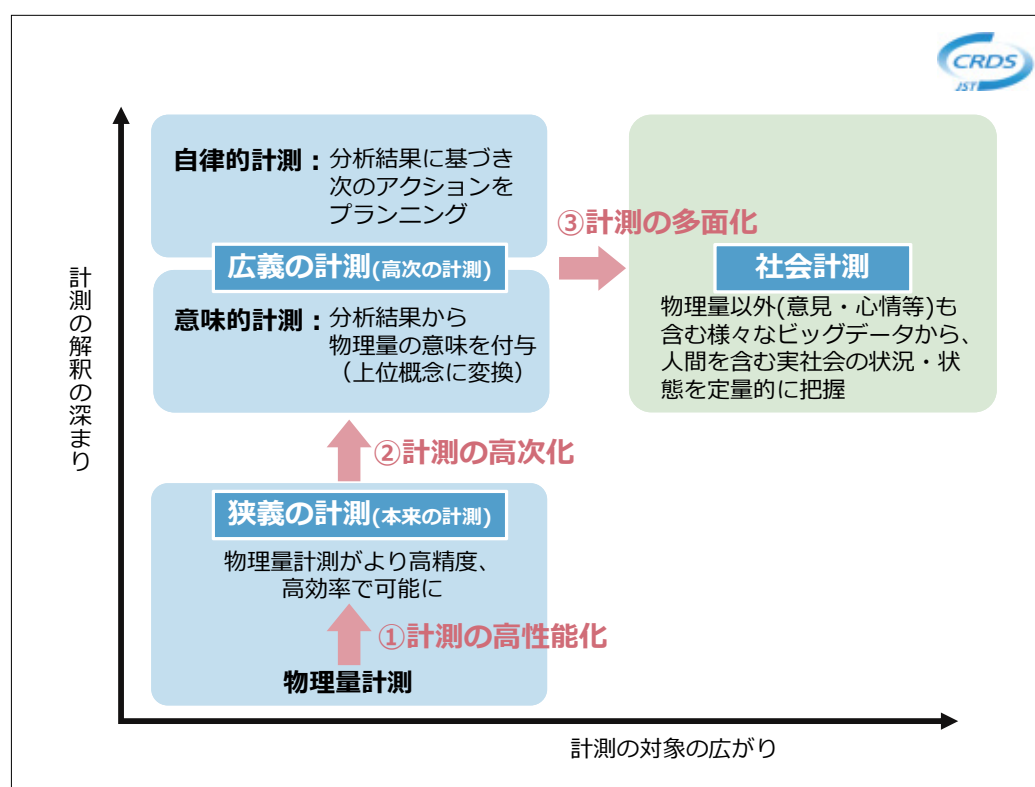


図 4-3 計測技術の発展における ICT の貢献

一般に、先端的な計測デバイスや装置を対象にした研究開発の多くは、測定対象の濃度、サイズや画像解像度、測定時間・頻度など様々な条件で極限の計測を目指しており、このときの計測データは多次元であるとともに多くのノイズを含むことが多い。したがって、このような多様、かつ大量の情報からの対象推定や複数測定情報の統合、事前知識による補完など、十分な計測結果を得るために高

度な推定処理が必要とされる。また、AI 基盤である高度な数理理論や統計理論に基づく機械学習やデータ構造・アルゴリズムの原理・技術を計測に応用するために、計測技術固有の課題に対してデータ科学側からの新たな研究開発も重要である。

(2) 研究のトレンド

(1) で述べた計測の高性能化、計測の高次化に関し説明する。まず、高性能化とは主として高精度化と自動化（効率化）を指す。高精度化では、機械学習を用いて、データ数により観測精度を高めたり、大量の画像データから規則性をつかみ画像・映像を高解像化したりすることが可能となっている。生命科学に画像処理・統計処理等を融合したバイオイメージ・インフォマティクス例では、細胞内の状況と現象を人間の眼よりも正確かつ安定的に可視化できるだけでなく、高解像度の観察画像が得られている。一方、自動化（効率化）では、様々なケースを対象とした計測手順の繰り返しを制御できるようにシステム化したり、計測対象・環境の条件に応じた計測手段やそのパラメータの選択を自動化したりといったことが進められている。このような網羅的探索は人間よりもコンピューターによる実行に適した処理であり、“計測インフォマティクス”による大きな効果が期待できる。

高次化には上記のとおり意味的計測と自律的計測がある。意味的計測は、計測結果として得られた物理量を分析することで、その意味を与える（上位概念に変換する）ことである。スペクトル計測において、信号の物理的要因を知るために必要なピーク分離を情報学的に行う試みはそのような例の一つであり、解釈の多義性を回避できると期待されている。また、地球観測・環境計測の分野で運用されているデータ統合・解析システム（Data Integration and Analysis System, DIAS）では、地球規模で各地域の観測データを収集・蓄積・統合・解析し、社会経済情報等とも融合することで、地球規模の環境問題や大規模自然災害等の脅威に対する危機管理に有益な情報へ変換している。一方、自律的計測は物理量の分析結果に基づいて、次のアクションのプランニングまで行うものである。データ同化（観測データとシミュレーションのすり合わせによる数値モデルの精度向上）による気象現象の予測、強化学習による車の運転制御や産業用ロボット制御の最適化などの例が知られている。

このように、既に身近で利活用されている“計測インフォマティクス”においては、大きく分けて次の2つの課題に取り組まれている。一つはベイズ推定、スパースモデリング、ディープラーニング、データ同化等の情報学・統計数理による物理量解析や大量データの迅速・高精度解析に必要な手法の開発である。関連研究分野としては、統計的モデリング、データマイニング、強化学習、転移学習、テキスト検索、情報検索、Web 情報解析、自然言語処理、音声解析、画像認識、情報抽出、最適化などがある。もう一つは、このデータ解析に係る技術基盤を具体的な計測課題に応用し、物質・材料やライフサイエンス・医療をはじめとした科学技術全般における新現象の発見、原理の解明や新たな知識獲得を目的とするものである。後者では、データ解析の専門家に加えて、情報、物理、化学、生物など様々な面から個別の先端計測を指向する者や、実際の計測デバイス・装置、情報処理のためのエッジコンピューティングデバイス・アーキテクチャに関する産官学の研究者、技術者との融合が重要である。

ハイスループット計測が当たり前になった現在、例えば二次元スペクトルマッピングのように大量の実験データが容易に得られるようになった一方、取得データの多くがほぼ手つかずのまま埋もれ、あるいは貯めることなく捨てられているというのが先端計測分野における負の側面とも言える。データ科学に基づき大量のデータ処理を可能にする“計測インフォマティクス”は、平均化やエラー評価のための統計処理と異なり、こうした埋蔵データを先端的かつ信頼性高く活用する新たな手段として期待される。この意味では、計測データを蓄積、さらには研究者間で共有することによって情報学的な解析をさらに効果的に行えるデータ蓄積・活用プラットフォームを構築することも重要な取り組みである。特に大型研究施設では、こうした利活用を前提とした計測データリポジトリシステムの開発に取り組む必要がある。

(3) 国内外のトピックス

諸外国では、米国を皮切りに欧州や中国、韓国でも大型プロジェクトが進んでおり、その中で計測の個別問題に取り組まれているものの、“計測インフォマティクス”に注力するための大規模な政府投資は見られない。一方、わが国では、この分野で世界に先駆けた大型プロジェクトとして、文部科学省科学研究費補助金「新学術領域研究」の「スパースモデリングの深化と高次元データ駆動科学の創成」が2013～2017年に実施された。その後2016年からは、JSTの戦略的創造研究推進事業CREST・さきがけ「計測技術と高度情報処理の融合によるインテリジェント計測・解析手法の開発と応用（情報計測）」が推進中である。

物質・材料の観点では、JSTの「情報統合型物質・材料開発イニシアティブ (MI²I)」、JST 未来社会創造事業の探索加速型「共通基盤」領域の重点公募テーマ「革新的な知や製品を創出する共通基盤システム・装置の実現」、CREST「実験と理論・計算・データ科学を融合した材料開発の革新」、さきがけ「理論・実験・計算科学とデータ科学が連携・融合した先進的マテリアルズ・インフォマティクス (MI) のための基盤技術の構築」、戦略的イノベーション創造プログラム (SIP) 第2期の「統合型材料開発システムによるマテリアル革命」、NEDOの「超先端材料超高速開発基盤プロジェクト」などにおいてMI研究課題の一部として進められている。

以上の取組みにおける成果の例として、熊本大学を中心とするグループは、放射光X線で取得した吸収スペクトルの解析法に、スパースモデリングによる情報抽出技術を適用することで、事前知識を必要とせず測定データのみから原子周辺の構造と原子の構造ゆらぎが解析できる手法を世界で初めて開発した。原子スケールの構造情報を入力して解析する従来のフーリエ変換法とは対照的に、事前知識の存在しない新奇材料の開発などに貢献することが期待される。また、東京大学などのグループによって、数千～数万に及ぶ膨大なスペクトルデータが、専門知識や職人技がなくても高速かつ高精度に解析された。具体的には、樹形図を作成する2種類の機械学習(階層型クラスタリングと決定木)を活用することによって、AIでデータを解釈し原子配列や電子構造の情報を獲得する新手法であり、従来の約2万倍の速さでスペクトルの解釈と予測が可能になった。X線吸収スペクトル解析の自動化、高効率化は、機械学習の一種である多様体学習および、データ類似度の概念を応用することでも実現している(高エネルギー加速器研究機構など)。

ライフサイエンス・医療分野では、画像取得技術の発展に伴い、画像解析へのAI適用が研究されている。例えば、光学顕微鏡においては、光学的分解能を超えたナノスケールの構造を観察するための技術として超解像顕微鏡が普及しつつあるが、生体内現象のダイナミックな計測に必要なレベルで高速に超解像画像を取得するためには、空間分解能を決める要因であるシグナル/ノイズ(S/N)比と時間分解能のトレードオフを破る必要がある。米国・カリフォルニア大学を中心とする研究グループは、機械学習を用いてS/N比の低い画像から高分解能の画像を生成することにより、従来より10倍以上高速な超解像画像の取得に成功している。また、従来は、透明な細胞内構造は通常の見視野だと観察が難しいため、蛍光プローブで標識することにより可視化するのが一般的であったが、特に再生医療の分野において、蛍光標識なしに画像評価を行いたいという要望が高まっている。実際に、機械学習を用いて、明視野画像から各組織を色分けした3D画像を再構築する手法がトップジャーナル(Nature Methods (2018))で報告され、注目されている。

2013年から2014年にかけて、脳機能の全容解明・全神経ネットワークの可視化(コネクトーム)を目指して、Brain/MINDS(日本)、Brain Initiative(米国)、Human Brain Project(欧州)といった大型プロジェクトが開始された。その中で、モデル動物の脳全体をイメージングするハイスループット測定技術が開発され、取得された膨大なデータの解析のためにデータ科学が活用されている。例えば、ATUM-SEM(Automated Tape-collecting UltraMicrotome Scanning Electron Microscopy)という手法では、固定した脳を薄くスライスして電子顕微鏡撮影を繰り返すことにより、2次元画像から3

次元像を再構築する。電子顕微鏡ではナノレベルの高分解能画像が得られる反面、すべての構造が可視化されるため、ディープラーニングを用いた画像解析で神経回路を識別することによって、膨大な量の画像からのコネクトーム解析を自動化している。一方、行動中の個体に対して脳全体の神経活動を計測することは非常に難しいが、機械学習を用いることで、脳表面の部分的な計測から全体の活動パターンを推測する、といった研究が行われている。

(4) 課題と方向性

①高レベルな学術的研究の推進

“計測インフォマティクス”における今後の学術的課題は、単なる既存のデバイス・装置と既存のAI技術との組合せではなく、おのおの基礎的原理にまで遡り、情報処理を指向するデバイス・装置原理の基礎研究、計測を指向する機械学習や統計的推定原理の基礎研究を推進していくことである。そのためには、特に既存の各研究分野のテーマや価値判断の枠にとらわれずに、いかに柔軟でかつ学術的レベルの高い研究を誘発する環境を提供していくかが重要となる。

②製品開発に向けた産業界との融合

今後は、産業界や市民の生活環境においても、学際融合的な新しい先端計測技術が広く使われていくと予想される。特に、わが国は、米国や中国などのように数億人を超える大規模市場を抱える地域とは異なり、自国内でGAFAに代表される大規模情報サービス産業を育成することが難しい。したがって、引き続き世界に先駆けてデバイスや装置、システムなどの優れた製品を開発して世界市場に提供していく経済活動が求められる。このような環境において、“計測インフォマティクス”の果たす役割は特に大きい。

③裾野拡大への活動強化

計測関連の研究は、首都圏や関西圏の大学や研究機関に留まらず、全国的広がりとは多様な枠組みにより広範な分野で取り組まれるようになってきている。また、多くの計測関連製品の開発、製造が地方の民間企業でも行われていることは見逃せない。このような先端的計測の高性能化への関心が高まる中、2018年4月に「計測インフォマティクス研究会」が人工知能学会の第2種研究会として設立された。ここでは、地方での研究会・セミナー実施に加えて、関連する国際会議、ワークショップなどの国内開催に向けた基盤作りを視野に入れた活動が進められている。

“計測インフォマティクス”に関する今日の研究開発は、わが国の強みである数理科学における地道で息の長い研究が基盤となっており、ここに一定の存在感がある。したがって、中長期的な視点で、関連する学術分野の進展やコミュニティの充実・多様化と、それらを担う研究者、技術者のネットワークを整備することによって、計測インフォマティクスの研究開発を持続的に支えられるようなシステムを構築することが重要である。その意味でも、上記の「計測インフォマティクス研究会」への期待は大きい。

(5) 参考資料

- ・ 科学技術振興機構 研究開発戦略センター 「計測横断チーム調査報告書 計測の俯瞰と新潮流」(CRDS-FY2018-RR-03) (2018年12月)

4.4 データ取得技術：センサー・ロボット（環境・社会計測）

(1) 定義・意義

高度なサイバーフィジカルシステム（CPS）を築くためには、人工知能などの新たな情報処理技術や高速大容量ネットワーク技術と並んで、現実の世界（フィジカル空間）の多種多様な情報（データ）を取得するセンシング技術が重要になっている。センシングデバイス（センサー）を必須とする IoT 応用分野は、自動運転、ロボット、健康・医療、スマート農業、社会インフラ管理、スマート工場などがある。例えば、自動運転には高性能ジャイロ스코プ、加速度センサー、高精度 GPS、LIDAR（Laser Imaging Detection and Ranging）、高解像度可視光イメージセンサー、赤外線イメージセンサー、マイクロ波レーダー、運転者の状態を把握するセンサー（視線センサー、アルコールセンサー、心拍センサーなど）が必要である。これには、MEMS（Micro Electro Mechanical Systems）センサーや化学センサーなどセンシングデバイスの高感度化、高信頼化、低消費電力化、小型軽量化、低コスト化、MEMS プロセス技術の高度化、プリントエレクトロニクス技術などの研究開発課題がある。

また、進化し続ける研究機器によるデータの量と質が飛躍的に向上したことを背景に高度なロボット技術を活用し、バイオ・ライフサイエンスや化学・材料開発分野の研究生産性の向上と良質のデータ取得を目指した技術開発も進められるようになってきた。すなわち、ロボットにより研究開発に必要な実験条件の設定、実験、生成物の分析・評価によるデータ取得などの作業を自動化するものである。これにより、研究に関わる多様なデータを取得するとともに、品質の高い多くのデータを得ることにより、人工知能（AI）が必要とする学習の高効率化や、研究開発効率の向上が期待される。

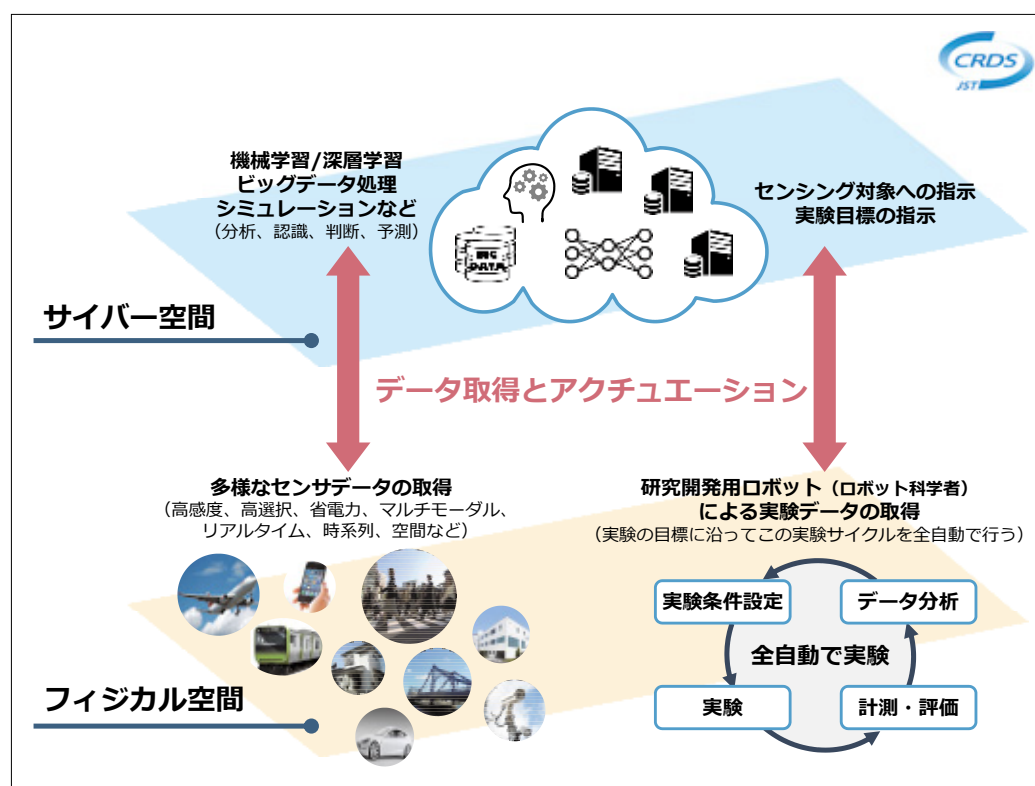


図 4-4 データ取得技術 [センサー・ロボット] の概念

(2) 研究のトレンド

ここでは、MEMS センサー・化学センサー（バイオセンサー）と研究開発用ロボットの研究トレンドについて述べる。

(a) MEMS センサー・化学センサー（バイオセンサー）

センサーとしては、画像・映像を取得するイメージセンサーや加速度、圧力、音などの多様な物理量を検出・測定可能な MEMS の利用が拡大しており、現在では世の中のセンサーの多くが MEMS 技術によるものである。スマートフォンを例にとると、そこにはマイクロフォン、加速度センサー、ジャイロスコープ、圧力センサー（気圧センサー）など、合わせて 10 個以上の MEMS センサーが用いられている。これらの MEMS センサーの技術は継続的に革新され、その結果、低コスト化と高性能化が進んでいる。例えば、加速度センサーの大きさは、最近の 5 年間でチップダイの面積にして約 1/10 になっている。低コスト化による爆発的な応用の広がり、スマートフォンがそうだったように、想像を超えた大きなパラダイムシフトをもたらす。今後は、自動運転に用いるジャイロスコープ、音声認識のためのマイクロフォン、資源探索等のための重力センサー（高感度加速度センサー）、ロボットの機械学習に用いる触覚センサーなどでの高性能化や高機能化が期待されている。

バイオセンサーは通常は上で述べた分子（またはイオン）認識部とトランスデューサ（信号変換部）とから構成される。人の健康を体内環境として直接モニターするには生体活動に関与する微量物質を高い選択性をもって検出・認識する必要があるため、生物由来あるいは生物に関連する物質がバイオ分子受容体（認識部）として開発されてきた。しかし、バイオ物質は生ものであり多くの場合センサーとして常温で保存することができないという問題があるため、どこでも使えるバイオセンサーという観点から、生体の主要高分子であるタンパク質、炭水化物（糖鎖）、核酸を用いた微量ターゲット物質受容体技術や、呼吸などの生体から出る多様なガスの検出技術研究が進められている。主要なトランスデューサとしては、電気化学デバイスと光検出デバイスがある。小型・軽量で簡便に使えるウェアラブルなセンサーとしては、電気化学デバイスが優れている。しかし、電位を検出するタイプのセンサーでは参照電極が必要となり多くの場合定期的にキャリブレーションを入れる必要がある。また電流検出を利用するタイプのセンサーは消費電力が課題になる。このため、消費電力が少なく安定な測定が可能な電気化学デバイスの開発が望まれている。光検出型の場合は、光源および消費電力が課題である。シリコンフォトリソグラフィや表面プラズマモンなどのナノ・マイクロフォトリソグラフィ技術の利用が期待される。

今後はこれらに加え、物理センサーや化学センサーと集積回路（IC）やプリントドエレクトロニクス（Printed Electronics）との組み合わせも重要になると考えられる。

(b) 研究開発用ロボット

機械（センサーやロボット）を活用した実験の自動化は効率に対する適応的な応答といえる。また、科学は再現性問題の危機に対応しており、この対応にもなる。これらは結局ハイスループットにデジタルデータが排出され、その処理が研究の鍵を握るということにもなる。

従来、大規模研究であるオミックス解析を中心に、ロボットによる自動化研究は行われてきた。これまで開発・上市された自動化装置は、溶媒・試料の分注ロボット、または調整されたサンプルを解析機に自動的に供給するオートサンプラー的なものが中心で、それぞれの作業に特化した単機能ロボットである。

2 つの実験技術、CRISPR 遺伝子編集および細胞の誘導多能性（iPSC）は、生物医学研究に新たな次元をもたらした。一方で、特定のがん細胞内の遺伝子型と表現型の関係の特徴付けには、人間の手で達成できない多くの CRISPR 処理とその組み合わせが必要になる。ターゲットの識別からトランスフェクション、分析まで、可能な限り多くのプロセスを自動化することは、効率と一貫性の観点から重要で、この根底にあるのは、マイクロフルイディクス技術の進展ともいえる。マイクロフルイディ

クスは、チップやフローチャンバーなどのデバイス上に、段階的、並行、または同時的に、多くの異なる実験室手順をミニチュアで統合することを容易にする。

これを受けて、再生医療の分野では細胞製剤生産の自動化が積極的に試みられている。その多くは、細胞・ラボウェア（試験管・試薬瓶・培養皿等）をハンドリングするロボットアームに細胞培養に必要な恒温槽・遠心機等をカスタマイズ自動化し、それに必要な外部軸と特殊治具をインテグレーションしたライン化方式である。

最近では、人の動作をコピーし、人がこれまで使ってきた道具を使い「作業」するロボット「汎用ヒト型ロボット（LabDroid：laboratory + humanoid）」が出てきた。LabDroid の特徴は、これまで自動化が困難であった、複雑で高度な手技を必要とする作業を自動化するのみならず、これまで個人のスキルとして暗黙知化していた作業を可視化・数値化し、数値化されたパラメータを最適化することにより、熟練者よりも高い精度で作業を何度でも再現することを可能としている。

材料開発においても、ロボットを使って全自動で材料合成実験および作製した材料の特性評価を行う試みが出てきている。このインテリジェントなロボットはロボットケミスト、ロボット科学者とも呼ばれ、化学材料開発、有機材料開発において導入が進められており、最近では無機材料開発へも適用する動きが出てきている。例えば、新たな反応性の有機材料の探索用のロボットが開発されており、約 1000 個の反応組合せの中のわずか 10% を合成することで、残りの 90% についても実験しないで 80% の精度で結果が予測できるという報告もある。また、機械学習などのデータ科学的手法を用いて目標仕様に基づいて材料の組成・構造・合成条件までロボットが予想して自動的に設定し、合成・評価・次の実験条件設定を自動的に繰り返して実行し、目標の成果（新規材料および各種の特性評価データ）を得ることができるロボットも研究されている。人間は多次元のパラメータを一度に解析することはできず、2 次元ないし 3 次元にパラメータを集約して実験を繰り返す必要があるが、機械学習では多次元のパラメータと特性との相関を一度に見ることができる。また、実験条件の設定には、研究者による思い込みのようなことは起こらないため、偏りのない探索を行うことができる。また、ベイズ最適化などの手法を用いることで、実験回数を桁違いに少なくすることもでき、人力を遙かに超える速さ・正確性・再現性で各種の実験データや特性評価データを取得しながら新規材料を創製することも期待されている。

(3) 国内外のトピックス

米国で MEMS 技術を約 25 年にわたり強力に支援しているのは、DARPA（Defense Advanced Research Projects Agency）である。DARPA が継続的に重要視しているのは、兵器のナビゲーションに必須の慣性センシング技術である。これに関する最新のプログラムは、Precise Robust Inertial Guidance for Munitions (PRIGM) program と呼ばれる。慣性センサーの 1 つであるジャイロ스코ープには MEMS の中核をなす重要技術が凝縮されており、その研究を行うことで MEMS 技術の本筋の研究に繋がっている。

欧州では Horizon 2020 のもと、いくつかの MEMS 関連プロジェクトが実施されている。多くは Industrial Leadership のためのプロジェクトに整理されている。MEMS スピーカ（https://cordis.europa.eu/result/rcn/198209_en.html）もそれらのプロジェクトの 1 つで開発されたものであり、事業化を目指すベンチャー企業 USound（オーストラリア）が注目を集めている。

中国では、MEMS やパワー半導体の大型研究拠点である SITRI（Shanghai Industrial μ Technology Research Institute）が 2015 年に上海に設立され、遅れていた 8 インチラインの立ち上げは 2017 年に完了したと思われる。半導体デバイスや半導体装置の国産化に多額の投資を続ける中国政府は、MEMS 関連の国内企業にも積極的に補助をしている。最近、米国との貿易紛争で中国国内のハイテク産業が危機にさらされていることが、政府投資を活発にさせている。

バイオ・ライフサイエンス分野の研究開発ロボットでは、New York Stem Cell Foundation は、Global Stem Cell Array を使用して、体細胞の解凍から再プログラミング、ソート、クローン選択、培養拡大、凍結までの 8 段階のロボットワークフローを自動化し、スケーラビリティとスループットを最適化しようとしている。さらに、エボテック社はハーバード大学と提携して iPSC ベースの工業化薬物スクリーニングプラットフォームを構築した。これらのシステムは、オルガノイドや他の派生培養への iPSC の分化に適応できる可能性がある。

最近、パナソニック社が iPSC 細胞の維持培養を自動化する専用ロボットを開発し、60 日間の人非介入の細胞培養に成功した。また、臨床研究が控えている、パーキンソン病を対象とする iPSC 細胞による再生医療も、産業界が積極的に細胞生産の自動化にも取り組んでいる。多くの細胞培養において、顕微鏡画像を教師データとして大量に得ることが可能であり、ディープラーニングのような AI 技術との相性がよいと考えられ、細胞培養の最適化と品質管理に AI 技術を活用する試みもスタートしている。特に、ヒューマノイドのような自由度と汎用性が高いロボット技術と組み合わせた場合、膨大な培養時のパラメータを効率よく最適化し、培養細胞の良し悪しを画像データから非破壊に判断することも可能であると期待されている。

産総研人工知能研究センターが中心となり細胞培養を自動化し AI と融合するプロジェクトがスタートしている（NEDO 次世代人工知能技術の社会実装を目指した先導研究：AI ×ロボットによる高品質細胞培養の自動化とオミックスデータの大規模取得）。さらに、理化学研究所多細胞システム形成研究センターにおいては、顕微鏡画像を教示データとし、AI による iPSC 細胞の維持培養を最適化するのみならず、網膜再生のための分化・誘導プロトコルの最適化を自律的に実施する AI ロボットシステムの研究開発がスタートしている。

同様にゲノム編集も大量の細胞株のゲノムへの一括ハイスループットスクリーニングに用いられ、データが膨大に産出する。

材料分野の研究開発ロボットとしては、英国では有機材料合成でこの構想に基づいた装置の開発（University of Glasgow）や、ロボットケミストの開発計画（University of British Columbia）が進められている。また、日本でも無機材料開発用に、円周上に成膜用の真空チャンバーや特性評価装置を複数配置し、それぞれのチャンバーへのサンプルの受け渡しをロボットアームで行う装置が試作されるようになってきた。

（4）課題と方向性

MEMS センサーの代表的な技術的挑戦は、チップ面積や消費電力を大きくせずに S/N 比を上げること、および現実的な方法で異なる要素を集積化することである。例えば、S/N 比を上げるには、極めて高い Q 値、超高精度加工、超高性能圧電膜、超高感度圧電抵抗などが必要である。これらは、現実的な解があるのかどうかを含めて、原理検討や材料開発から取り組むべき中長期的な課題である。

現在、スマートフォン、自動車、ヘルスマonitoring デバイス、ゲーム機器などに幅広く用いられているセンサーはおおそ物理センサーであるが、バイオマーカーの検出、呼気の分析、空気中汚染物質の濃度測定などを低コストで行うためには、ソリッドステート化学センサーが必要である。実用的なソリッドステート化学センサーとして、ISFET（Ion Sensitive FET）、触媒燃焼型ガスセンサーなどがある。しかし、必要な性能を発揮するにあたっては、前者は使用前のキャリブレーションを必要とするため連続使用には向かず、後者はセンサーを高温にして表面を清浄かつ活性に保たなくてはならないため低消費電力化が難しい。このような化学センサーは明らかに IoT 用途に向かない。IoT 用途に長期間安定して使えるソリッドステート化学センサーの新しいアイデアが待たれる。

センサー・MEMS の研究開発は産業に直結し、その進歩や展開は速い。このような技術分野においては、従来型の研究開発プロセス、例えば、競争的研究費によってアカデミックを支援し、生み出

されたシーズが論文等を通じて企業に伝わり、企業による社内研究開発あるいはアカデミアとの共同研究を通じて実用化されるというプロセスは、速度の点からも成功確率の点からも十分な機能を果たせなくなっている。米国においては、アカデミアでの研究とベンチャー企業の開発とが極めて近い。欧州では、フラウンホーファー研究機構、Leti（フランス）、IMEC（ベルギー）、CSEM（スイス）、VTT（フィンランド）などの研究拠点の役割が大きい。これらの研究拠点には、先行投資して開発、蓄積された技術群を基に複数の企業が集まり、効率よく研究開発が進められている。日本では、大型プログラムによって順調に立ち上がったものの、その後の基盤的支援が行われずに縮小される研究拠点もある。産業に直結した研究開発を行うべき研究拠点に対しては、従来の競争的資金を主とする研究がそぐわないことから、欧州の研究拠点のように全体の3割程度の基盤部分を公的資金によって支える、という方式に変更していくような議論も必要であろう。

ロボットによるライフサイエンスの自動化を推進する上で、複数種類のロボット（専用ロボット、ヒューマノイド、搬送ロボット等）と各種解析機器との連携が問題である。また、全てのロボットが理解できる共通言語によるプロトコルの記述と、セマンティックにログやデータ記述を記述し、AIと連携するシステムの構築が課題である。そのためには、試薬・遺伝子／タンパク質等の分子名称を表記揺れなく取り扱い、作業オントロジーによる概念区分の導入が必須となる。

ロボットシステムは人との協働作業により進化し、より大きな知識を生み出す。このプロセスを推進するにはロボットの特性やユーザーインターフェースを熟知し、ロボットによる研究作業を指導し、データ共有の価値を啓蒙する新たな職種が必要となる。そのような人材供給を含めた全体設計が必要となる。

(5) 参考資料

- W. Chappell, R. H. Olsson III and R. G. Polcawich (DARPA), "Enabling the Next Generation of MEMS Technology", The 2018 Solid State Sensors, Actuators, and Microsystems Workshop, Hilton Head, South California, USA, June 3-7, 2018, pp. 1-2.
- 田中秀治「IEEE MEMS2018の論文から見える集積化センサの将来」『次世代センサ』(2018): 38-41.
- 南部谷俊介他「糖鎖固定化電界効果トランジスタによるインフルエンザウイルスから分離したヘマグルチニンの検出」『Proceedings of the Chemical Sensor Symposium』58 (2015): 91-93.
- 梶原正靖他「キャピラリー電気泳動法を用いた修飾核酸アプタマーのセレクション」『電気泳動』(2015): 88-90.
- Hirotaka Minagawa et al., "Selection, Characterization and Application of Artificial DNA Aptamer Containing Appended Bases with Sub-nanomolar Affinity for a Salivary Biomarker," *Scientific Reports* 7, no. 1 (2017). DOI:10.1038/srep42716
- Nozomu Yachie, Robotic Biology Consortium, and Tohru Natsume, "Robotic Crowd Biology with LabDroids," *Nature Biotechnology*, Vol. 35, No. 4, (2017) pp. 310-312. DOI: 10.1038/nbt.3758
- Shun-Ichiro Iemura and Tohru Natsume, "One by one sample preparation method for protein network analysis," *Protein interactions*, Vol. 15, 978-953-51-0244-1, InTech (2012) pp. 293-310. DOI: 10.5772/37931
- Shuhei Konagaya, Takeshi Ando, Toshiaki Yamauchi, Hirofumi Suemori, and Hiroo Iwata, "Long-term maintenance of human induced pluripotent stem cells by automated cell culture system," *Scientific Reports*, Vol. 5, Article No. 16647, (2015) pp.1-9. DOI: 10.1038/srep16647
- Jarostaw M. Granda, Liva Donina, Vincenza Dragone, De-Liang Long and Leroy Cronin,

“Controlling an organic synthesis robot with machine learning to search for new reactivity,”
Nature 559, 377-381 (2018).

- 科学技術振興機構 研究開発戦略センター 「科学技術未来戦略ワークショップ報告書 多様な安定相からの高機能材料の創製」(CRDS-FY2018-WR-11) (2019 年 3 月)

5 国内・海外の関連政策

ここではデジタルトランスフォーメーションに関係する政策上のキーワードとして、「オープンデータ」、「オープンサイエンス」、「人材育成」、「知的財産」等に着目し、日本と諸外国における関連政策や制度・事業の動向をまとめる。

5.1 日本

(1) 科学技術全般に関わる政策

科学技術基本計画に掲げられた目標を達成へ向けて、毎年度、「統合イノベーション戦略」が策定されている。2019年度の統合イノベーション戦略では、「デジタル化の流れにおいて、最も重要なのはデータである。表面的、皮相的な大量のデータの収集、活用から、質の高い産業社会活動のデータ、さらには、論文、特許等の成果に直接関係するデータに限らず研究開発の過程で得られた幅広いデータの収集、活用が極めて重要となっている」とする一方で、「今後、現実世界のアーキテクチャやプラットフォームの構築、必要な研究開発データや「深層データ」が集められる基盤づくりが競争力を左右することになる。」（第I部1総論(2)）という危機感を持った記述がされている。

(2) IT戦略に関わる政策

データに関連する政策としては、科学技術にとどまらず、IT戦略が大きな役割を果たしてきた。時間的な流れに沿って概観する。

- ・ 高度情報通信社会推進本部の設置（1994年8月）

日本において、データ基盤を含む情報通信の総合政策の取組みはここからスタートしたと言える。この推進本部は「高度情報通信社会推進に向けた基本方針」（1995年、1998年改定）を決定し、3つの行動原則（民間主導、政府による環境整備、国際的な合意形成に向けたイニシアティブの発揮）を示した。2001年1月に現在のIT総合戦略本部に移行した。

- ・ 高度情報通信ネットワーク社会形成基本法（IT基本法）（2000年11月制定）

世界最高水準の高度情報通信ネットワークの形成、電子商取引の促進、電子政府化の推進、人材の育成、安全性や個人情報の保護等を基本方針としている。このIT基本法の方針に沿って、翌2001年に「e-Japan戦略」及びその戦略を実現するための「e-Japan重点計画」が決定された。その後、重点計画は2008年まで更新されてきたが、その間に「電子政府推進計画」も作られた。

- ・ 世界最先端IT国家創造宣言（2013年6月）

オープンデータ・ビッグデータの活用推進、農業や周辺産業の高度化、ワーク・ライフ・バランス等、これまでe-Japan戦略や新たな情報通信技術戦略で掲げた主要政策をより具体的な計画とするために、府省庁の縦割りを打破して横串を通すことを狙った。また指標を使ってPDCAサイクルやガイドラインを積極的に用いることを示した。その後毎年改定されることとなった。

- ・ 官民データ活用推進基本法（2016年12月）

「オープンデータ」、「AI」、「IoT」、「クラウドコンピューティング」等の言葉を法律の中ではじめて定義した。また行政手続きのオンライン化、マイナンバーカード利活用等、幅広く具体的な内容となっている。

ここでのオープンデータとは、国や地方公共団体等が保有する公共データが、①営利目的、非営利目的を問わず二次利用可能なルールが適用されたもの、②機械判読に適したもの、③無償で利用できるもので、公開されるものを指す。

- ・ 世界最先端デジタル国家創造宣言・官民データ活用推進基本計画（2018年6月）

世界最先端 IT 国家創造宣言の名称を変え、毎年改定されている。ここではブロードバンド・ネットワーク環境などインフラの整備は順調に進んだが、電子政府やオープンデータへの取組みはまだ不十分との自己認識がされている（2019年計画・第1部I-2）。

2019年の計画では、Society 5.0時代にふさわしいデジタル化の条件として、以下を挙げている。（1）国民の利便性を向上させるデジタル化、（2）効率化の追求を目指したデジタル化、（3）データの資源化と最大活用につながるデジタル化、（4）安全・安心の追求を前提としたデジタル化、（5）人にやさしいデジタル化。

- ・ AI戦略2019、バイオ戦略2019、統合イノベーション戦略2019（2019年6月）

統合イノベーション戦略推進会議は日本のAI分野とバイオ分野の競争力を高めるために、2019年6月にこの二つの戦略を策定した。この二つに共通するのは「データ基盤整備の強化」と「人材育成の強化」（次項に説明）である。さらにこの二つを包含する形で「統合イノベーション戦略2019」がまとまった。

AI戦略2019では、重点5分野（健康・医療・介護、農業、国土強靱化、交通インフラ・物流、地方創生）における、AIの活用のためのデータ連携プラットフォームやネットワークの本格稼働をめざす。さらに収集するビッグデータの品質確認、保証に資する取組を実施するとしている。

バイオ戦略2019では、市場領域や科学の発展という観点から健康医療、バイオ素材、育種等について必要なデータ基盤を整備していく。2020年度までに全体設計（アーキテクチャ、標準化を含む）を策定する。

(1)で触れた統合イノベーション戦略2019では、上記二つの戦略にあるデータ基盤だけでなく、オープンサイエンスの文脈におけるデータの重要性を述べている。すなわち、研究データの管理・利活用の方針・計画の策定、研究データの機械判読可能化、諸外国の研究データ基盤との連携等を急ぐことを求めている。

(3) 人材育成に関わる政策

以上に述べたような科学技術やIT戦略に係わる事業を推進するためには、デジタルトランスフォーメーションの関連技術に精通した多数の人材が必要となる。それだけでなく、新しい価値の創造や技術革新（イノベーション）、起業・新規事業化、産業基盤を支える技術の維持発展、第三次産業を含む多様な業界で活躍が期待される理工系人材が多くの産業部門において強く求められている。

総務省の「ICT人材育成事業」（2016～17年）ではICT技術者を育成するためのICTスキル総合習得プログラムを開発した。

「成長分野を支える情報技術人材の育成拠点の形成（enPiT）」事業（2016年～）は産学連携による課題解決型学習等の実践的な教育の推進により、大学における情報技術人材の育成強化を目指す。大学間のネットワークによる学部学生に対する実践的教育と、IT技術者の学び直しの推進、の二つを柱にしている。

「データ関連人材育成プログラム（D-DRIVE）」事業（2017年～）は博士課程学生や博士号取得者に対してAIを含むデータサイエンティスト研修を実施して、多様な職場に高度な研究人材を送り出すことを狙っている。これまでにコンソーシアムを形成する拠点5大学を指定している。

「未来価値創造人材育成プログラム」（2018年～）は、Society 5.0に対応した高い専門性と実践的でハイブリッドな工学系人材の養成等をめざして、人材育成事業である「超スマート社会の実現に向けたデータサイエンティスト育成事業」と、その教育システム拠点形成事業である「科学技術の社会実装教育エコシステム拠点の形成事業」の二本柱からなっている。

前述した「AI戦略2019」、「バイオ戦略2019」（2019年）ではともにデータを取り扱う人材の育

成に力点を置いている。

「AI 戦略 2019」では、「数理・データサイエンス・AI」をデジタル社会の「読み・書き・そろばん」になぞらえて、小学校～大学までのリテラシー教育を充実させるとともに、社会人の再教育にも取り組むとしている。2025 年目標として、小中高校で 100 万人 / 年、大学・高専で 50 万人 / 年、社会人エキスパートを 2 千人 / 年を育成する計画である。さらに大学等の優れた教育プログラムを政府が認定する制度（「数理・データサイエンス・AI 教育認定制度」（仮称））も構築する。

「バイオ戦略 2019」においては、バイオとデジタルの融合を担うバイオ系データサイエンティストについて、必要とする人材像、その人材像に応じた育成・採用目標、そのための施策について 2020 年度をめぐって検討するとしている。

（4）知的財産に関わる政策

政府のイノベーション戦略や民間企業の研究開発戦略の上においても、オープンイノベーションの深化に伴い、研究開発成果の権利化と秘匿化を適切に使い分ける「オープン・アンド・クローズ戦略」の重要性が増してきている。一方、ビッグデータや AI がイノベーション政策の重要ファクターになっていることを反映して、データ（情報）や営業秘密などを財産として保護するため、いくつかの法律が生まれた。

「個人情報保護法改正」（2015 年）では、蓄積された個人情報を「匿名加工情報」に加工することによって、「ビッグデータ」として企業等が利用しやすくした。一方、情報収集側には「利用目的の明示」、「第三者提供の際の本人同意」など、義務が細かく定められた。

「次世代医療基盤法」（2017 年）では、改正個人情報保護法上での本人同意について特則を設ける⁹のために、匿名加工医療情報の取扱い規制を定め、医療情報の利用促進をめざした。これにより、匿名加工医療情報については本人の同意なく第三者への提供が可能となった。

（5）競争戦略に関わる政策

デジタル化が進むグローバル市場に対する国としての競争戦略やルール整備が求められてきた。そこで政府は、成長戦略の一環として、デジタル市場のルール整備によって、市場における公平な競争、データ流通の活性化、民間企業のデジタルトランスフォーメーションの推進、5G 時代に即した高度な通信環境の確保など、社会全体のデジタル化の推進に向けた次のような施策を講じるとしている。

- ・ 専門組織の設置

競争政策の迅速かつ効果的な実施を目的として、デジタル市場の評価、競争政策の企画、立案、関係機関との総合調整を担う「デジタル市場競争本部」が内閣に設置された（2019 年 9 月）。

- ・ デジタル・プラットフォーマー企業への対応

デジタル・プラットフォーマー企業の独占・寡占化を防ぎ、デジタル市場における取引の透明性を求めて、「デジタル・プラットフォーマー取引透明化法」（仮称）を法制化予定（2020 年国会提出予定）。

- ・ 個人情報への対応

個人情報の望ましくない使用について国内外無差別に防止策を適用する他、匿名化情報の一層の活用を促進する等の対策を講じるために、「個人情報保護法」改正をおこなう（2020 年国会提出予定）。

これらと並行して、データ流通のいっそうの促進をはかるために、経済産業省は日本企業の IT 環境刷新を求める報告書（いわゆる「DX レポート」）とガイドラインを発表した。この背景には急速

9 第三者への個人情報提供にあたり、改正個人情報保護法ではいわゆるオプトイン（本人の同意）が必須であるが、本法ではいわゆるオプトアウト（通知された本人が停止を求めないこと）を丁寧におこなう（初診時に書面で通知する）ことも認める。

に進むデジタル化の波に日本が取り残されかねないとの危機感がある。

(6) オープンデータに関わる政策

オープンデータとは、国や民間が保有しているデータの中で、広く無償で公開し、その二次利用を妨げないものを指す。日本においては国の電子政府推進に沿って国の統計調査データの公開と利活用が進められてきた。その中核がオープンデータ推進である。以下、公的データのオープン化について経緯をまとめる。

- 2000年11月 「IT基本戦略」の策定。電子政府の実現において、行政（国・地方公共団体）内部の電子化、官民接点のオンライン化、行政情報のインターネット公開、利用促進、の3点を掲げた。これがいわゆるオープンデータ推進の原点にある。
- 2012年7月 「電子行政オープンデータ戦略」の策定。それに沿ってIT戦略本部に設置された「電子行政オープンデータ実務者会議」で府省の保有するデータのオープンデータ公開に向けた作業を開始。
- 同上 「オープンデータ流通推進コンソーシアム」の設立。産学官が共同してオープンデータの利用ルールや技術に関する検討、普及・啓発活動を開始。2014年、一般社団法人「オープン&ビッグデータ活用・地方創生推進機構（VLED）」に改組。
- 2013年6月 「オープンデータ憲章アクションプラン」の策定。英国でのG8サミットにおける「オープンデータ憲章」の合意¹⁰を受けて、各府省情報化統括責任者（CIO）連絡会議において決定され、「キー・データのデータセット」13と「ハイバリュース・データセット」14の公開スケジュールを公表。
- 2014年10月 政府データカタログサイト DATA.GO.JP 正式運用開始。
- 2016年5月 「【オープンデータ 2.0】官民一体となったデータ流通の促進」。これまでの諸施策の基本的な考え方等を継承しつつ、課題解決型オープンデータの推進の具体的な実現を目指し、これまでの取組を更に強化させていく。国連のSDGsの取組みにも言及している。
- 2019年1月 “Data Free Flow with Trust”（DFFT）。世界経済フォーラム年次総会（ダボス会議）にて、安倍晋三首相がスピーチで、単なるデータ流通ではなく、トラストのある自由な流通を提唱した。

(7) オープンサイエンスに関わる政策

オープンサイエンスとは研究者が研究を実施した際に用いた研究データ、および成果である論文データを公開し、研究者間で共有することを指す。背景には、学術の知識基盤であるはずの論文誌の価格上昇等が、研究者・研究機関の予算を圧迫するようになってきたことがある。これまで学界での活動が地道に行われてきたが、オープンデータ化やビッグデータの取扱いが本格化してきたことに伴って、オープンサイエンスの機運も少しずつ高まっている。

オープンサイエンスは論文誌の電子化と共有の動きから始まって、2013年には Research Data Alliance (RDA) のような国際的な動きに発展している。データ共有に関しては、天文学、物理学など大規模な連携が必要とされる実験分野では早くから国際的な共有が進んでいる。具体的にオープンサイエンスが実行されている例として次が挙げられる。

10 主要10項目の議題の一つに「オープンデータ」が取り上げられ、①原則としてのオープンデータ、②質と量、③すべての者が利用できる、④ガバナンス改善のためのデータの公表、⑤イノベーションのためのデータの公表の5原則を掲げた。外務省「オープンデータ憲章（概要）」、2013.6.18

- ・ 生命科学分野：JST バイオサイエンスデータベースセンター（NBDC）
- ・ 地球環境分野：東京大学地球観測データ統融合連携研究機構を中心としたデータ統合・解析システム（DIAS）
- ・ 海洋分野：海洋研究開発機構（JAMTEC）海洋観測データ
- ・ 物質・材料分野：物質材料研究機構（NIMS）の物質・材料データベース（MatNavi）

さらに最近では研究不正の抑止手段として研究データ公開を推奨する考え方も出ている。

技術面では、論文の識別子として Digital Object Identifier（DOI）¹¹ が事実上の標準として用いられ、引用関係の把握にも役立ってきている。近年は研究データに対する DOI の付与を行うための環境も整備されている。また同様に、研究者を対象とした永続的な識別子 ORCID¹² が利用され始めている。

以下にオープンサイエンスに関わる主な動きをまとめる。

- | | |
|-------------|--|
| 2013 年 3 月 | Research Data Alliance (RDA) 設立。当初は欧州委員会 (EC)、米国立科学財団 (NSF) 等が中心となり、その後、137 カ国、約 8,800 人が登録している (2019 年 8 月時点)。「障壁なき研究データ共有」をスローガンとしている。 |
| 2015 年 5 月 | 内閣府「我が国におけるオープンサイエンス推進のあり方について」を公表。第 5 期基本計画におけるオープンサイエンス対応の重要性を明記。 |
| 2016 年 2 月 | 文科省科学技術・学術審議会が「学術情報のオープン化の推進について（審議まとめ）」を公表。公的研究資金による研究成果を原則公開すべきとした。 |
| 2016 年 7 月 | 日本学術会議「オープンイノベーションに資するオープンサイエンスのあり方に関する提言」を公表。①研究分野を超えた研究データの管理およびオープン化を可能とする研究データ基盤の整備、②研究コミュニティでのデータ戦略の確立、③データ生産者およびデータ流通者のキャリア設計、を重要視している。 |
| 2017 年 12 月 | 内閣府に「国際的動向を踏まえたオープンサイエンスの推進に関する検討会」設置。 |
| 2018 年 6 月 | 上記検討会が「国立研究開発法人におけるデータポリシー策定のためのガイドライン」公表。 |
| 2019 年 3 月 | 上記検討会が「研究データリポジトリ整備・運用ガイドライン」公表。 |
| 2019 年 10 月 | 上記検討会が「研究データ基盤整備と国際展開に関する戦略」を公表。国全体の研究データ管理及び利活用に関する基本方針（ナショナル・データ・ポリシー）を示す。我が国の研究データを共有・公開するための基盤システム、国レベルでの研究データ戦略やデータポリシー、国際的な対応等について報告している。特に競争的研究資金制度において、研究データを適切に管理・共有・公開するために「データ・マネジメント・プラン（DMP）」を導入することを推奨した ¹³ 。 |

11 インターネットで使用される Uniform Resource Locator (URL) はデータの所在を直接示すため、時間とともにデータが消滅・移動した時には追従できない。そのためデータの場所を恒久的に示す標識として DOI が用いられるようになった。DOI は国際 DOI 財団が指定した DOI 登録機関名と、その登録機関が個々の論文や研究データに付与した一意の番号の組み合わせから構成される。日本の登録機関はジャパンリンクセンター (JaLC) である。

12 Open Researcher and Contributor ID. 論文や研究データに紐付いた研究者の永続的な個人識別子である。同姓同名の人を区別したり、改名した人を同一視したりすることが可能である。https://orcid.org (2020年3月1日閲覧)

13 これに先立ち、国立研究開発法人日本医療研究開発機構 (AMED) は DMP の提出を義務化した (2018年3月)。

(8) 参考資料

- ・ 知的財産戦略本部「知的財産政策ビジョン」(2013年6月7日)
- ・ 首相官邸「デジタル市場のルール整備」 https://www.kantei.go.jp/jp/singi/keizaisaisei/portal/digital_rule/ (2020年3月1日閲覧)
- ・ 経済産業省「DX デジタルトランスフォーメーション レポート～ITシステム『2025年の崖』の克服とDXの本格的な展開～」、2018年9月 https://www.meti.go.jp/shingikai/mono_info_service/digital_transformation/20180907_report.html (2020年3月1日閲覧)
- ・ 経済産業省「DX推進ガイドライン Ver. 1.0」、2018年12月 <https://www.meti.go.jp/press/2018/12/20181212004/20181212004.html> (2020年3月1日閲覧)
- ・ 科学技術振興機構 科学技術情報委員会「わが国におけるデータシェアリングのあり方に関する提言」(2015年4月)
- ・ バイオサイエンスデータベースセンター「データベース統合を通じて新たな知識へ」 <https://biosciencedbc.jp/> (2020年3月1日閲覧)
- ・ 一般財団法人リモート・センシング技術センター (RESTEC)「データ統合・解析システム」 <http://www.diasjp.net/> (2020年3月1日閲覧)
- ・ 国立研究開発法人 海洋研究開発機構データベース <http://www.jamstec.go.jp/j/database/> (2020年3月1日閲覧)
- ・ 国立研究開発法人 物質・材料研究機構「MatNavi N I M S 物質・材料データベース」 <http://mits.nims.go.jp/> (2020年3月1日閲覧)

5.2 米国

世界的なデジタル化の潮流の中、米国でも連邦政府全体としての対応が優先課題と認識されており、関連インフラや制度の整備が進められている。その中で、デジタル時代における国家安全保障や産業競争力の優位性確保に不可欠な、AI やコンピューティング技術を活用したデータ駆動型科学に関する取り組みも進んでいる。

2018 年 3 月に発表された「大統領の政策運営アジェンダ」では、連邦政府機関が横断的に取り組む優先項目の一つとして「戦略的資産としてのデータの活用」を掲げ、連邦政府全体のデータ戦略を整備して、米国の市民、企業、研究者が政府データを利用しやすくするとともに、政策立案やイノベーションにおけるデータ活用を向上させるとしている。これを踏まえて、2019 年 6 月には「連邦データ戦略」とその「アクションプラン（案）」が作成され、オープンデータを中心とした連邦政府のデータ戦略を調整する「データ会議」の設立や、データの管理と使用のための倫理枠組みの策定、AI の研究開発のためのデータ資源の改善などが盛り込まれた。さらに 2020 年 1 月には同戦略の実施のための「アクションプラン 2020」が確定した。「アクションプラン」は今後定期的に更新される予定である。また、大統領府が各連邦政府機関に対して通達する「2021 年度研究開発予算優先事項」（2019 年 8 月）においても「データの活用」が重点項目の一つとして挙げられ、データのアクセシビリティとセキュリティの改善、AI を含む新興技術の活用、データのスキルを有する労働力の育成等に優先的に取り組む方針が示された。

分野別の研究開発戦略でも、データの活用が強調されている。2018 年 10 月に発表された「先進製造における米国リーダーシップ戦略」では、デジタル化を反映した新たな製造技術やシステムを筆頭課題に挙げている。特に、堅牢な産業用モノのインターネット（IIoT）、広範な製造プロセスに適用できる機械学習アルゴリズム、および統合された情報システムに自動接続する機械ツールなどに関してイノベーションが必要とし、優先すべき技術課題としてスマート / デジタル製造、高度な産業用ロボット、AI 対応インフラ、サイバーセキュリティを特定している。

AI に関しては、2019 年 2 月に大統領府主導で「米国 AI イニシアチブ」が打ち出されており、研究開発、人材育成、基盤整備（データ、インフラ、規制、標準化等）への集中投資と、国際枠組みにおける米国 AI 企業への市場開放と国益確保の両立という方針が掲げられている。また、2019 年 6 月にはオバマ政権時に策定された国家 AI 研究開発戦略の改訂版が発行された。同改訂版は従来の戦略（研究開発、人材、倫理・セキュリティ等）を踏襲した上で、「官民パートナーシップ拡大」を新たな取組事項として追加している。

これらの先進製造や AI に係る戦略においては、特に専門人材やスキルを持った労働力の開発といった観点で人材育成の重要性が述べられている。一方で、2018 年 12 月に発表された「STEM 教育戦略計画」では、初等・中等教育も含めた長期的な STEM 人材育成の方針がまとめられている。同戦略計画では、推進方策の一つとして「コンピューター・リテラシーの育成」を掲げ、デジタル・リテラシーと安全なサイバー利用スキルの向上、全ての教育段階および科目における計算科学的思考の養成、教育用デジタルプラットフォームの整備などを推進すべきとしている。

このような政権の考え方を踏まえつつ、各連邦政府機関では、既存および新規の研究開発を加速させている。

情報科学技術分野における機関横断型の枠組み「ネットワーキング・情報技術研究開発（NITRD）」では、大規模データ管理、コンピューティング、インテリジェントロボティクスなどのプログラム構

成領域（PCA）を定め、毎年機関ごとの予算割り当てを調整するなど、戦略的な研究開発投資を進めている。2020年度からはPCAの一つに「AI」が新たに設定され、この分野への注力が顕在化している。2020年度のNITRD予算に対する大統領府の意向として、全体で55億ドル、うちAI関連で9.7億ドルが予算教書ベースで示されている。なお、いずれの金額もDODおよびDARPAのAI関連予算は非公開のため含んでいない。こうした戦略的投資の下、各機関において様々な研究開発プログラムが実施されている。

表 5-2 情報科学分野における連邦政府機関の研究開発予算の規模（単位：億ドル）

省庁・機関	2020年度予算※1 (予算教書ベース)	(うちAI関係)
国防総省(DOD) (DARPA含む)	15.7※2	--- ※3
国立科学財団(NSF)	13.0	(4.9)
国立衛生研究所(NIH)	11.6	(2.0)
エネルギー省(DOE)	9.8	(1.6)
国立標準技術研究所(NIST)	1.7	(0.2)
航空宇宙局(NASA)	1.4	(0.0)
国土安全保障省(DHS)	0.7	(0.3)
その他	1.2	(0.6)
計※4	55.1	(9.7)

出典：NITRD Supplement to the President's FY2020 Budget を元に CRDS 作成

※1 実際の予算編成・審議は議会が行うため、確定した数字ではない

※2 DOD、DARPA の AI 関連予算は含まない

※3 DOD、DARPA の AI 関連予算は非公開

※4 端数処理のため合計は一致しない

データ駆動型の研究開発に関しては、横断的・総合的な戦略は策定されておらず、各機関で個別に取り組まれている。例えば NIH は 2018 年 6 月に「データサイエンスのための戦略計画」を発表している。これは NIH が支援する研究から生まれたデータの保存、管理、標準化、公開を、インフラや人材の観点も含めて総合的に推進するためのロードマップとなっている。

研究データの収集という面では、いわゆるオープンサイエンス推進の文脈において、2003 年に NIH、2011 年に NSF から先行的に公的資金による研究成果（データ）の共有促進のための方針が示され、2013 年の科学技術政策局（OSTP）の指令により連邦政府機関全体へ拡大した。一方、データの保管や共有・アクセスに関しては分野別、機関別の多様なリポジトリが充実しているが、これらを横断した統合的な研究データリポジトリ整備を推進するような政策は示されていない。米国全体では、大学や研究機関によって構成される全米データサービス（NDS）や非営利団体であるオープンサイエンスセンター（COS）などが研究データ活用のための環境やツール、フレームワークを提供しており、連邦政府もこうしたボトムアップの取り組みを適宜支援している。2020 年 1 月には OSTP により、連邦政府から助成を受けた研究成果のデータを管理・共有するためのリポジトリの望ましい特性についてのパブリックコメントの募集が開始されている。

これら連邦政府機関による研究成果を社会実装につなげてイノベーションを実現するためには、民

間セクターと効果的に連携していくことが重要となってくる。特に米国においては GAFA に代表される IT プラットフォーマーの存在感は極めて大きく、連邦政府はこうした企業の人材・資金・データの活用も念頭に、官民パートナーシップの拡大による新たな米国全体の研究開発エコシステム構築を優先課題の一つとしている。

また、G20 や G7 等の国際枠組みを通じてデジタル経済の障壁緩和やデータ流通促進に取り組み、米国・メキシコ・カナダ協定（USMCA）におけるデジタル貿易の強化やデータ流通への国の介入を極力排除した日米デジタル貿易協定の締結など、企業のイノベーションを阻害しないルール作りを積極的に進めている。他方で、これらのプラットフォームによるデータ寡占が適切な市場競争を妨げているとの指摘もあり、司法省や連邦取引委員会による調査と監視が強まりつつある状況も見られる。

5.3 欧州連合（EU）

（1）重要政策

2014 年 11 月に欧州委員長に就任した Jean-Claude Juncker 氏の元、EU は 2015 年 5 月にデジタル単一市場戦略 (Digital Single Market Strategy : DSM) を発表した。EU では移動の自由や関税撤廃、通貨統合等により市場の単一化を進めてきたが、デジタル分野でもそれを実現しようとするものである。EU 域内にデジタル単一市場を構築することは、Juncker 氏が掲げた、今後 5 年間で欧州委員会が取り組む成長のための 10 の優先課題の一つであった。

同戦略は、①欧州全体の消費者・企業のデジタルグッズ・サービスへのアクセス改善、②デジタルネットワーク・サービスにとってより適した環境の創出、③デジタル経済・社会の成長可能性の最大化、という 3 つの柱と、各柱に位置づけられる全 16 個の主要施策によって構成されている。

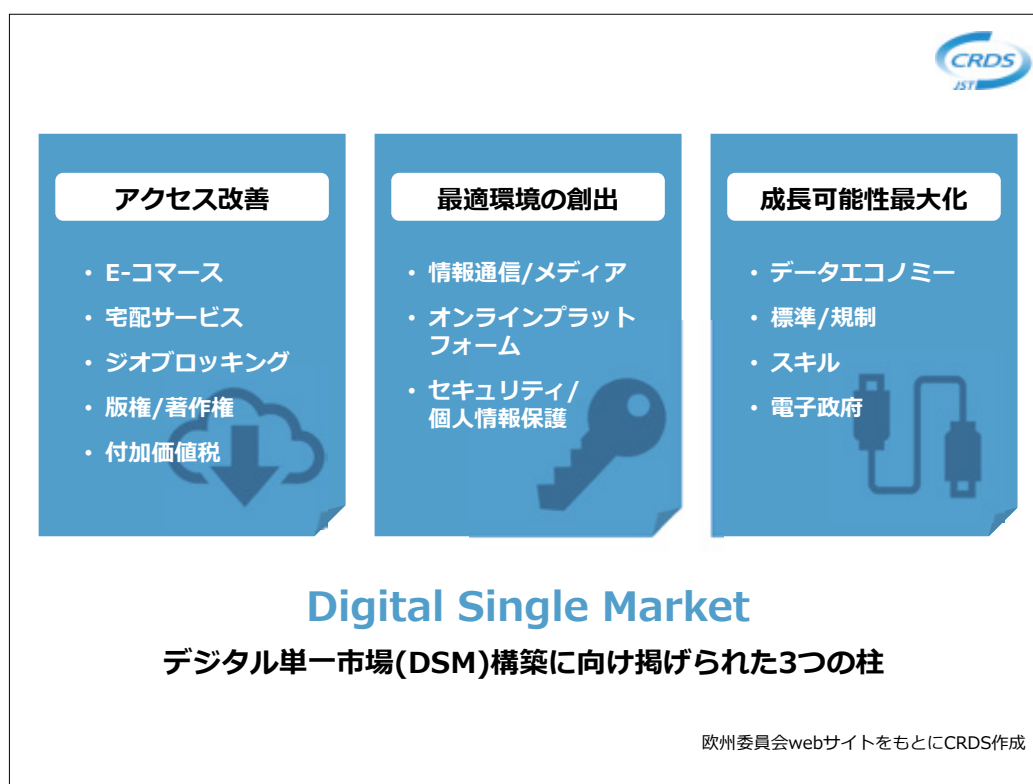


図 5-3 EU のデジタル単一市場 (DSM) 構築に向け掲げられた 3 つの柱

2017 年 5 月には本戦略の中間評価が行われ、今後さらなる取り組みが必要な 3 つの分野として、「データエコノミーの最大限の活用」「サイバーセキュリティ強化による欧州の資産保護」「オンラインプラットフォームの推進」が特定された。Juncker 氏の着任以降、DSM に関する 30 の法案が提案され、28 が成立に至っており、デジタル市場の単一化が着実に進められている。

2019 年 12 月 1 日より Juncker 氏の後任として新たに欧州委員長として就任した Ursula von der Leyen 氏は、2019 年 7 月に「A Union that strives for more My agenda for Europe」と呼ばれる政治的ガイドラインを発表した。これは、彼女の就任期間である 2019 年～ 2024 年の 5 年間で取り組む 6 つの優先課題を示したものである。優先課題の一つに「デジタル時代に適応する欧州 (A Europe fit

for the digital age)」があり、この中で、就任後 100 日以内に AI の人間的・倫理的意味合いに関する欧州の統合的アプローチを法制化することや AI への投資重点化、さらには委員会の完全デジタル化などが示されている。

(2) 欧州クラウドイニシアチブ

デジタル単一市場政策の一環として、EU は 2016 年 4 月に「欧州クラウドイニシアチブ」を発表した。このイニシアチブの中で、「欧州オープンサイエンスクラウド (EOSC)」を構築する方針が打ち出されている。EOSC では、科学、技術、人文社会科学の 170 万人の欧州の研究者と 7,000 万人の専門家が、欧州の広範な国家、地域、機関の公的研究インフラから、増大する膨大な量の公開データやその他のリソースに自由にアクセスできるようにすることを目指している。2020 年までの完成を目指し 6 億ユーロを投じるとしており、2018 年 11 月には最初のサービスやリソース群へのアクセスを可能にするポータルサイトがオープンした。

また、高性能コンピューターの供給も重要課題として位置づけられており、高性能コンピューティング共同事業 (EuroHPC JU) という官民共同の取り組みが 2018 年 11 月に始まった。これは世界トップクラスの汎欧州スーパーコンピューター新設を目指すものであり、2022 年～2023 年に世界 5 位以内に入るエクサスケール・コンピューターを 2 機 (うち最低 1 機は欧州の自前技術で) 作るという目標が掲げられている。2019 年、2020 年の 2 年間で欧州委員会および共同事業への参加国による公的投資 10 億ユーロと民間投資 4.2 億ユーロが提供される。2021 年以降は、後述の「Horizon Europe」、「デジタル・ヨーロッパ」から公的な資金提供を受ける予定である。

(3) 研究開発支援施策

2014 年から 2020 年までの 7 年間を対象とする研究開発枠組みプログラム「Horizon 2020」の「産業リーダーシップ」という柱において、ICT は 6 つのキーテクノロジーのうちの 1 つに指定されており、7 年間で約 58 億ユーロの予算が措置されている。支援対象となる具体的な研究開発分野としては、AI、5G、次世代インターネット、サイバーセキュリティ、ロボティクスなどがある。

また、2021 年からの 7 年間を対象とする次期研究開発枠組みプログラム「Horizon Europe」(全体予算として 941 億ユーロが提案されており現在審議中) の三本の柱の一つである「地球規模課題と欧州の産業競争力」の中に、「デジタル・産業・宇宙」というクラスターが設けられている。量子技術を含む重要デジタル技術、AI・ロボティクス、次世代インターネット、先端コンピューティング・ビッグデータといった分野の研究開発に重点投資が行われる予定である。

加えて、高性能コンピューター (HPC)、AI、サイバーセキュリティ等の機能強化に必要なインフラ構築を目的とした新プログラム「デジタル・ヨーロッパ」を 2021 年から開始すべく調整が進んでいる。現在提案されている予算は 7 年間で 92 億ユーロであり、スーパーコンピューターに 27 億ユーロ、AI に 25 億ユーロ、サイバーセキュリティ・信用に 20 億ユーロを充てるとされている。

(4) 参考資料

- European Commission, “A Digital Single Market for Europe: Commission sets out 16 initiatives to make it happen”
https://ec.europa.eu/commission/presscorner/detail/en/IP_15_4919

5.4 英国

英国では2017年11月にビジネス・エネルギー・産業戦略省（BEIS）から「産業戦略：将来に適応する英国の建設」が新たに発表された。名称は産業戦略であるが、同戦略は単なる産業振興のための施策や提言に終始するものではない。2030年までに英国を世界最大のイノベーション国家にすることを目指して科学技術政策にも着目しており、生産能力向上などの長期構想を示した内容である。

この新産業戦略は、科学技術・イノベーション政策も包含する内容となっており、英国がグローバルな技術革命を主導できる領域として「AI・データ経済」をはじめとする4つの「グランド・チャレンジ」を特定し、2018年5月には各グランド・チャレンジの具体的な目標として「ミッション」を明らかにした。これらミッションは特定の問題に焦点を当て、政府、産業界、様々な団体など英国内の多様なステークホルダーが団結して現実には人々の生活を変えることを目指している。表5-3に、グランド・チャレンジおよびミッションについて示す。

表5-4 英国の産業戦略におけるグランド・チャレンジとミッション

グランド・チャレンジ	ミッション
AI・データ経済	データ、AI、およびイノベーションを用いて、病気の予防、早期診断、および慢性疾患の治療を2030年までに変える
高齢化社会	富裕層と貧困層の格差を縮め、2035年までに今よりも少なくとも5年間長く人々が健康で自立した生活を送れるようにする
クリーン成長	2030年までに新しい建物のエネルギー利用を少なくとも現在の半分にする 2030年までに低炭素の産業クラスターを少なくとも1つ確立し、2040年までに世界初となる温室効果ガス純排出量ゼロの産業クラスターを確立する
未来の輸送手段	英国をゼロエミッション車（ZEV）のデザインと製造の最先端に位置付け、2040年までに自動車とトラックすべての新車を事実上排出ゼロにする

AI技術は2030年までに英国のGDPの10%（2,320億ポンド）にあたる経済効果を生むと見込んでいる。2018年4月に英国政府と関連産業界は「AIセクター協定」を締結した。本協定では、AI分野で官民合わせて10億ポンドの投資を実行することをはじめ政府と産業界で取り組むべき事項が盛り込まれた。本協定の締結から1年が経過し、次のような成果が生まれている。

- ・ グランド・チャレンジやセクターディールの実行に責任を持つAIオフィスを政府内に創出
- ・ データ・倫理・イノベーションセンターの設立
- ・ 今後5年で1,000名の博士課程修了生を出すため全国で新たに16の博士教育センターが設立

5.5 ドイツ

(1) 政府のデジタル化推進政策

2013年に発足した第三次メルケル内閣（2013-2017）の連立与党政権公約に、デジタルアジェンダ 2014-2017（Digital Agenda 2014-2017）の策定が明記された。連邦経済エネルギー省（BMWi）、連邦内務省（BMI）、連邦交通デジタルインフラ省（BMVI）を所管省とし、経済や産業のインフラ構築だけでなく、教育と研究のデジタル化、市民生活の安全性と労働の革新などに関する包括的なガイドラインとして翌年発表された。その後2016年に後継となるデジタル戦略（Digital Strategy）に更新されたが、ドイツのデジタル化政策の基本方針はこのデジタルアジェンダに基づいている。これに先立ち、ドイツ工学アカデミー（acatech）と連邦政府の有識者会合である研究連合（Forschungsunion）は、本格的なIoT時代に備え、第四次産業革命に対する実行戦略提言（Recommendations for implementing the strategic initiative INDUSTRIE 4.0）を2013年に出した。これを受けて施策化された製造業のデジタル化を目指す「未来プロジェクト インダストリー 4.0（Zukunftsprojekt Industrie 4.0）」にとどまらず、輸送・交通、エネルギー供給、医療・介護、金融などの産業セクター横断的にデジタルプラットフォームの構築を推進する施策も実行されている。

(2) 科学技術基本政策「ハイクテック戦略」

科学技術イノベーションの基本政策は「ハイクテック戦略（High-Tech Strategy）」で、2006年から施行されている。2014年に出された第三期にあたる「新ハイクテック戦略（New High-Tech Strategy）」では、第三次メルケル内閣の連立与党政権公約を受けてデジタル化を6つの重点分野の1つに特定した。2018年の第四期「ハイクテック戦略 2025（High-Tech Strategy 2025）」では重点分野として特出しされなかったものの、全ての重点分野基礎となるインフラとしてデジタル技術を位置づけている。

第四期ハイクテック戦略とほぼ同時期に連邦政府は「人工知能戦略（Künstliche Intelligenz Strategie）」を発表、研究開発の推進協議会としてプラットフォーム（Platform Lernende Systeme）を構築、産学官が連携し、AI、ビッグデータ、ソフトウェアの各領域の知識共有を図り、課題解決に取り組んでいる。

また、研究開発分野のデジタル化については、2019年にBMBFから指針となる「デジタルの未来（Digitale Zukunft）」が出されている。さらに2018年、16の州と連邦政府の文部科学大臣会合（GWK）でドイツ研究データインフラストラクチャ（NFDI）構築が決まった。連邦と州が共同でNFDIに助成を実施、参加コンソーシアムの公募が2019年に始まったところである。NFDI構築の目的は、従来の研究データは分散的で時限的に保存されていたが、これを共通の基盤上に集積して「使えるデータ」にすることで研究開発を推進するものである。2019年から2028年の10年間に9,000万ユーロ/年を限度額に助成が予定されている。計画では30の大学や研究機関を単位としたコンソーシアムを採択し組織横断的なデータ収集と利用機会の提供ができるようにする。公募のレビューはドイツ研究振興協会（DFG）が担当し、GWKがDFGの評価に基づいて採択を決める。

(3) 具体的な研究開発助成プログラム

BMBFによる主要な研究開発助成プログラムとしては、2007-2017年に実施された「ICT2020-イノベーション創出のための研究」があげられる。ICT分野の基礎技術、電子技術とマイクロシステム、ソフトウェアとデータ処理・通信技術とネットワークが対象領域となり、10年間で34億ユーロが投資された。基礎的な研究だけでなく、サービスプラットフォーム構築や中小企業の研究開発支援も含む包括的なファンディングである。2013年以降は上記に触れたインダストリー 4.0 プロジェクト

トのスタートもあってデジタル化における中心的な施策として位置づけられていた。現在は、医療分野のデータ科学助成プログラム（German Medical Informatics / 2016-2025 年）や生活支援ロボットの開発など市民生活に直結したデジタル技術（Bringing Technology to the People/2016-2020 年）への助成が実施されている。

5.6 中国

(1) 科学技術基本政策

中国の科学技術における最上位基本政策ともいえる「国家イノベーション駆動発展戦略綱要（2016～2030年）」では、「イノベーション及びオリジナリティの強化」の項目で、「データ資源」などのプラットフォームを早急に整えることを掲げている。「国家イノベーション駆動発展戦略綱要」等の上位政策を受けて策定された国全体の科学技術実施政策ともいえる「科学技術イノベーション第13次5年計画（2016～2020年）」において、次世代情報技術関連では「インテリジェント製造とロボット技術、次世代情報通信技術、マン・マシン融合に向けた情報通信技術、量子制御と量子情報技術」などと共に、「ビッグデータ技術」が重点領域として挙げられている。

(2) 官民共同推進体制

中国における情報科学の研究開発は、特に応用研究において産業界の存在が大きい。通称BATと呼ばれる百度（バイドウ：Baidu）、阿里巴巴集団（アリババ：Alibaba）、騰訊（テンセント：Tencent）などの企業における情報技術の応用研究開発が盛んである。現在では、上記3社に加え、科大讯飞（アイフライテック：iFlytek）と商湯集団（セNSTタイム：SenseTime）が政府により次世代人工知能オープンイノベーション・プラットフォームに認定され、政府の後押しにより官民共同で研究開発を推進している。ここでは、バイドウを自動運転、アリババをスマートシティ、テンセントを医療画像認識、アイフライテックをスマート音声、セNSTタイムをAI画像処理技術の各領域における国家次世代人工知能オープンイノベーション・プラットフォームに認定している。これらの企業にそれぞれの領域のプラットフォームを整備させ、各領域のデータを政府から認定された一社に集約することにより、より強力な情報産業力を生み出している。

これらの施策は、2017年7月に国務院から発表され、2030年までに中国がAIにおいて基礎研究から産業全てにおいて世界をリードする立場となることを掲げている「次世代人工知能発展計画（通称：AI2030）」を受けて科学技術部により2017年11月に組織された「次世代人工知能（AI）発展計画及び重大な科学技術プロジェクト始動会」により遂行されている。

(3) データ科学関連の倫理規定

一方で、情報科学におけるガバナンス面でも動きがあり、2019年5月、科学技術部と北京市政府が支援する北京智源人工智能研究院（Beijing Academy of Artificial Intelligence, BAAI）が「北京AI原則（Beijing AI Principles）」を発表した。「人間のプライバシー、尊厳、自由、自律性、権利が十分に尊重されるべき」など、AIの研究開発に関する指針を説明している。同年6月には、科学技術部が「責任あるAI技術」の原則を述べた「新世代人工知能ガバナンス原則（Governance Principles for the New Generation Artificial Intelligence）」を発表している。ここでは、「人工知能研究開発者は無論、使用者、管理者等その他関係者も含む各方面の関係者は、高度な社会的な責任感と自律意識を持って、法令及び倫理道德と標準規範を厳格に守り、人工知能を利用して違法な活動を行うことを防止するべきである」と述べられている。

6 まとめ

1989年の冷戦終結ののち、情報通信革命の流れに伴い、国際化、情報化という大きなフェーズチェンジが起こった。2006年以降に、ビッグデータ、サイバーフィジカルシステムズ（CPS）、IoTといった概念が立て続けに出現した。当初は、インターネット上のサービスに集まるデータを収集・解析し、そのサービスを改良・強化するに過ぎないようなサイバー空間に閉じたパイプラインであった。現在では、本報告書の各事例で見て取れるように、センサーを含むさまざまなIoTデバイス、あるいはそれらを実装した機械・ロボットが導入され、実世界（フィジカル空間）からリアルワールドデータを収集し、ディープラーニング（深層学習）を含む機械学習の解析結果に基づいて、新たな科学的知見の発見や、実世界のシステムにフィードバックをかけるような応用へ広がっていることがわかる（デジタルトランスフォーメーション）。分野にもよるがこうした動きは第3次AI革命（2012年～）以降に顕著であり、まだまだ新しい動きと言える。

21世紀には、絶えず変動するオープンな複雑系システム（ヒト等の生命システムから環境・社会システムまで）を研究対象として取り上げることが増えており、そういったシステムは確率論で論じることが合理的である（論じざるを得ない）ことから、データ駆動型、デジタル技術と相性が良く、今後もこの方向は進展していくことが予想される。

各節の事例を俯瞰すると、こうしたデジタル技術は、科学研究における新たな発見、研究開発等における時間とコストの削減、スマート・持続可能な社会の構築、といった目的のために使われ始めていることが見て取れる。特に、第2章で取り上げた自然科学の研究分野では創薬や材料開発プロセス等においてデジタルトランスフォーメーションが着実に取り入れられ始めている一方で、第3章のような社会システムとより深く関係し、自然科学と人文学・社会科学の知見を総合的に活用する研究分野においてはまだまだ要素毎の進展に留まっていることもわかる。

こうしたデジタルトランスフォーメーションによって、マクロな潮流として、3つのつながりが進行・加速していくことが読み取れる。

1つ目は基礎的な研究と社会とのつながり（仮説の立証と社会における実装）である。基礎研究から見出された知見や技術シーズは、実用化と小規模な実践を経て社会へ実装されたのち、改めて社会の中でその意義や効果が評価・検証され、新たな課題の抽出、仮説の設定へとつながり、それらが再び基礎研究へと還元されるような循環構造を持っている。このサイクルが、デジタル技術やビッグデータの出現・活用によって、それ以前と比較して非常に早く回るようになってきている。

2つ目は、デジタル化に伴うディシプリン（学術の専門体系）の横断によるつながりの加速である。その背景には例えば、様々なデジタル計測機器等から排出されるデータを統合解釈することが可能になってきていることが挙げられる。デジタル時代の科学技術研究は、着実にチーム型、オープン型、アジャイル型へとようになっていく。そういった世界的な事象から、日本の伝統的な研究文化は、すべてではないものの取り残されている。

3つ目は、社会や産業の場におけるサービス、価値が互いにつながりを見せ始めていることである。例えばエネルギーに関しては、気候変動緩和型の低消費エネルギー社会へと向かい、ヘルスケアでは予防による健康寿命の延伸・医療費の削減といったように、各課題の主体となる行政、企業、家庭、個人のデータが存在している。エネルギーや交通システム、ヘルスケアのような分野では特に、国民・消費者・需要家行動を考慮したシステムへの転換が図られていることが特徴である。これらのデータを掛け合わせることで、今後新たな価値が創出される可能性が示唆されている。これは産業構造の変革にもつながる。

これら3点のみに鑑みても、研究開発や研究成果の社会実装、教育・人材育成のあり方の変革の必

要性が認められる。

一方で課題も浮かび上がる。現在の AI ブームの引き金となった、第 3 世代 AI の中核であるディープラーニング技術の登場によって、さまざまなタスクに対し大幅な精度向上を達成し、上記の 3 つの進展が加速してきたが、限界・問題も指摘されている。具体的には、①学習に大量の教師データや計算資源が必要、②理論的に未解明で、学習結果の信頼性が確保できていない、③実世界環境・文脈・常識等の理解や適応ができていない、といった問題である。さらにこれはデータ取得側の課題として、既存のデータ群（ラベル・アノテーション含む）には信頼性や様式不統一の問題が大きい。

このようなデジタル社会において無視できないのが、大量データ収集のプラットフォームやキーデバイスを持つ GAFA の存在である。中国では、政府方針によりある産業分野データについては特定の企業に独占的に集約させ、AI によるソリューションの実用化を加速させようとしている。AI を扱う技術者が全世界的に不足している中、GAFA や中国 AI リード企業 5 社（Tencent、Baidu、Alibaba、iFlytek、SenseTime）が囲い込もうとする状況もあり、この分野でいかに日本の存在感を示していくかは大きな課題である。

また、本報告書の各事例の課題には、およそ共通した項目が並ぶ。各研究分野におけるデータ連携プラットフォーム創出や人材育成、異分野連携である。これらは大学・国研や科学技術のあり方そのものが関係してくる。国際標準化戦略などの出口戦略も重要である。

以上に述べたような、本報告書から得られた大きな潮流を踏まえた上で、以下では共通的に抽出された今後の方向性と課題を、研究開発システム、社会実装システム、教育・人材育成・開発システムの 3 つの視点から、これまでの CRDS の知見も活用しつつまとめた。

1. 研究開発方法論の変革

● 問題解決のための両輪となるデータ駆動アプローチと計算理論的アプローチ

現状、本報告書の各事例で見て取れるように、様々なデータベースからデータを抽出する方向性と同時に、センサーを含むさまざまな IoT デバイス、あるいはそれらを実装した機械・ロボットが実世界（フィジカル空間）からデータを収集し、ディープラーニング（深層学習）を含む機械学習の解析結果に基づいて、新たな科学的知見の発見や、実世界のシステムにフィードバックをかけるような応用へ広がっている。

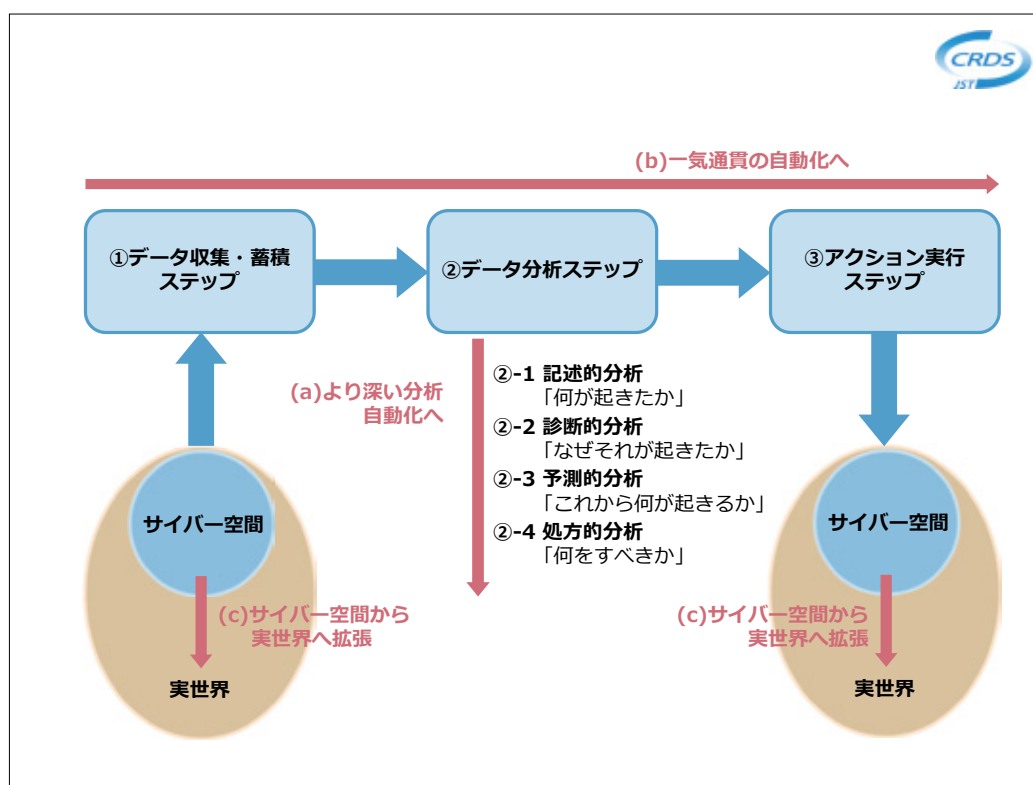


図 6-1 問題解決パイプラインとその技術的発展の方向

このような過程で得られる信頼性の高い実世界データに基づく AI は、日本が強みを出し得る方向性である。

技術的には容易に多様な大量の情報が得られるようになってきたが、第一線の研究や社会実装に必要な量と質の情報を得るための金銭的・労力的コストは従前と変わらず、研究実施にあたっての障壁となっている。また、第3世代 AI をもってしても、実世界環境・文脈・常識等の理解や適応ができていないこともデータ駆動型研究の進展を妨げている。現時点では研究現場やフィールドから発生するビッグデータによる知識の構造化はヒトが介在せざるを得ず、シソーラス、アノテーションやオントロジー構造等の辞書を作成することが求められている。科学技術・研究開発にかかる文書等をコンピューターに理解させ、自動で概念や文脈等を推測することが可能になれば、デジタル化した科学技術知識を社会に実装することができるようになる。後に述べる、マルチモーダル計測技術の標準化やデータプラットフォームの構築が重要になっている。

将来的に、実世界で起きる現象・活動の状況の精緻かつリアルタイムの把握・予測、膨大な選択肢の網羅的検証、大規模複雑タスクの自動実行等が可能になり、科学の方法論の革新、産業構造・社会システムの転換等がもたらされ得る。科学の方法論の革新としては、高度な計測機器やセンサー、実験ロボットが高信頼のデータをハイスループットに産出し、これらのデータに基づき人工知能が仮説の立案（相関・因果の提示）を行い、全自動あるいは半自動でモデルに落とし込み（データ駆動型アプローチ）、次世代スーパーコンピューターによって仮説の検証を実行する（計算理論的アプローチ）、また実験（観測・計測）データが追加（データ同化）されて、このサイクルを回す、といった科学技術基盤も考えられる。

● 計測技術、センサー・ロボット技術の重要性（AI 技術との融合によるデータ取得・収集）

実世界の様々な課題（顧客課題や社会課題）や科学的問題に対して、実世界から得られるビッグデー

タを AI 技術で解析することが、課題解決の有効な手段になっている。こと科学技術、研究開発に関して、デジタルトランスフォーメーションの中心にあるのは、計測技術等データを取得する技術の進展である。

狭義の計測技術はセンシングに相当する。しかし、昨今、センシングとその後の分析（AI）やプランニングとの融合が起こりつつある。この融合によって、計測技術をより広義のものとして捉えるべき流れが生まれている（図 6-2）。このような AI 技術との融合による計測技術の高度化は、計測機器の高度化に直結するだけでなく、計測技術を様々な課題解決や価値創出につなげることにもなっている。

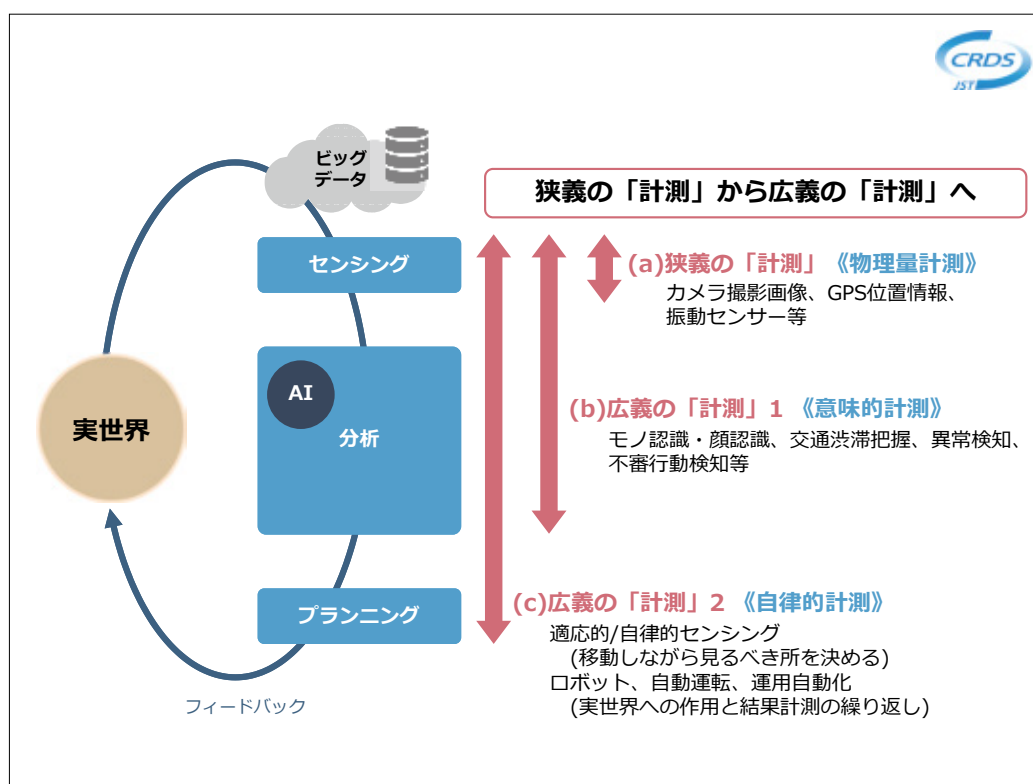


図 6-2 「ビッグデータ×AI」型ソリューションにおける「計測」の位置付け

計測技術自体（つまり計測されたアナログ情報のデジタル処理）が機械学習とあいまって、様々な進化を遂げている。もちろん機器からの産物であるビッグデータの解析は言うまでもない。

計測技術そのものについては、その高度化×多面化×広域化によって、実世界の状況・状態を広く深く早い把握を可能にすることが技術チャレンジの方向性であろう。計測の多面化では、例えば、人間の五感を超える「超五感」や、人間の内面や意思の把握等が目指されつつある。計測の広域化は、小型、低消費電力、安価、メンテナンス容易等の価値と連動して実現されていく。計測の高度化は、意味的計測や自律的計測を実現していくものである。そして、それら計測の高度化×多面化×広域化の総合的な形として社会計測が発展していく（社会計測では人間の行動情報に関するプライバシー保護にも配慮が必要になる）。個々の要素技術として新たな計測手法・機器の創出・発展とそれらを統合した広義の計測技術の高度化の両面から取り組みが必要である。

繰り返しになるが、信頼性の高い実世界データに基づく AI は、日本が強みを出し得る方向性である。日本は計測技術などでは強みを有するところも多く、マルチモーダルな計測から産出されるデータを

つないでいくことが期待される。実際に組み合わせて利用すると精度不足、技術同士がかみ合わないといったことが生じるため、要素技術のブラッシュアップ、すり合わせ、標準化を意識した技術開発が重要である。そのためには少量のサンプルでの検討を可能にする微細加工技術や、工程の高速、高精度、自動化に向けた制御工学の分野との連携が必要不可欠である。

以上を充足するためには、分野毎に研究環境や計測機器群をオープンプラットフォーム化して、オンデマンドで目的のデータをとる環境の構築も視野に入れるべき時代となっている。

● オープンサイエンス・オープンイノベーション型研究プラットフォームの構築（知の活用システムの変革）

ほぼすべての科学研究にデータ解析のセンスが必要になっている。技術の進展が非常に早く、研究のあり方が変容していく時代において、単独の研究室、単独の研究チーム、単独の研究所でできることには限界がある。今後もこの方向は進んでいくとみられ、機能単位としての研究モジュールが柔軟に組み合わせることが求められていく時代になっている。

例えば、生物に関わる分野では、生命の複雑性、遺伝子型や表現型の階層の非線形性、環境要因による敏感な変動などに起因してデータ取得に関する課題が顕在化していることから、国際的なビッグプロジェクトや研究所において、プロジェクト等内にオープンソースのクラウドベースのデータプラットフォームを置くことは当たり前のこととなっている。同時に、研究所にコアファシリティ（技術プラットフォーム）を置き、ラボや研究グループにサービスを提供して研究を進めるスタイルが広がっている。

大学や国研の研究所あるいは企業の研究所においてもデジタルトランスフォーメーションに対応した研究環境の構築が必要となっている。基礎研究から社会実装までを展望した一貫通貫の研究開発が求められる。従来型の研究開発プロセス、例えば、競争的研究費によって学術界の研究を支援し、生み出されたシーズが論文等を通じて産業界へ伝わり、企業による社内研究開発あるいは学との共同研究を通じて実用化されるというプロセスは、速度の点からも成功確率の点からも十分な機能を果たせてなくなっている。米国においては、学の研究とベンチャー企業での開発とが極めて近い。欧州では、フラウンホーファー研究機構などの公的研究拠点の役割が大きい。これら研究拠点には、先行投資して開発、蓄積された技術群を基に複数の企業が集まり、効率よく研究開発が進められている。日本では、大型プログラムによって立ち上がったものの、その後の基盤的支援が行われずに縮小せざるを得ない研究拠点も多い。産業に直結した研究開発を行うような研究拠点に対しては、従来の競争的資金を主たる原資とする研究がそぐわない。欧州の研究拠点のように、例えば全体の3割程度の基盤部分を公的資金によって支える、という方式に変更していく議論も必要であろう。

研究の現場毎、個々に独自のデータセットを確立することは非効率的であり、国内でいくつも似たようなデータセットが同時多発的に構築されることは望ましくない。また、AIは発展途上の技術であるため合目的な学習データセット構築方法が確立されておらず、コンソーシアムによる取り組みでも利害関係者が多すぎる場合、かじ取りが困難である。

研究データの流通や活用を推進している団体 FORCE11 は 2014 年にオープンデータについて留意すべき原則として、①発見可能性（Findability）、②利用可能性（Accessibility）、③相互運用可能（Interoperability）、④再利用可能性（Reusability）の4点を挙げた。オープンデータを作成・流通させる際に求められるガイドラインを示している。

科学研究が多くの分野が高度に複雑化・高コスト化していく中で、もはや日本で完結する研究分野も相当限定的である。元来、サイエンスに国境はないが、デジタル化の時代においてはいっそう国際連携を模索していくことが求められる。一方で、地政学的な理由により、必ずしもグローバルにオープンな研究開発関係が構築・継続するか不透明な時代でもあることには留意が必要である。

また、オープンサイエンスは科学技術の分野によって様相が大きく異なることも留意が必要である。その分野の研究開発者が置かれている環境や文化の違い、産業界との関係の違い等によって、オープンサイエンスへの対応が変わってくる。下図はそのような違いを示す一例である。

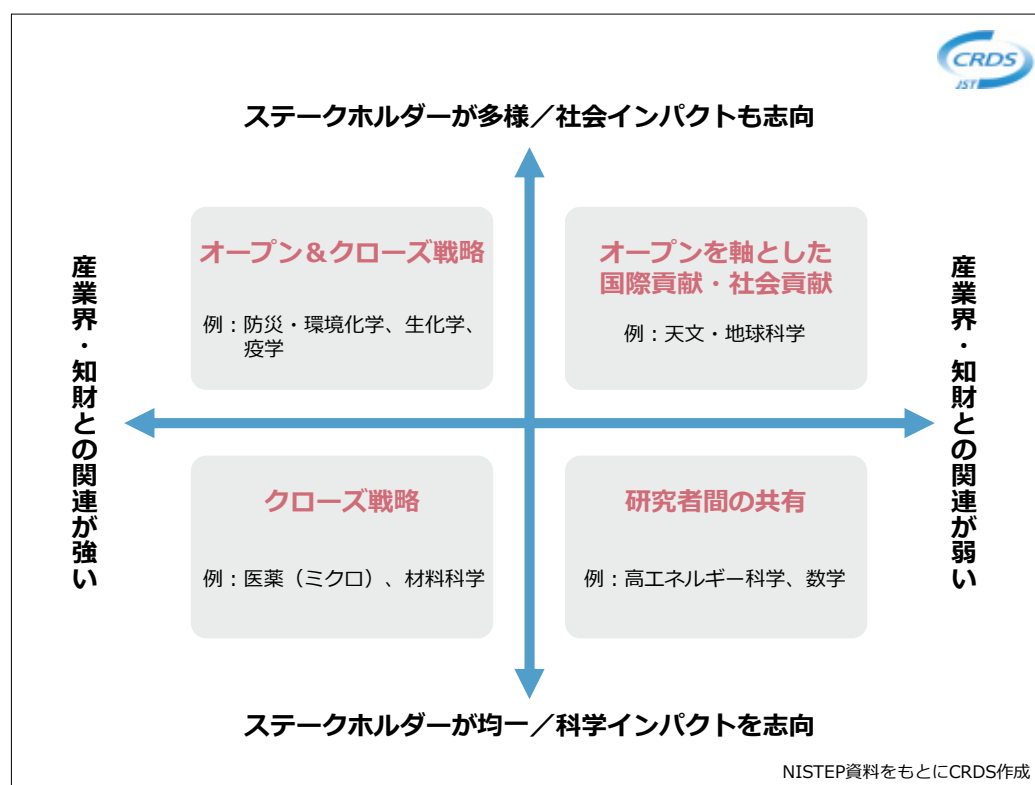


図 6-3 オープンサイエンスに関する戦略のちがい（例）

以上の動向を踏まえた上で、過去 30 年間で、日本の研究室、研究所（企業含む）の研究活動・行動様式はどれだけ変わったのだろうか。いずれの研究分野においても、科学の発展段階にともなって研究の対象はより複雑化し、高度・高価な実験機器と、膨大な実験試行回数、それに伴って生じる膨大な研究データを処理するコンピューティング能力の保有が、これらの研究を進展させ、世界の研究競争において決定的な差をもたらすことが自明となりつつある。さらに、最近ではロボットの進展により、研究に必要な膨大な実験試行回数を自動化する、いわばロボット実験機器の導入が始動している。にもかかわらず、今日の日本の研究における日々の活動・行動様式は、多くの機関であまり変わっていない。

これには二つの原因がある。一つは、昨今指摘されるように、日本の科学技術研究費が頭打ちの状況にあるなか、研究者人口は少しずつ増加してきた。また、1 件当たりの研究費を減らして採択率を上げてきた。上述のように一研究単位あたりの研究に用いなければならないリソース・コストは増えているはずなのに、一人あたりの研究予算は減る構造にある。もう一つは、研究の方法論、やり方そのものにある。すなわち、主要国では高度な実験機器やコンピューター、そして現在では研究データや解析ソフト、そして実験方法の標準化（サンプル標準や測定標準）も含め、ハブ・コア拠点を設けて共有化・共通化する仕組み構築が早くから進んできたが、日本では研究室ごとに、PI 単位ごとに申請された研究費によって導入するスタイルが長く常態化し、研究室や研究分野がタコソボ化した面がある。また、世界の各地でこのように共有化・共通化されたハブは、異分野連携や産学官連携の共

創の場・土壌として大きな存在感を発揮するようになり、いわば現在の IT プラットフォーマーさながらの R&D プラットフォーマーとして、研究の推進力・競争力の源泉となっている。

以上のように研究開発手法に変革が起こると、科学のあり方そのものが問われる時代にもなっていることに留意しなければならない。「仮説を立て、実験を通じて検証する」という実験科学において、これまで仮説生成は人間が行っており、研究のデザインや手法に創造性が発揮されていると言える。

現状、研究者が仮説を立て（研究をデザインし）、実験・検証の部分を AI やロボット等技術が実行し、新しい発見を支援するということまでできている。一方で、仮説の立案を人間が行うのではなく、データから AI が行うというアンバイアスな仮説立案という考え方も出てきている。

将来的にデジタル技術の統合により、仮説の立案や発見、その背後にある法則性を自律的に見いだすということになるのだろうか。このような時代に検証可能性や反証可能性という概念はどう考えれば良いのだろうか。科学の定義を再構築して共有することが必要になるかもしれない。

2. 人に寄り添う環境・社会に向けた研究開発（社会システムの変革）

人間は AI に取って代わられるのではないかと、多くの人が失業するのではないかと、個人情報やプライバシーの侵害、監視社会化するリスクが顕在化する等社会の不安も増しつつある。また、敵対的生成ネットワーク GAN（Generative Adversarial Networks）等、ディープラーニングに基づく生成モデルが開発されたことによって、新しい生成系アプリケーションが次々に生み出されている（たとえば、あたかも本物のような実在しない画像・映像の生成、他者の動作のコピー、線画の自動着色・写実画変換、見えない面まで推測した 3D モデル生成等）。本物と見紛うフェイク動画等の生成が可能になり、SNS を介した世論誘導・政治干渉の危険性が増大している。写真や動画が証拠として信じられるのかという問題も生じてくる。

このような社会との関係から生まれた AI の新たな問題に対処するため、AI 社会原則・倫理ガイドラインが国・世界レベルで議論・策定されるようになった。日本では、「人間中心の AI 社会原則（2019 年 3 月 統合イノベーション戦略推進会議）」が、欧州では「Ethics guidelines for trustworthy AI（2019 年 4 月 欧州委員会）」等がすでに公開されている。

また、企業が遺伝子データや健康データ等の個人情報を用いた研究・開発を行う場合にも、倫理の問題は避けて通れない。

一方で、このような技術やデータの活用により社会・個人が恩恵をこうむることも多い。気候変動問題や社会保障問題等においては、個人の満足と社会システムの最適化を両立することが求められ、このためには人の行動変容を促す方策（情報や金銭的インセンティブなど）も必要となる。社会システムの一部として関与する人の行動に係る研究は今後一層重要になる。すなわち、「人（集団）の行動のモデル化」が必要であり、行動予測のみならず行動変容の方策に関する行動科学の研究も必要である。このような領域は自然科学のみならず、ヒトに関する研究の歴史がある人文・社会科学、例えば行動経済学、社会心理学などとの学際的な分野融合研究が必須となる。

翻っている国際社会を見ると、民主主義、市場原理に対する価値観が一定せずグローバリズムの規範が弱まっているとも捉えることが出来る。地政学および各国政治も、世界経済、各国経済も大きな不確定要因を持っており、80 億人弱の人類の価値観が大きな変動にさらされている。この価値観の変動は基本的には情報発信者が一部の人に限られていたという時代からすべての人が情報発信者に容易になることができる（つまりデジタルトランスフォーメーションの流れ）こととも関係している。様々な情報に簡便にコンタクトできようになったことから、個人の価値観が多様になり、あるべき、ありたい社会の姿が定まらなくなっているとも言える。個人のプライバシーに関してをみても欧

州、米国、中国で国としての制度が異なっているのではなかろうか。日本国民はいったいどういう社会を望むのか。日本はどのような社会を構築しようとしているのか見えにくい時代でもある。このような現代においては、自然科学への問いかけは、社会や個人との価値観の共有と人文・社会科学との協同という方向で問題設定がされ直しつつある、とも考えられる。

何より重要なのは、第5期科学技術基本計画においても謳われているように、多様なステークホルダーが双方向で対話・協働し、それらを政策形成や知識創造へと結び付ける「共創」を推進することである。それら取組の具体化と実行が政策、研究者に浸透することが求められている。

こうした異分野連携や社会との対話を実装するには、法制度や倫理的課題も考慮しながら社会システムを設計・構築し継続的に進化させるフレームワークが必要になる。進化的社会システムのデザインフレームワークは、現在の社会システムを計測しデータを集め、そのデータを分析して要件を明らかにし、要件に基づき社会システムを設計・構築して社会に提供し、新しくなった社会システムをまた計測する、というスパイラル状のデザインプロセスによって社会システムを進化させるものである。

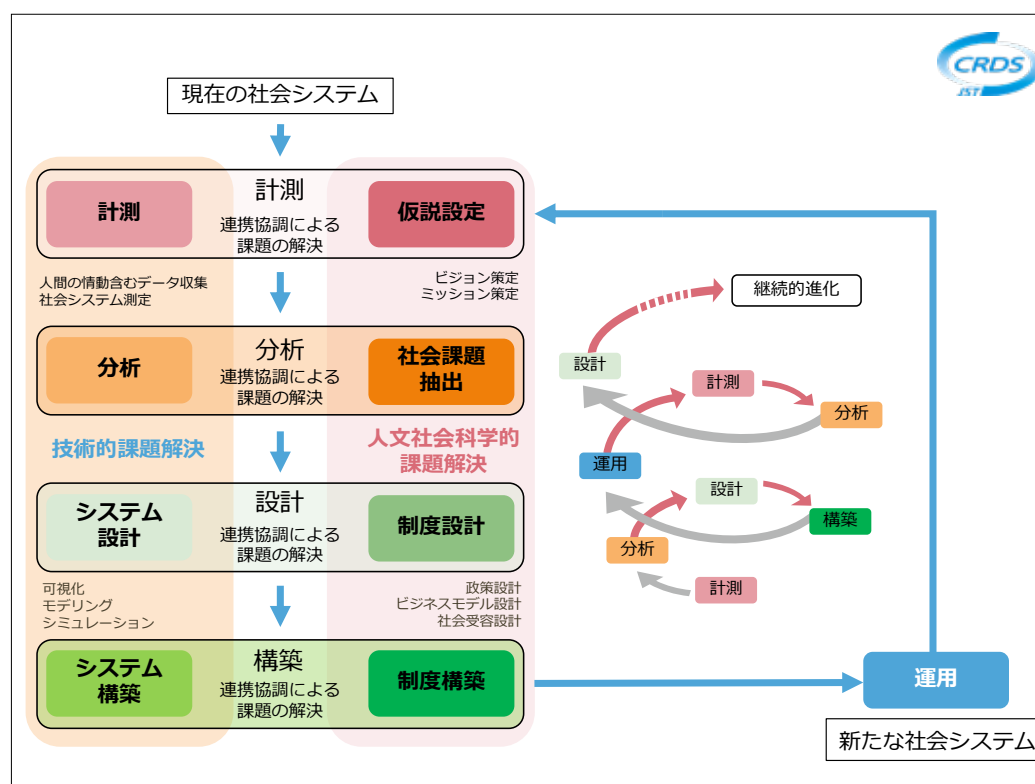


図 6-4 進化的社会システムデザインのフレームワーク

3. 大学や社会における教育体系の見直し（教育システムの変革）

様々な研究分野において、データ科学者、インフォマティシャン、デジタル人材の不足が課題として取り上げられている。これは、技術の進展による社会の変化に日本の教育システムが追いついていなかったことが考えられる。特に、デジタルネイティブ世代より上の世代、すなわち科学者、研究者のボリュームゾーンにそのような人材が不足していることが示唆される。現在の各部局にそのようなポストがほとんどないことも問題である。

人文学・社会科学を含む学際、異分野連携、融合・横断という言葉が何度も登場する。すべての分野の研究者にデジタルの素養は必須となっているが、大学の教育体系がそのようなになっていない。旧

来の大学教育、大学院教育の枠組みのあり方、例えば理学部、工学部、薬学部、医学部といった枠組み自体がすでに時代の進歩に追いついていないのではないかという考え方もある。技術の変化があまりにも早いため、社会制度の改革、人間の能力・スキル獲得のスピードが追いつかないことも問題になっている。社会として人材の再教育システムが重要になってくるだろう。

以上のような課題を踏まえて、世界各国では STEM 教育が重視されている。Science : Technology : Engineering : Mathematics それぞれの頭文字を取って、科学・技術・工学・数学の教育分野を総称した言葉であり、2000 年代に米国で始まった教育モデルである。ハイテク職種の適格者が不足しているという懸念からスタートしたが、学習や能力開発について統合的で学際的なアプローチへと進化している。新しいアプローチには、実社会での応用を通じて学術的な概念を教えることが含まれ、学校、地域社会、職場での公式・非公式の学習が包含される。大学学部における STEM 教育だけでなく、大学への入口である K-12 教育から出口である産業界の意向までを考慮した一貫性のある教育を志向しており、STEM 学習の改善、若者・市民の STEM 教育への参加・関与の拡大、大学学部と大学院教育の改善等が進んでいる。

デジタル・リテラシーを高め、デジタル人材を強化するには、必然的に日本の教育・研究・産業活動全体の関与が必要となる。

参考資料

- ・ 科学技術振興機構 研究開発戦略センター 「研究開発の俯瞰報告書 統合版（2019 年）～俯瞰と潮流～」(CRDS-FY2019-FR-01)（2019 年 7 月）
- ・ 科学技術振興機構 研究開発戦略センター 「Beyond Disciplines — JST/CRDS が注目する 12 の異分野融合領域・横断テーマ（2018 年）—」(CRDS-FY2018-RR-02)（2018 年 8 月）
- ・ 科学技術振興機構 研究開発戦略センター 「— The Beyond Disciplines Collection — 科学技術イノベーション政策における社会との関係深化に向けて 我が国における ELSI/RR の構築と定着」(CRDS-FY2019-RR-04)（2019 年 11 月）
- ・ 科学技術振興機構 研究開発戦略センター 「戦略プロポーザル 自然科学と人文・社会科学との連携を具体化するために—連携方策と先行事例—」(CRDS-FY2018-SP-01)（2018 年 10 月）
- ・ 科学技術振興機構 研究開発戦略センター 「戦略プロポーザル 進化的社会システムデザイン ～自然科学と社会科学の連携協調による持続可能な社会の実現～」(CRDS-FY2019-SP-01)（2019 年 7 月）
- ・ 科学技術振興機構 研究開発戦略センター 「調査報告書 計測横断チーム調査報告書 計測の俯瞰と新潮流」(CRDS-FY2018-RR-03)（2018 年 12 月）

■ 報告書作成メンバー ■

(監修・編著) 島津 博基、永野 智己 (企画運営室)

- 1 島津 博基 (企画運営室)
 - 2.1 井上 貴文、中村 輝郎、島津 博基 (ライフサイエンス・臨床医学ユニット)
 - 2.2 中村 輝郎 (ライフサイエンス・臨床医学ユニット)
 - 2.3 山原 恵子 (ライフサイエンス・臨床医学ユニット)
 - 2.4 宮下 哲、永野 智己 (ナノテクノロジー・材料ユニット)、
嶋田 義皓 (システム・情報科学技術ユニット)
 - 2.5 大平 竜也 (環境・エネルギーユニット)
 - 3.1 辻 真博 (ライフサイエンス・臨床医学ユニット)
 - 3.2 宮菌 侑也 (ライフサイエンス・臨床医学ユニット)
 - 3.3 桑原 明日香 (ライフサイエンス・臨床医学ユニット)、荒岡 礼 (ナノテクノロジー・材料ユニット)
 - 3.4 中村 亮二 (環境・エネルギーユニット)
 - 3.5 長谷川 景子 (環境・エネルギーユニット)
 - 3.6 尾山 宏次 (環境・エネルギーユニット)
 - 3.7 高島 洋典 (システム・情報科学技術ユニット)、尾山 宏次 (環境・エネルギーユニット)
 - 3.8 原田 裕明 (科学技術イノベーション政策ユニット)
 - 4.1 福島 俊一 (システム・情報科学技術ユニット)
 - 4.2 福島 俊一 (システム・情報科学技術ユニット)
 - 4.3 八巻 徹也 (ナノテクノロジー・材料ユニット)、宮菌 侑也 (ライフサイエンス・臨床医学ユニット)
 - 4.4 馬場 寿夫 (ナノテクノロジー・材料ユニット)、茂木 強 (システム・情報科学技術ユニット)、
島津 博基 (ライフサイエンス・臨床医学ユニット)
 - 5 新田 英之、澤田 朋子、長谷川 貴之、原田 裕明、山村 将博 (海外動向ユニット)
 - 6 島津 博基 (企画運営室)、青木 孝、高島 洋典、福島 俊一 (システム・情報科学技術ユニット)
- (企画・構成) 魚住 まどか、小松崎 美奈、梅原 千慶、根上 純子、
鈴木 ソフィア 沙織、中山 智弘 (企画運営室)

※お問い合わせは下記までお願いいたします。

CRDS-FY2020-RR-01

The **Beyond Disciplines** Collection

デジタルトランスフォーメーションに伴う 科学技術・イノベーションの変容

令和2年4月 April, 2020

ISBN 978-4-88890-665-4

国立研究開発法人科学技術振興機構 研究開発戦略センター

Center for Research and Development Strategy, Japan Science and Technology Agency

〒102-0076 東京都千代田区五番町 7

電 話 03-5214-7481

E mail crds@jst.go.jp

<https://www.jst.go.jp/crds/>

©2020 JST/CRDS

許可無く複製／複製することを禁じます。

引用を行う際は、必ず出典を記述願います。

No part of this publication may be reproduced, copied, transmitted or translated without written permission.

Application should be sent to crds@jst.go.jp. Any quotations must be appropriately acknowledged.

ISBN 978-4-88890-665-4

