

『バイオDX』による科学的発見の追究

AI・ビッグデータの活用を中心とした生命科学研究のデジタルトランスフォーメーション（『バイオDX』）により、未踏の生命現象解明に挑む！！

近年の生命科学研究における DX

ビッグデータ・AI

AIを用いた画像解析や新薬探索等の技術が進展。2020年11月には、AIによるタンパク質構造予測性の飛躍的向上が話題に。



(Callaway, Nature, 2020)

自動化

化合物合成や代謝工学などにおいて、実験の自動化（ロボット利用）、ハイスループット化の研究発表が目立つように。



IBM RXN for Chemistry (<https://rxn.res.ibm.com/rxn/robo-rxn/welcome>)

従来の生命科学研究

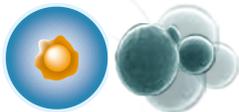
人間の認知能力、作業量による制約。偶然の発見に期待。



©2016 DBCLS TogoTV

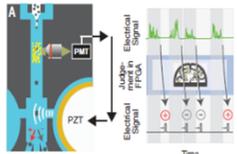
我が国の強み

再生医療研究やバイオ生産技術等



©2016 DBCLS TogoTV

世界を先導する計測・機械化技術



(Ota et al., Science, 2018)

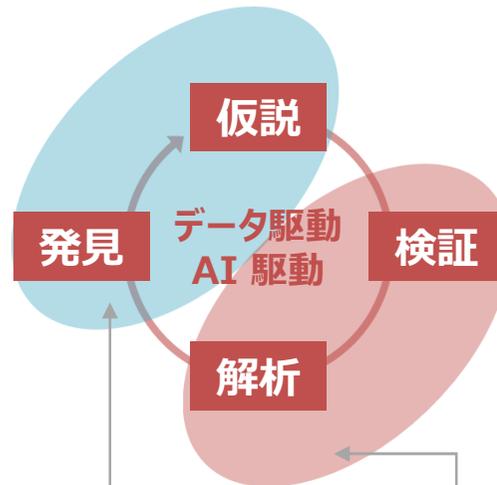


(Yachie et al., Nat. Biotechnol., 2017)

人間の認知能力では辿り着かなかったかもしれない科学的発見

質の高いビッグデータの取得・解析とAI駆動による結果
解釈・推論の両輪のアプローチにより、**人間の認知能力・バイアスを越えた科学的アプローチ**を可能に！

データ・AI駆動 × 生命科学研究



生命科学研究のDXにより創る未来像



© 2016 DBCLS TogoTV

『複雑な生命システムを解明』

『予防、検査・診断、創薬および治療技術の効率的な創成』

『研究者が真に創造的な活動に取り組むことのできる環境』

AI技術による解析結果解釈・モデリング・仮説生成

自然言語処理技術等に基づく、**既存の領域知識の抽出・活用**に加え、能動推論により実験結果を解釈し、**人力では為し得ない網羅的な理論シミュレーション及び解釈妥当性評価**を効率的に推進！

データの『量』と『質』の確保

マルチモーダル解析や分析装置の自動化等、**均質かつ大量のデータを取得するための技術**、あるいは**実験計画を自動で柔軟に判断し遂行するAI技術**の開発等により、データ駆動型研究を加速！

令和3年度戦略目標

1. 目標名

『バイオ DX』による科学的発見の追究

2. 概要

AIによる画像解析、新薬探索、タンパク質三次元構造予測等、AI・ビッグデータの活用を中心としたデジタルトランスフォーメーション(DX)は、生命科学研究の在り方に大きな変革をもたらそうとしている。ヒト等の生命システムは、絶えず変動するオープンな複雑系システムであることから、生命科学研究において、従来の生化学的・分子生物学的研究手法と並び、データ駆動型、AI駆動型の研究手法が必須となりつつある。

我が国では、AIを用いた自動計測装置の開発や、生命科学研究の実験ロボットの開発など、世界を先導する計測・機械化技術を有している。また、自然言語処理技術による既存の領域知識獲得や強化学習を通じた自動仮説生成、能動推論についても研究が進められており、良質・大量データ取得やビッグデータ解析だけでなく、新たな知見の獲得やモデリングの自動化が期待される場所である。これら技術の活用によるデータ駆動型、AI駆動型研究の推進により、生命科学研究の新しいスタンダードを確立できるものと考えられる。

膨大な背景知識やビッグデータの適切な利活用は人間の認知能力や認知バイアスを越えた未知未踏な現象の発見をもたらす大きな可能性を秘めている。本戦略目標では、データ駆動、AI駆動型の生命科学研究の推進により、従来の方法論や人間の認知能力では辿り着かなかった科学的発見を目的とする。加えて、本研究領域の推進により、研究の自動化から研究プロセス全体の自動化・自律化へとつながる礎を築き、研究者が真に創造的な活動に取り組む事のできるニューノーマルな研究環境の創成へとつなげる。

3. 達成目標

本戦略目標では、データ取得の在り方を含めた新しいデータ駆動型・AI駆動型の生命科学研究を推進し、従来の方法論、人間の認知能力では辿り着かなかった科学的発見を目指す。

具体的には、以下の3つの達成を目指す。(※)

- (1) 多様、大量のデータセットから新しい生命現象やその法則の発見
- (2) データ駆動型研究に資する革新的なデータ取得・解析技術の開発
- (3) 仮説生成、推論の自動化等、生命科学現象に関する実験自動化を越えたAI駆動型研究の手法の確立

※ただし、(2)、(3)は(1)の達成を目的とするものであり、(2)、(3)の技術や手法開発そのもののみを目的とする課題は認めない。

4. 研究推進の際に見据えるべき将来の社会像

3. 「達成目標」の実現を通じ、複雑な生命システムを解明するとともに、生命科学研究のパラダイムシフトを起こし、以下に挙げるような社会の実現に貢献する。

- ・ 複雑な生命システムの解明が可能な社会
- ・ 予防、検査・診断、創薬及び治療技術の効率的な創成が可能な社会
- ・ 研究者が真に創造的な活動に取り組む事のできる社会

5. 具体的な研究例

(1) 多様、大量のデータセットから新しい生命現象やその法則の発見

自動化機器や高度な分析機器等の活用により、均質、リアルタイム、マルチモーダルなデータを取得し、AI を活用することで、人間によるデータ取得・解析の質的・量的限界を超えて、新たな生命現象の発見とそのモデル化を可能とするための研究を行う。具体的には以下の研究等を想定。

- ・ 大規模反復試験により、偶然に起こる生命現象をモデル化する研究
- ・ 細胞や組織の活動を、リアルタイム・自動で計測し、さらには摂動を加える等により、疾病の数理的理解等、生命現象の解明につなげる研究
- ・ DBTL [設計 (Design) - 構築 (Build) - 試験 (Test) - 学習 (Learn)] サイクルの自動化・高度化により、細胞内の代謝システムを数理的に理解し、物質生産への応用に資する研究
- ・ DMTA [デザイン (Design) - 合成 (Make) - 試験 (Test) - 解析 (Analysis)] サイクルの自動化・高度化により、創薬の新規標的探索等に資する研究

(2) データ駆動型研究に資する革新的なデータ取得・解析技術の開発 (※)

マルチモーダル解析や分析装置の自動化等、均質かつ大量のデータを取得するための革新的技術や実験計画を自動で柔軟に判断し遂行する AI 技術の開発等により、データ駆動型研究の加速に資する基盤技術研究を行う。具体的には以下の研究等を想定。

- ・ イメージング情報とオミクス情報の融合解析等により、細胞や組織、個体から非侵襲でリアルタイムのデータ取得を可能とする技術の開発
- ・ フロー合成法、インライン分析、ロボット等と AI の組み合わせにより、従来法では実現不可能な化合物合成やアッセイ系構築、予測を可能とする技術の開発
- ・ 得られた実験データや文献データ等から、次に施行すべき最適な実験条件を導き出す AI 技術の開発

(3) 仮説生成、推論の自動化等、生命科学現象に関する実験自動化を越えた AI 駆動型研究の手法の確立 (※)

生命科学における研究プロセスにおいて、人間の認識能力・バイアスを本質的に越えた発見に向けては、実験データに基づく解釈・モデリングの網羅性及びそれらの妥当性の評価が重要となる。これらのアプローチには様々な AI 技術の開発が不可欠であり、生命科学分野にお

けるこれまでの膨大な知の蓄積を踏まえた新たな科学的発見の在り方を確立する。具体的には、以下の研究・開発を想定。

- ・自然言語処理技術に基づいた、先行研究からの領域知識の抽出と、それらを利用した研究方針策定を補助するシステムの開発
- ・実験結果の自動解析結果に基づいた、自動仮説生成・能動推論の手法の研究
- ・これらのシステムや手法を研究開発する中で得られる、科学的発見につながるモデルを網羅的に自動評価することが可能となるシステムの開発

※（２）、（３）は（１）の達成を目的とするものであり、（２）、（３）の技術や手法開発そのもののみを目的とする課題は認めない。

6. 国内外の研究動向

（国内動向）

科学技術振興機構（JST）が未来社会創造事業「共通基盤」領域において平成 30 年度から実施する「ロボティックバイオロジーによる生命科学の加速」では、ロボットを活用した研究自動化について、一部取り組まれている。平成 28 年度戦略目標「急速に高度化・複雑化が進む人工知能基盤技術を用いて多種膨大な情報の利活用を可能とする統合化技術」や、平成 25 年度戦略目標「分野を超えたビッグデータ利活用により新たな知識や洞察を得るための革新的な情報技術及びそれらを支える数理的手法の創出・高度化・体系化」等において、生命科学分野における AI 研究やビッグデータの利用研究が一部取り組まれている。また、日本医療研究開発機構（AMED）が平成 29 年度から実施する創薬等ライフサイエンス研究支援基盤事業において、共通基盤技術の自動化や AI 創薬に関する研究が進められている。

令和元年には「ラボラトリーオートメーション研究会」が発足する等、生命科学分野を含めた研究手法における DX が、新たな研究領域として確立しつつある。

（国外動向）

平成 24 年に畳み込みニューラルネットワーク（CNN）を多層ニューラルネットワークとして用いた画像解析技術が登場し、平成 28 年頃から網膜や皮膚がんの画像診断にディープラーニングを用いた論文が発表され、注目を集めるようになった。こうした技術の進展の結果、イメージング画像のハイスループットな取得や超解像化、三次元化が可能になった他、平成 30 年には、FDA（米国食品医薬品局）が世界で初めて自律型 AI 診断システムを承認した。

平成 26 年に敵対的生成ネットワーク（GAN）が登場し、本物に見えるが実在しないデータ（フェイク画像等）の自動生成といったことができるようになった。様々な進展を経て、同じ「生成モデル」として平成 29 年には、変分オートエンコーダ（VAE）に基づき化合物を生成する技術が登場した。こうした技術が創薬に展開されている。令和 2 年には、大手製薬企業による AI を活用して創製された新薬候補化合物の臨床試験の開始が公表された。

平成 30 年頃から、AI との相性の良さから、化合物合成や代謝工学などにおいて、実験の自動化（ロボット利用）、ハイスループット化の研究発表が目立つようになってきている。

令和 2 年 11 月には、英国に拠点を置くディープマインド社が、タンパク質構造予測「アルファ・フォールド・システム」を発表し、AI 活用による生命科学研究の可能性の拡大が示された。

7. 検討の経緯

「戦略目標の策定の指針」（令和元年 7 月科学技術・学術審議会基礎研究振興部会決定）に基づき、以下のとおり検討を行った。

(1) 科学研究費助成事業データベース等を用いた国内の研究動向に関する分析及び研究論文データベースの分析資料を基に、科学技術・学術政策研究所科学技術予測センターの専門家ネットワークに参画している専門家や JST 研究開発戦略センター (CRDS) の各分野ユニット、AMED のプログラムディレクター等を対象として、注目すべき研究動向に関するアンケートを実施した。

(2) 上記アンケートの結果及び JST-CRDS の俯瞰報告書や「AI×バイオ DX 時代のライフサイエンス・バイオメディカル研究」報告書等を参考にして分析を進めた結果、生命科学研究におけるデジタルトランスフォーメーションが重要であるとの認識を得て、注目すべき研究動向「生命科学研究の DX」を特定した。

(3) 令和 2 年 11 月に、文部科学省と JST は共催で、当該研究動向に関係する産学の有識者が一堂に会するワークショップを開催し、注目すべき国内外の動向、研究や技術開発の方向性、研究期間中に達成すべき目標等について議論した。ワークショップにおける議論や JST-CRDS のライフサイエンス・臨床医学ユニット等の提案を踏まえ、本戦略目標を作成した。

8. 閣議決定文書等における関係記載

「AI 戦略 2019」（令和元年 6 月 11 日統合イノベーション戦略推進会議）

別表 1、4-1

- ・ AI による科学的発見の研究（具体的取組内容：細胞レベルでの実験検証を対象として、仮説生成、実験計画立案、実験の自動実施、結果の検証などを行う AI サイエントリストの開発）

「統合イノベーション戦略 2020」（令和 2 年 7 月 17 日閣議決定）

第 II 部第 2 章 3 ①

- ・ 全国規模で研究開発をシームレスに連動させ、その活動を継続できる環境の実現に向け、

AI、ロボット技術を活用した実験の自動化などスマートラボの取組や、遠隔地からネットワークを介して研究インフラにアクセスし分析等を実施する取組の推進、高速通信ネットワークの整備、大規模な計算資源の徹底活用、研究データ等の効果的・効率的な創出・共用・利活用環境の整備等、研究開発環境と研究手法のデジタル転換を推進する。更に、オープンサイエンスの推進を図る。

第Ⅲ部第2章（1）②

- ・最適な研究設備・機器へのアクセスの確保に向けた、各施設・設備のネットワーク化、共用プラットフォーム構築を進めるとともに、研究室における AI・ロボット等の活用によるスマートラボラトリ化を促進する。

第Ⅲ部第5章（2）

- ・バイオとデジタルの融合を全ての土台とし、生物活動のデータ化等も含めてデータ基盤を構築しそれを最大限活用することにより産業・研究を発展させることで、世界最先端のバイオエコノミー社会を実現

「バイオ戦略 2020」（令和 2 年 6 月 26 日統合イノベーション戦略推進会議）

3 3. 1 ②バイオとデジタルの融合

- ・市場領域・科学の発展に必要なビッグデータ収集・バイオデータ基盤構築の方向性と持続可能な方策を提示。測定機器やバイオデータ基盤を活用し、我が国の強みや匠の技を AI 化。これを支える研究開発人材等を、質を重視して育成

9. その他

本戦略目標に関連する施策として、未来社会創造事業「共通基盤」領域「ロボティックバイオロジーによる生命科学の加速」において、研究の自動化に向けた取組が行われることから、本戦略目標との技術連携等により効率的・効果的な研究推進が期待される。

本戦略目標の下で行われる研究によって得られたデータや開発された技術等については、領域内外における適切な活用や共有を推進するとともに、科学研究費助成事業の新学術領域研究・学術研究支援基盤形成「先端バイオイメージング支援プラットフォーム」（平成 28～令和 3 年度）や金沢大学のナノ生命科学研究所の「世界トップレベル研究拠点プログラム（WPI）」事業及び AMED「創薬等先端技術支援基盤プラットフォーム（BINDS）」事業等の世界を牽引する計測・解析技術の研究開発基盤との連携が期待される。

また、本戦略目標では、研究体制の中にデータサイエンス人材を適切に組み込むことを応募の要件とすること等により、生命科学分野におけるデータサイエンス人材の育成も図ることとする。