

バイオデータ連携・利活用に向けたガイドブック

令和5年6月

目次

1. はじめに	2
1.1. データ連携・利活用を取り巻く環境	5
1.2. バイオデータの連携・利活用の現状と課題	7
1.3. 本ガイドブックの位置づけ	9
2. データ連携・利活用に関する基本概念	12
2.1. データ連携・利活用とは何か	13
2.1.1. データ連携・利活用で取り扱うデータ	13
2.1.2. データ連携・利活用に関わるプレーヤ	15
2.1.3. データ連携・利活用が生み出される場	16
2.2. データ連携・利活用を成功に結び付けるメソッド	24
2.2.1. データ連携・利活用によって生み出す価値を明確化	24
2.2.2. 生み出す価値を最大化するためのバリューチェーンを設計	27
2.2.3. 生み出す価値の最大化の障害となる課題を抽出	28
2.2.4. 課題を解決するためのデータ連携・利活用の実証を推進	31
2.3. データ連携・利活用を効果的に進めるためのヒント	32
2.3.1. フォアキャスト・バックキャストの双方向のアプローチ	32
2.3.2. データ連携の加速に繋がるコミュニケーション	33
3. データ連携・利活用に向けたデータの管理・運用方法	39
3.1. データの整理と可視化	41
3.2. データシェアリングポリシーの策定	45
3.3. データマネジメントプランの策定	47
3.4. データ連携・利活用における契約等手続き方法の策定	50
3.5. データの成果化	52
3.5.1. 標準化、データクレンジングとデータ形式の設計	52
3.5.2. 事業・分野間連携	54
3.5.3. 事業終了後のデータの取扱い	55
3.6. プレーヤ間の利害関係調整機能の整備	59
3.7. セキュリティやトラストの確保	60
3.8. 研究対象者とのコミュニケーション	67
<コラム：データ連携・利活用におけるマネジメント体制の構築>	69
(別添) バイオ戦略におけるデータ関連取組一覧	72

1. はじめに

経済社会のあらゆる面でデータを起点とした価値創造が注目され、その基盤となるデジタル化とデータ連携・利活用の重要性が飛躍的に高まっている。バイオ分野は生命現象という共通の科学的知識に立脚しながらも多種多様な産業上の出口を有しており、質の高いデジタルデータの創出とその連携・利活用が新たな科学的知識の創出や産業競争力の強化につながるものが強く期待される。

このような期待にも関わらず、我が国におけるバイオ分野のデータ連携・利活用の取組は低調であり、またその連携の範囲は小さいままに留まっている。この原因の一つには、「バイオ」として表現される分野であっても実際には多種多様な専門領域の集合であり、専門領域ごとの標準的なデータ連携の取組の方法論が異なること、いわば「常識」が異なることによる実務の障害となっているとの指摘がある。

この「バイオデータ連携・利活用に向けたガイドブック」（以下「本ガイドブック」という。）は、上記の指摘を踏まえ、多種多様な専門領域を有するバイオ分野のデータ連携・利活用を進めるに際し、関係者間の共通認識の醸成に資するよう、押さえておくことが望ましい基本的な考え方をまとめたものである。

本ガイドブックは、データ連携・利活用に取り組む上での基礎的な背景や概念、実務の基本的な考え方、実務を進める上での基盤となるデータの管理・運用にかかる基本的な考え方、をそれぞれ整理している。バイオデータ連携・利活用の実務に関わる方々に幅広く参照いただけると幸いである。

なお、本ガイドブックは、「バイオデータ連携・利活用に関するガイドライン 中間とりまとめ」（以下「旧版」という。）において示した内容を踏まえ、内閣府戦略的イノベーション創造プログラム「スマートバイオ・産業・農業基盤技術課題」でのデータ連携・利活用の実践を通じて得られた知見を旧版に反映させたものであり、旧版の増補改訂版に相当する。

本ガイドブックの位置づけ	
目的	<ul style="list-style-type: none"> ・ バイオデータの連携・利活用促進に向けた関係者間での共通認識の醸成に資することを目的として、考え方や進め方、特に関係者間で調整が生じやすい論点とその対応策等を示す。 ・ 本ガイドブックに則った事前調整・対応をデータ提供者・利用者・仲介者等の関係者が図ることにより、バイオデータの連携・利活用がもたらす成果の最大化を図る。
想定する読者	<ul style="list-style-type: none"> ・ 国研・大学や民間企業の研究者などの実務者 →関係者との調整やバイオデータの連携・利活用の具体的な進め方の参考として活用いただきたい。 ・ 企業等でデータ利活用の推進戦略等をリードする立場の方（R&D 部門等の責任者等） →決裁権者としてバイオデータ連携・利活用の意義や重要性、進め方について理解を深め、組織単位でのデータ連携に向けた投資意欲を高めるとともに、組織内のリテラシー向上や情報発信等に活用いただきたい。

図表 1 本ガイドブックの位置づけ

知りたいこと		対象となる章
データ連携・利活用に取組む上で、 その意義や背景について理解を深めたい	➡	P2~7「1.はじめに」をご覧ください -1.1 データ連携・利活用を取り巻く環境 -1.2 バイオデータの連携・利活用の現状と課題 -1.3 本ガイドブックの位置づけ
データ連携・利活用に取組む上で、 基本的な概念から理解を深めたい	➡	P13~23「2.1.データ連携・利活用とは何か」をご覧ください -2.1.1 データ連携・利活用の対象となるデータ -2.1.2 データ連携・利活用に関わるプレーヤ -2.1.3 データ連携・利活用が生まれる場
データ連携・利活用を成果創出につなげる ための方法論について理解を深めたい	➡	P24~31「2.2.データ連携・利活用を成功に結び付けるメソッド」 をご覧ください -2.2.1 データ連携・利活用によって生み出す価値を明確化 -2.2.2 生み出す価値を最大化するためのバリューチェーンを設計 -2.2.3 生み出す価値の最大化の障害となる課題を抽出 -2.2.4 課題を解決するためのデータ連携・利活用の実証を推進
データ連携・利活用に向けて、効果的に 取組を進めていくためのヒントとなる考え 方を知りたい	➡	P32~38「2.3.データ連携・利活用を効果的に進めるためのヒン ト」をご覧ください -2.3.1 フォアキャスト・バックキャストの双方向のアプローチ -2.3.2 データ連携の加速に繋がるコミュニケーション
自機関が保有するデータの棚卸し・整理 から取組を進めていきたい	➡	P41~44 以下の章をご覧ください -3.1 データの整理と可視化
データの棚卸し・整理を踏まえて、データ 連携・利活用の方針を定めたい	➡	P45~51 以下の章をご覧ください -3.2 データシェアリングポリシーの策定 -3.3 データマネジメントプランの策定 -3.4 データ連携・利活用における契約等手続き方法の策定
継続的にデータの連携・利活用の成果を 創出していきたい	➡	P52~68 以下の章をご覧ください -3.5 データの成果化 -3.6 プレーヤ間の利害関係調整機能の整備 -3.7 セキュリティやトラストの確保 -3.8 研究対象者とのコミュニケーション

図表 2 知りたいことと対象となる章

1.1. データ連携・利活用を取り巻く環境

我が国は、「第5期科学技術基本計画」¹において、「サイバー空間とフィジカル空間を高度に融合させたシステムにより、経済発展と社会的課題の解決を両立する人間中心の社会」である「Society 5.0」を、我が国が目指すべき社会として提唱した。

Society 5.0では、知識や情報が共有されず分野横断的な連携が不十分であるという情報社会(Society 4.0)における課題が解決され、IoT²で全ての人とモノがつながり様々な知識や情報が共有され、新たな価値を生み出すことを目指している。このため、データの利活用がロボットや人工知能(AI)と並び重要な基盤技術に位置付けられている。

一方で、「第6期科学技術・イノベーション基本計画」³においては、Society 5.0の実現に向け、

- あらゆる分野でIT化を進めていたものの、既存の業務の効率性の向上を目指す取組が中心となり、諸外国のようなデータ連携・活用による新たなビジネスモデルの創出が不十分である
- 各組織が異なるシステムでネットワークを閉鎖的に利用している現在の状況では、分野を跨いだリアルタイムでのデータ収集・分析・活用を行う環境が整っていない

といった課題が指摘されている。

こうした状況を踏まえ、Society 5.0のビジョンを実現するためのデータ利活用に焦点を当てた、「包括的データ戦略」⁴が策定されている。この中で、データは智慧・価値・競争力の源泉であるとともに、課題先進国である日本の社会課題を解決する切り札と位置付けられ、現実世界をサイバー空間で再現(デジタルツイン)することと、個人、民間企業、国家のニーズを踏まえ、新たな価値を創出することの両面が求められている。

また、内閣府の「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」⁵では、研究データ基盤システムや関係する制度環境の整備、先進的なデ

1 <https://www8.cao.go.jp/cstp/kihonkeikaku/5honbun.pdf>

2 Internet of Things

3 <https://www8.cao.go.jp/cstp/kihonkeikaku/6honbun.pdf>

4 https://www.soumu.go.jp/main_content/000756398.pdf

5 <https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/index.html>

ータマネジメントの推進、研究データ基盤の管理・利活用に関する考え方を示している。

欧米に目を向けると、米国の医学研究分野における研究資金配分機関である国立衛生研究所(National Institute of Health、以下NIH)がデータの運用・管理方針を定めた「Data Management and Sharing Policy」⁶を公開した。EUでは、オープンサイエンスに向けたデータ公開の指針を表した「FAIR原則」⁷、健康研究におけるデータ管理とデータガバナンスを改善するために開始された「FAIR plus プロジェクト」⁸など、実証プロジェクトを通じてその結果を踏まえたデータ連携・利活用に向けた制度・仕組みが整備されつつある。こうした各国で行われている取組、生じている動向を適切に把握しながら、国内の取組を進めていくことが重要である。

6 <https://sharing.nih.gov/data-management-and-sharing-policy>

7 <https://biosciencedbc.jp/about-us/report/fair-principle/>

8 <https://www.fairplus-project.eu/>

1.2. バイオデータの連携・利活用の現状と課題

バイオ分野においても、イノベーション志向のバイオデータ連携・利活用を本格化し、バイオエコノミーの実現に向けたバイオ関連市場の拡大等を加速させていくことが不可欠である。すでに以下のような動向が見られている。

- マテリアルズインフォマティクスや医療分野でデータ連携・利活用が新たな価値を生み始めている。また、新たなモダリティにおける新薬の開発やその後のドラッグリポジショニングや新たな副作用の課題などを評価し続けるため、臨床での継続的な評価のインフラが必須となってきた。
- バイオ分野と幅広い専門領域との関連性の強化を通じ、新たな市場獲得やこれまでにないビジネスモデルの創出が期待されている。

一方で、バイオデータの連携・利活用を阻害する要因としては、「バイオ戦略2020(基盤的施策)」⁹で示されているものを含め、以下のものが考えられる。

- 過去のバイオ戦略に基づいた様々な研究が行われ、データベースが連携なく散在している。また、データベースの継続的な再評価・アップデートが十分になされない。
- データベースの統合を試みるも、目的を明確にした統合的な戦略がなく、イノベーションに必要なビッグデータはほとんど構築されず、産業界やアカデミアの情報も集まらず、主要な国際連携に参画できていない。また、諸外国との比較検討をした上での我が国の弱みと強みの全体を評価できていない。
- 省庁ごとの予算で設計されたデータベースが科学的な評価の観点からは他省庁所管分野でも活用できる場合もありうるがその調整が困難である。

これらの背景には、バイオ分野が、健康・医療、ヘルスケア、農林水産業、食品、工業製品、環境など多くの専門領域に関連していることが挙げられる。一般的には、専門領域ごとに、あるべきデータの取得方法、精度等に対する考え方の違いが存在することから、専門領域が特に多いバイオ分野では、専門領域ごとの考え方の違いが及ぼす影響は他の分野より大きい¹⁰。例えば、フィジカル空間での実験により得られるデータに関して、分子生物学のように再取得が容易な専門領域もあれば、コホートやフィールドワークのようにデータ再取得が困難な専門領域もある。また、実験条件が標準化しづらい植物を扱う専門領域では、再実験

9 https://www8.cao.go.jp/cstp/bio/bio2020_honbun.pdf

10 <https://biosciencedbc.jp/about-us/report/strategy-for-database-integration/01.html>

とデータ再利用のいずれを選ぶかの判断が大きく異なる。実験方法や条件の多様性とそれによる再現性の確保は、生物学と化学・工学との間で大きく異なる。このような事情が、生物学に端を発する専門領域と、化学・工学との間のデータ連携を難しくしている。ほかにも、健康・医療領域・創薬領域と食品領域との間には、医療行為・医薬品と食品に期待される効果量の違いや、ヒト試験における評価項目の違い¹¹が存在する。それらが積み重なり、健康・医療領域とのデータ連携のハードルは、創薬領域よりも食品領域にとって高いものとなる。これまでも基礎科学のベースとなる分子情報、例えば、DNA や RNA の塩基配列¹²、タンパク質の立体構造¹³、発現情報¹⁴などについては、論文発表時に公開と共有が進んでおり、世界で共通のリポジトリに登録する動きも確立されてきた。一方で、コホート研究などで得られるデータは個人情報を扱う際の問題があり、インフォームド・コンセントのとり方の違いなどの理由で共有が進んでいない部分もある。さらに、今後の気候変動を踏まえた環境の変化との関係性の分析とその結果に基づく戦略構築や食料安全保障の問題も今後の課題となるが、この観点での政策の方向性を議論するうえでこれまで想定していない課題が発生し得る。

上記の背景からこれまでバイオ分野では、特定の専門領域の研究データを除き、データ連携・利活用の取組は十分に進んでこなかった。また、データ分析による新たな課題への取組も生まれなかった。データ連携・利活用の起点と方向性が定まらず、取組が進まないため、ノウハウが蓄積せず、連携・利活用が更に進まなくなるという悪循環が発生している。

しかしながら、バイオ分野は、生命現象を起点としているという共通点に基づく分野間の類似性から、共通の研究開発技術が応用できるケースが多数見られることも踏まえて、データ連携・利活用によって所管省庁の分担を超えた、科学的かつ今後の新たな産業育成も踏まえた新たな付加価値が創出される大きな可能性を秘めていると言える。例えば、培養技術は古くからバイオ分野の根幹をなす技術の一つであるが、育種や物質生産での活用に加え近年では食分野において細胞性食品（いわゆる培養肉）、医療分野においては人工臓器に適用され、また、近年実用化が進むゲノム編集技術は、農作物の育種、医療分野ではゲノム医療に適用される等、分野間連携が進んでいる事例も多数見受けられる。

11 健康・医療領域や創薬領域では客観的な疾病症状が評価項目となるのに対し、食品領域では「対象者自身がどの程度不快に感じているか」といった主観的な指標も評価項目となりうる。

12 <https://www.insdc.org/>

13 <http://www.wwpdb.org/>

14 <https://www.proteomexchange.org/>

また、本来こういったデータベースの基盤は、国としての資産であり、国益につながるものである。データの管理方法・利活用の方法においても安全保障及びサイバーセキュリティの観点からの対応が必要である。さらに単なる研究データだけではなく、産業化することを見据えたサプライチェーンを考えたデータの取扱いについて議論する必要もある。

上述した状況を踏まえて、バイオ分野においては、データ連携・利活用が新たな価値を生み出すことを理解し、コストを超えるモチベーションにつなげ、将来的なデータ利活用のポテンシャルも視野に入れたデータ連携に踏み出すことが期待される。

1.3. 本ガイドブックの位置づけ

バイオ分野のデータ連携・利活用を加速させるためには、データ保有者・利用者をはじめとしたステークホルダが相互の立場・思惑を理解しながら、共通認識を醸成することが重要である。様々な専門領域に跨るバイオデータを取り扱い、立場や属性が異なる関係者と調整を図ることは容易では無い。様々な側面から共通認識を醸成することが必要となる。

本ガイドブックでは、関係者間での共通認識の醸成に資することを目的とし

て、バイオデータの連携・利活用の考え方や進め方、特に関係者間で調整が生じやすい論点とその対応策等を示している。本ガイドブックに則った事前調整・対応をデータ提供者・利用者・仲介者等の関係者が実践することにより、バイオ分野のデータ連携・利活用がもたらす成果の最大化を図ることが期待される。特に、下記の5点を通じて、好循環を構築することを目指す。

- ① 新たな経済的価値を生み出すためのデータ連携・利活用を実証により促進する。
- ② 自らの専門領域と異なる領域との調整に加え、これまで軽視されがちであったデータ連携・利活用に必要なリソースの確保やインセンティブの設定に関するノウハウを蓄積し、調整コストの低下(ローコスト化)を図る。
- ③ データ連携・利活用が生み出す経済的価値を明示することで、関係者のモチベーションを向上させ、連携・利活用を更に拡大する。
- ④ 今後のデータ戦略を考えるにあたり、一律に法律や制度等の専門家で決定できないことも様々存在することから、研究分野からの課題抽出や整理ができるような場の創出を図る。
- ⑤ 研究課題を生かす制度設計や ELSI の課題の構築も並行して行う実務の定着を図る。

本ガイドブックが対象とする読者は特に下記に該当する方を想定している。

- データ連携に携わる実務者(国研・大学や民間企業等の研究者等)
- 企業等でデータ利活用の推進・戦略策定をリードする立場
(R&D 部門等の責任者等)

本ガイドブックを踏まえて、実務者が、データ連携・利活用を現場で推進する当事者として、関係者との調整や連携・利活用の具体的な進め方を認識し、R&D 責任者が決裁権者として、データ連携・利活用の意義や重要性・進め方を正しく理解し、組織単位でデータ連携に向けた投資意欲を高め、リテラシー向上や情報発信等を促進させていくことが重要である。

また、研究代表者をサポートする立場である研究資金配分機関の担当者も本ガイドブックが想定する読者である。関係者が密接に連携しながら取り組むことが重要である。

なお、これまでのバイオ分野のデータ連携・利活用は、専門領域に閉じた個別の課題への対応が中心であり、それに伴い、得られる知見の適応範囲も限定的

であった。したがって、領域横断的な課題の解決を射程に収めた、バイオ分野を俯瞰したガイドブックを策定するためには、データ連携・利活用のあり方について、実証を通じて具体的な事例を更に収集・分析しつつ、現状に即した最適解を導き出すという段階的なアプローチが必要である。このため、本ガイドブックは最終版ではなく、今後の進展等、必要に応じて改定を検討するものとする。

2. データ連携・利活用に関する基本概念

1. では、バイオデータの連携・利活用において生じている課題や今後取り組むべき重要性を述べた。データ連携・利活用に対する捉え方は、プレーヤの立場や目的によって様々であり、取組の中で関係者間での認識齟齬が生じることがしばしば生じる。「先端的な取組」、「特定のプレーヤを対象とした取組」などと難しく捉えるのではなく、データ連携・利活用が広く一般的な取組として普及していくことが期待される。

こうした実情を踏まえ、2.1 では、データ連携・利活用に対するイメージの具体化に資することを目的として、「データ連携・利活用ではどのようなデータがやり取りされるのか」「データ連携・利活用に関わるプレーヤは誰か」「データ連携・利活用のやり取りはどのような場で生まれるのか」を示す。2.2 では、取組の目的等に応じて取捨選択してデータ連携・利活用の実践に活用いただくことを目的に、データ連携・利活用を進める上で期待されるアクションや留意事項を示す。2.3 では、データ連携・利活用の実践に向けて、常に意識しておくことが期待される考え方として、「バックキャスト・フォアキャスト双方向のアプローチの重要性」や「関係者間でのコミュニケーションの大切さ」を示す。データ連携・利活用の捉え方はプレーヤの立場や目的によって様々であり、関係者間での共通認識を醸成するための一つの考え方として参考にしていきたい。

なお、国際的な連携や、国際的な評価を通じた投資や産業連携を呼び起こすために、本ガイドブックに網羅しきれない諸外国のルール等も踏まえ、国際的な標準的制度との整合性を議論することも求められる場合もある。したがって、全てこのガイドブックに記載しきれているものに議論が限定されるものではないことには十分に留意されたい。

2.1. データ連携・利活用とは何か

データ連携・利活用に取り組む上で、まずは「連携を行う主体（誰が）・連携するモノ（何を）・連携の進め方（どうやって）」から共通認識の醸成を図ることは重要である。特に、複数のプレーヤ間で取組を推進する場合は、認識のずれが生じやすく、取組初期からこうした認識合わせを行うことが大切である。

2.1.1. データ連携・利活用で取り扱うデータ

1.2 で述べたように、バイオ分野は、医療、健康、農林水産、食品など、数多くの専門領域を包含する分野である。それ故、バイオデータには、組織の競争力を左右するため秘匿性が高いもの、公共性が含まれるため公開されやすいものなど、さらに、データから将来知的財産が生成される可能性があって、かつデータの段階で確保しないと知的財産を国内で守ることができない場合のデータの取扱い¹⁵、個人情報の保護（利活用によって影響のある不利益な措置につながるか等の事項の検討）、及び安全保障の観点から共有の範囲を議論しなければならぬものなど様々な特性を有するデータが含まれる。

データ連携・利活用で扱うデータは、データの特性やオーナーの意思によってその共有範囲が定められるべきであり、押し並べてすべてのデータを開示・共有するというわけではない。

図表3では、データの属性を「単一の組織が創出したデータ」「複数組織間のアライアンスによって創出されたデータ」「政府研究開発プロジェクトで創出されたデータ（創出者が単一組織、または複数組織かは問わない）」に切り分け、具体的にどのようなデータが共有されやすいのか、または共有されにくいのか、難しいのかを整理した。例えば、「単一の組織が創出したデータ」について、オープン・クローズ戦略の一環で、自組織の競争力を高めるために共同研究相手に出すデータは比較的共有されやすいと想定されるが、創薬開発の過程で創出される臨床データなど、個人情報に該当するデータは、個人情報保護の観点から、共有は難しい¹⁶。また、「複数組織間のアライアンスによって創出されたデータ」

15 創薬のシーズ研究などにおいて、実験段階のサンプルの設計情報によって製造してもらう事業者情報提供する場合など、特許権利化する前に日本国外にしかそのような事業者がない場合、研究者だけの力では対応できないと考えられることから政策的な対応が求められる可能性がある。現場の実務者と政策担当者の継続的かつ率直な対話が望まれる。

16 なお、研究現場からの課題を個人情報保護法の改正や新たな立法についての議論の際に各種会議に共有できていないといった課題も指摘されている。法律家だけで議論せず現場の意見を適切に吸い上げる仕組みづくりが望まれる。

について、特定の業界全体で抱える課題解決に資するデータや企画・標準化等を策定するために必要なデータは、公共性が高く共有されやすい。他方、組織が自費で取得したデータや個社のノウハウを駆使して取得したデータの場合は、組織の競争力維持の観点から共有されにくい。さらに、「政府研究開発プロジェクトで創出されたデータ(創出者が単一組織、または複数組織かは問わない)」について、政府研究開発プロジェクト開始当初より共有を前提として取得されたデータは共有されやすいが、政府研究開発プロジェクトに参加する組織が研究開発の過程で取得したデータなど、組織の競争力に資するノウハウとして共有する場合は、共有されることは難しい。このように、データはその特性に応じて共有しやすさが様々であり、データ連携・利活用においても、共有しやすいデータから連携してくることが大切である。一方で、今後は、関係者間でデータの取扱いに関する事前協議踏まえて共通認識の醸成を図ることで、共有されにくい/難しいデータも含めたデータ連携・利活用が進んでいくことが期待される。なお、具体的な進め方については、データシェアリングポリシーの策定(3.2)、データマネジメントプランの策定(3.3)などを参照いただきたい。

	比較的共有しやすい	グレーゾーン	共有されにくい/難しい
単一の組織(下記等)が創出したデータ ・国研 ・大学 ・研究機関 ・公設試 ・民間企業 ・民間研究機関	<ul style="list-style-type: none"> ・オープンクローズ戦略の一環で、競争力を高めるために共同研究相手に出すデータ ・サプライチェーンの中で上流・下流で生産などに必要なデータ 【具体例】 <ul style="list-style-type: none"> ・国研のデータベースに登録されている医療データ ・国研が保有する種苗関連のリファレンスデータ 	<ul style="list-style-type: none"> ・業界横断の課題解決に資するが、個社の競争力に関わる可能性のあるデータ 	<ul style="list-style-type: none"> ・個社の競争力に関わるデータ ・各機関が有する個人に係るデータ(個人情報保護の観点から共有が困難) ・海外流出が懸念されるデータ 【具体例】 <ul style="list-style-type: none"> ・大学が有する医療関連データ(創薬開発の臨床データ等) ・国研が有する種苗関連データ
複数組織間のアライアンスによって創出されたデータ	<ul style="list-style-type: none"> ・業界団体の活動で取得したデータ/業界横断の課題解決に資するデータ ・規格・標準化を構築する際に必要となるデータ 	<ul style="list-style-type: none"> ・業界横断の課題解決に資するが、個社の競争力に関わる可能性のあるデータ ・政府研究開発プロジェクトに参加する機関間のアライアンスにより創出したデータ 	<ul style="list-style-type: none"> ・業界横断の課題解決に資するが、個社の競争力に関わるノウハウを駆使して取得したデータ ・個社の研究開発の過程で取得した失敗データ
政府研究開発プロジェクトで創出されたデータ(創出者が単一組織、または複数組織かは問わない)	<ul style="list-style-type: none"> ・政府研究開発プロジェクトで取得したデータ(プロジェクト開始当初より共有を志向しているデータ) ・業界横断の課題解決に資するデータ 【具体例】 <ul style="list-style-type: none"> ・国プロで大学が取得した人の健康データ 	<ul style="list-style-type: none"> ・政府研究開発プロジェクトに参加する特定個社が研究開発の過程で取得したデータ 【具体例】 <ul style="list-style-type: none"> ・政府研究開発プロジェクトで企業が取得した微生物資源データ 	<ul style="list-style-type: none"> ・政府研究開発プロジェクトで個社の競争力に関わるノウハウを駆使して取得したデータ ・個社の研究開発の過程で取得した失敗データ 【具体例】 <ul style="list-style-type: none"> ・政府研究開発プロジェクトで企業が取得した作物の遺伝子型データ

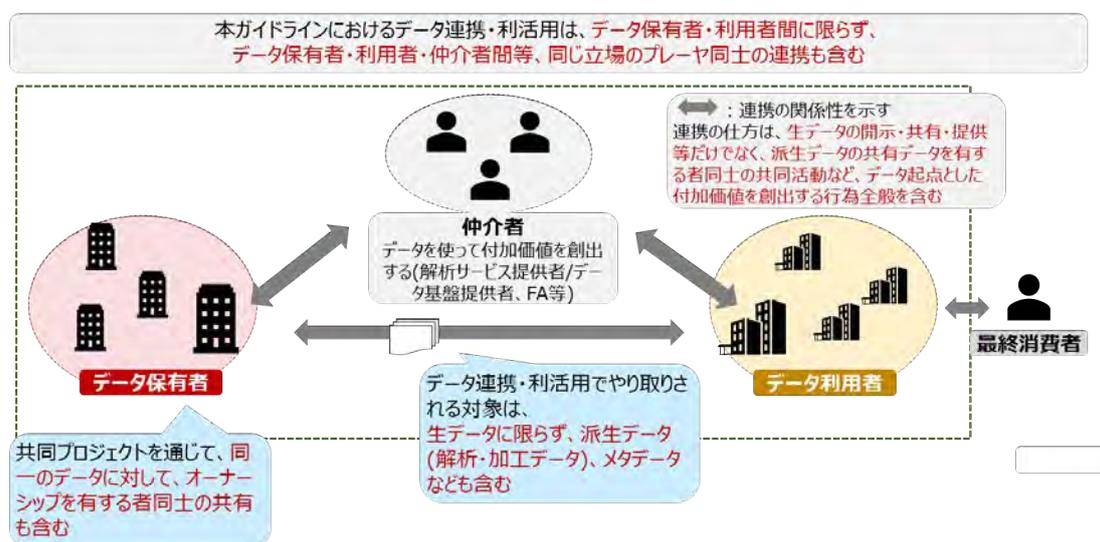
図表 3 データ連携・利活用の対象となるデータの整理

2.1.2. データ連携・利活用に関わるプレーヤ

前述の整理を踏まえると、データ連携・利活用には様々な考え方があるが、「2者以上の複数のプレーヤによるデータを起点とした協働行為」とも捉えることが出来る。協働行為とは、必ずしもデータ保有者が他者に対して、一方的にデータを開示・共有・提供するだけでなく、データ保有者が自身で有するデータを競争力の源泉として、他者に対して解析・加工した結果を提示することや、データを有する者同士が共同で新たなノウハウを生み出すことなど、データを起点とした付加価値を創出する行為全般を指している。

図表 4 で示すように、データ連携・利活用に関与するステークホルダには、データを取得し連携を推進するデータ保有者、データ保有者からデータ提供を受けるデータ利用者、データ連携を促す仲介者が存在する。データ連携の対象となるデータの種類の種類は、研究成果をそのまま記録した生データ、生データを解析・加工した派生データ、データに関する情報であるメタデータなどが想定される。

データ連携・利活用に関与するプレーヤや扱うデータは取組の目的に応じて異なり、連携の形は多種多様であるが、いかなる形であれ、データ連携・利活用の取組が増えていくことが期待される。



図表 4 データ連携・利活用のイメージ

また、データ連携を実現するには、研究の段階からある程度出口となりうる関連研究との更なる連携による応用研究の可能性や産業実装を想像した関連分野の情報も集めたうえでどのようなデータとすべきか、新しい技術に関して

は、研究の実施と並行しながら倫理や制度の議論も行うことが望ましい場合もある¹⁷。

2.1.3. データ連携・利活用が生み出される場

従来のデータ連携・利活用は、データ保有者が利用者候補に直接働きかけて連携の可能性を探り、保有者・利用者が1対1でデータのやり取りを行う形が主流であった。近年、データ連携・利活用が生まれる場は多様化しており、誰もが自由にアクセス可能なデータ集積基盤や、複数のプレーヤがデータを活用したサービス開発を行うサービス基盤など、N対Nのやり取りを通じてデータ連携・利活用の創発を目指す取組も見受けられる。

データ連携・利活用が生まれる場は図表5で示すように5つの階層で整理される。

- ・個々のプレーヤにより創出されたデータ・データベースを踏まえて、保有者・利用者ごとにやり取りが行われる場
- ・RDF化技術等を活用してデータが標準化・集積した場
- ・データの付加価値を目指して解析が行われる場
- ・データの利用用途の検討等を通じて、サービスの開発を促進させる場
- ・国研・企業等が保有するデータベース・解析基盤、プレーヤが集積した場

その上で、場として自立する際にその仕組みとしてのガバナンス¹⁸をどう考えるかも継続して検討する必要がある。

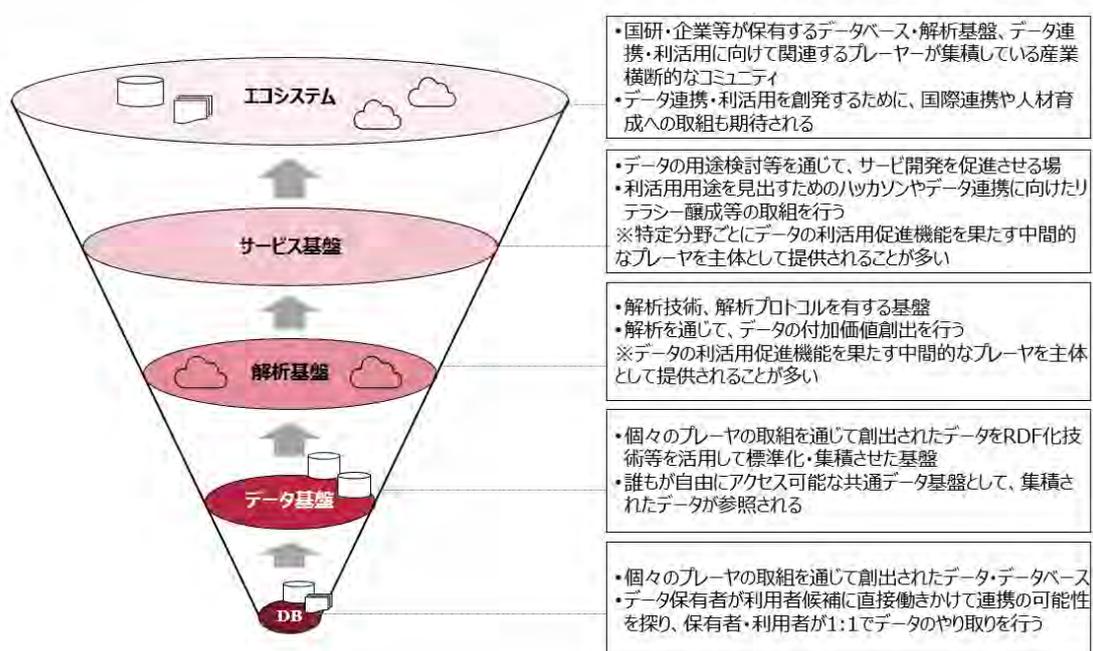
データ連携・利活用が生まれる場が多様化していることも踏まえて、先端的な取組が行われている国の研究開発プロジェクトをはじめ、動向を常に意識して知見を養うことも大切である。バイオデータの連携・利活用を行うためには、バイオ分野の専門知識の他、データベース構築、データ解析といったデータを正しく扱うための技術的な知見や、サービスを開発するための事業検討の知見など、

17 欧州では、単純に法律家と一部の有識者で法制度を議論するのではなく、European Health Data Spaceの実装に向けて、実際の取り組みと並行してデータ共有の在り方を議論するTEHDAS - Towards the European Health Data Space Joint Actionのプロジェクトを実施し各種調査を行い次々に各地域での実証の結果から得られている報告書がでている。<https://tehdas.eu/results/>

18 THADASの「Options for the minimum set of services for secondary use of health data in the EHDS」なども1つの考え方を整理している。また、日本人を対象とする研究に関するデータの取得にかかるインフォームドコンセントについて、単純に指針を項目ごとに遵守するのではなく、ルールや倫理の徹底のための場の全体のガバナンスをどう考えるかも重要である。この点、WMAの台北宣言にもバイオバンク等のデータベースの組織としてのガバナンスにおいて留意すべき項目の記載がある（WMA台北宣言2016）

多様な知識・知見が求められる。情報収集やリスキリングなどを通じて、リテラシーを高めていくことも大切である。

なお、各プレイヤー独自の取組だけではリソースに限界があるため、今回の取組を参考としつつ、その限界を考えながら、こういった活用における支援を行う仕組みの議論も必要と思われる。例えば、米国 NIH に設置される専門部署である Center for Information Technology のような支援組織がその例として想定される¹⁹。



図表 5 データ連携・利活用が生み出される場

具体的な事例として、上記のデータ連携・利活用を生み出すための場に対応する、SIP 第2期 バイオ・農業での実際の取組、情報・システム研究機構データサイエンス共同利用基盤施設に設置されたバイオデータ研究拠点で検討されている計画を紹介する。

¹⁹ European Health Data Space の実証プロジェクトの tehdas においても各種プロジェクトに関して最低限のIT技術の支援の必要性が指摘されている。 <https://tehdas.eu/results/tehdas-suggests-minimum-technical-services-for-the-european-health-data-space/>

＜DBCLS のデータ基盤の取組＞

(「バイオ・デジタルデータ統合流通基盤の構築(4A)」のデータ流通基盤に該当)

SIP 第2期バイオ・農業のコンソーシアムの取組では、国研、大学を中心に取得されたバイオデータを RDF 化技術により標準化・集積させた共通データ基盤を構築している。本データ基盤は、SIP 第2期バイオ・農業の関係者に限らず自由にアクセスすることができ、データ取得が容易に行えるようになっている。この基盤にライフサイエンス統合データベースセンター(DBCLS)が SIP 第2期バイオ・農業を通じて取得したデータを組み込むことによって、プロジェクトデータの利用者が、自身が所望するデータに容易にアクセスすることが可能となる。

また、プロジェクト終了後に公開されるデータでもプロジェクト期間内はメンバー間のみでの共有する場合がありますことを考慮し、クローズドデータを扱えるようにしている。具体的には、ドロップレットを用いた有用微生物スクリーニング技術を開発しているグループのデータをクローズドデータとして格納している。さらに、関連する微生物に関する培地や表現型などのオープンな情報を取得できるようにしていることにより、次の実験条件の設計などを容易に行うことができる。今後は、既存の培地情報を比較・解析することによる新しい培地の設計などの応用も期待される。

＜ちとせ研究所のデータ連携・利活用に向けた取組＞

(「食によるヘルスケア産業振興(2B)」の取組に該当)

SIP 第2期バイオ・農業 「食によるヘルスケア産業振興(2B)」は、データ駆動によるヘルスケア産業の創出に向けて、腸内マイクロバイオーームデータを中心としたデータ基盤²⁰を踏まえて、食・健康サービスの開発や統合健康栄養食品の民間認証等を推進している。

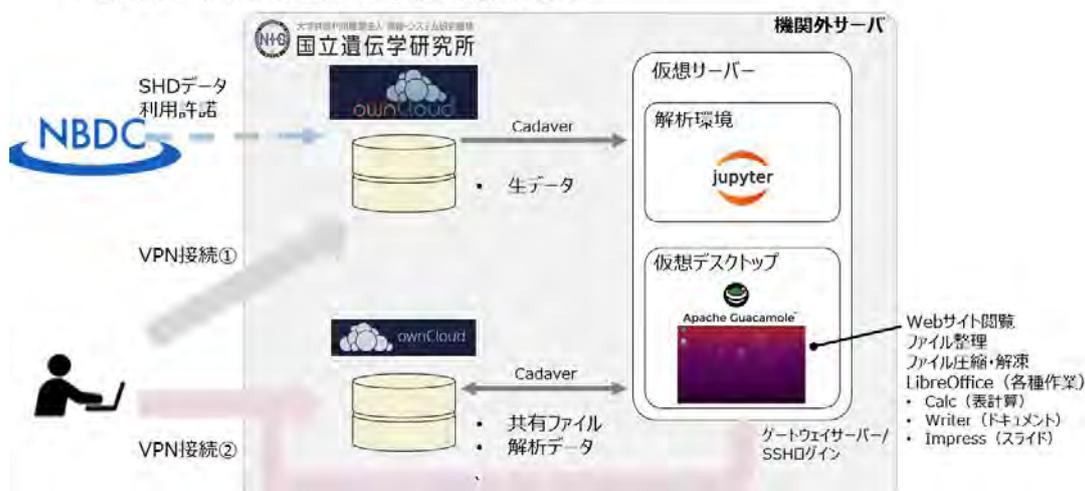
参画する民間企業である株式会社ちとせ研究所(ちとせ研究所)は、国立遺伝学研究所と共同で国立研究開発法人科学技術振興機構バイオサイエンスデータベースセンター(NBDC)より認められている機関外サーバーである国立遺伝学研究所のスーパーコンピューター上にデータ保管・転送・解析・結果出力プラットフォームを構築した。この解析プラットフォームは、セキュリティ・アクセス権限等に配慮して管理されており、参画する研究機関、大学、企業等が利用可能な

20 「食によるヘルスケア産業振興(2B)で集積されたデータの共有を目的としたデータベース」
<https://gr-sharingdbs.biosciencedbc.jp/shd>

基盤²¹となっている。ちとせ研究所はこの解析プラットフォームを単に提供しているのみならず、企業等との共同での解析結果創出に向けた支援等を実施しており、データ解析者のスキルに合わせた利活用が実証されている。この実証を通じて、ちとせ研究所は、データ取得と事業構築の間には大きなギャップがあることを明らかにした。そこで、データ連携・利活用を通じたデータ駆動によるヘルスケア事業の事業化・社会実装を専門的に支援する立場として、SIP 第2期バイオ・農業に参画しているヘルスケア企業、食品企業等に対して、データの利活用用途を見出すための場（ハッカソン・アイデアソン等）の企画・実施やデータ利活用リテラシー醸成に資するデータ解析手法のノウハウ構築・提供のみならず、実証支援等を通じたエコシステム形成を行うことを目的として、産学官を繋ぐデータ利活用促進のための取組を進めている。

遺伝研スパコン解析PF

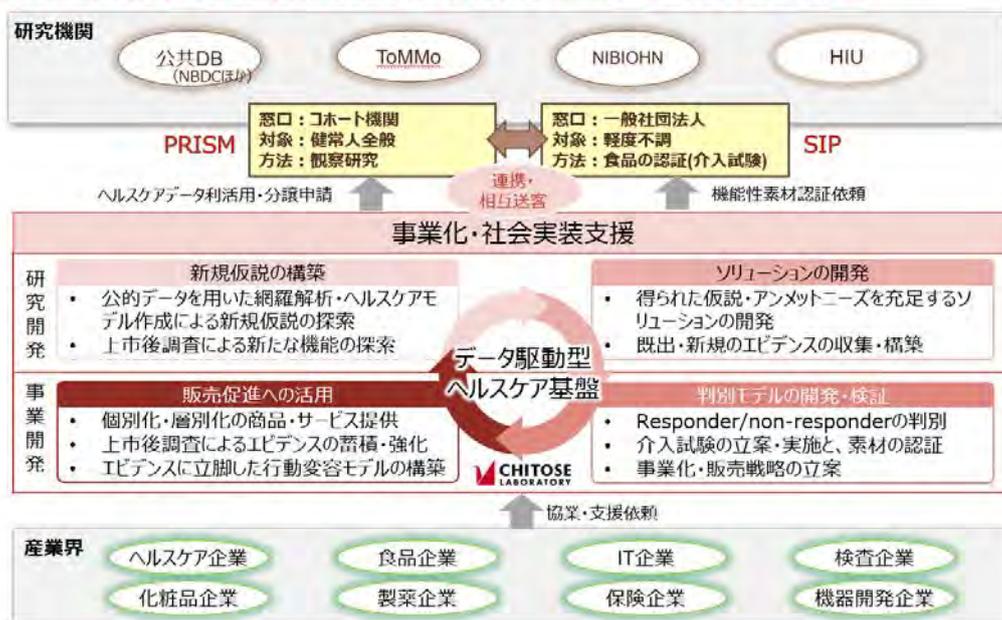
- NBDCで認められている機関外サーバとして、国立遺伝学研究所のスパコン上に解析基盤を構築し、企業（10社）に提供しました
 - ・ 機関外サーバを使うメリット1: サーバ環境依存による解析ツール動作不具合を回避できること。
 - ・ 機関外サーバを使うメリット2: セキュアに解析できること。



図表 6 ちとせ研究所が遺伝研と共同で開発した解析プラットフォーム

21 「国立遺伝学研究所とちとせグループ マイクロバイオームデータ利活用促進に向けたデータベース機能拡張について提携を開始」
<https://prtmes.jp/main/html/rd/p/000000026.000034251.html>

実証事例：SIP-PRISM連携で実現するバイオ戦略実現のための取り組み



図表 7 ちとせ研究所を主体としたデータ連携・利活用に向けた取組

注：図表 7 の略称は下記を示す

- ・ ToMMo(東北メディカル・メガバンク機構)
- ・ NIBIOHN(国立研究開発法人 医薬基盤・健康・栄養研究所)
- ・ HIU(学校法人 電子開発学園 北海道情報大学)

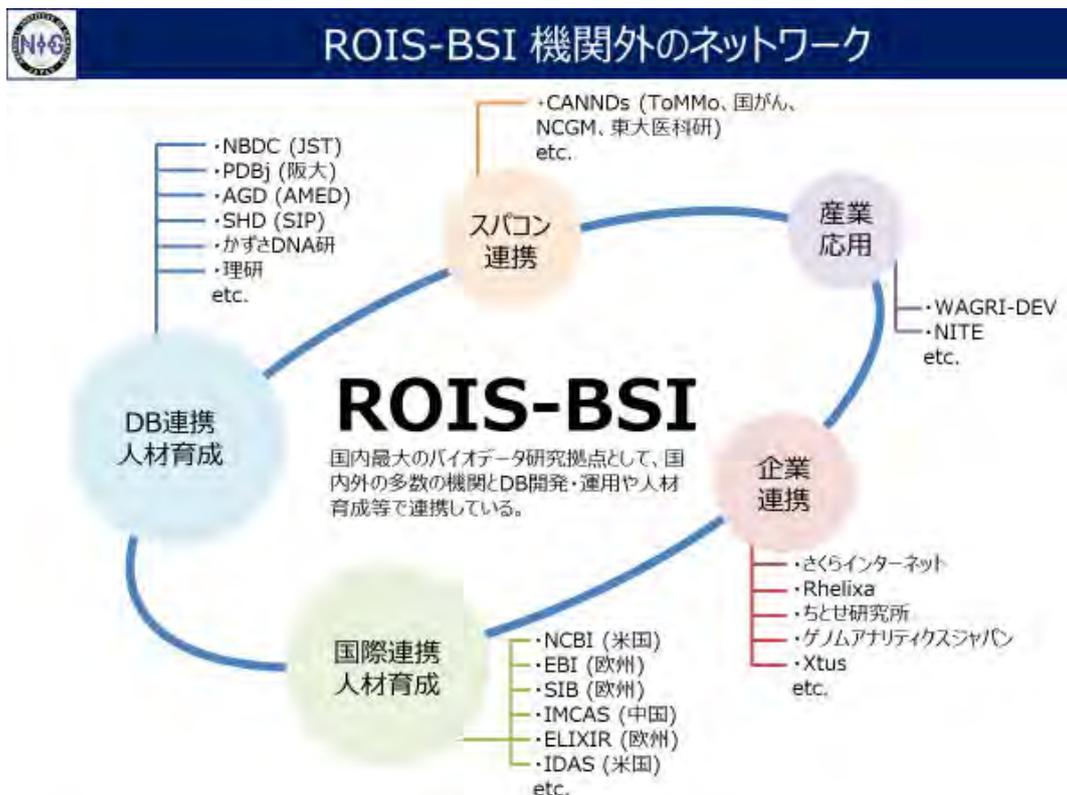
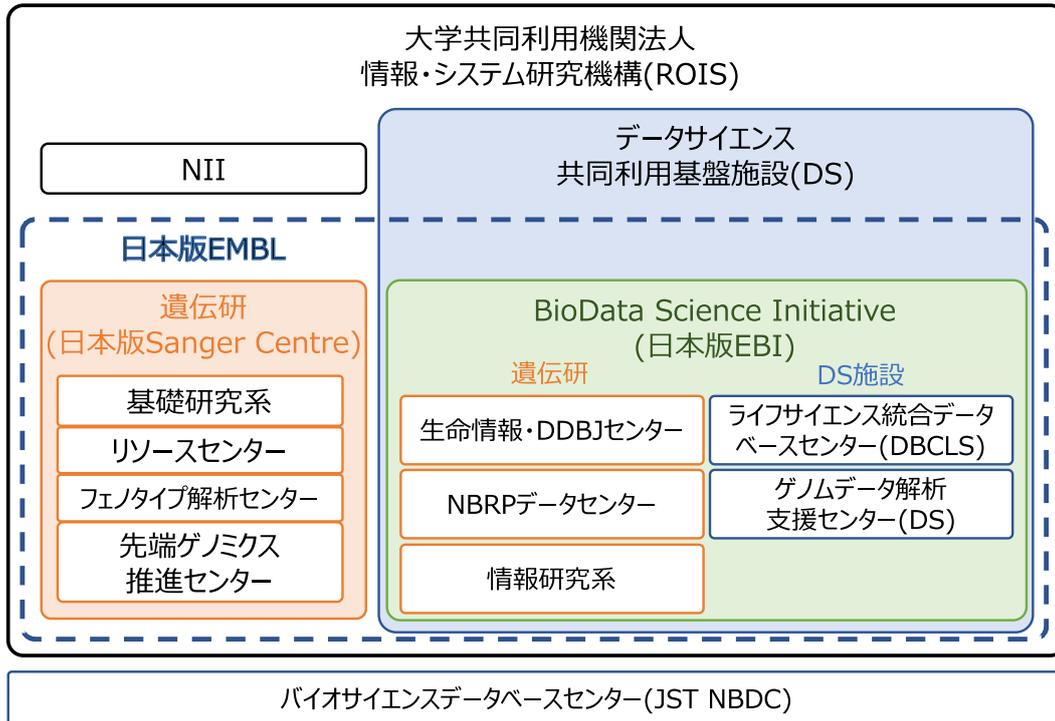
<情報・システム研究機構データサイエンス共同利用基盤施設

バイオデータ研究拠点(BSI: BioData Science Initiative) >

国立遺伝学研究所内の生命情報・DDBJセンター(DDBJ:日本DNAデータバンク)と、データサイエンス共同利用基盤施設内のDBCLSが中心となり、日本のバイオデータ情報基盤の中核拠点として2022年度末にバイオデータ研究拠点(BSI)が設立された。BSIではビッグデータレポジトリを扱う生命情報・DDBJセンターとデータ統合基盤技術を開発するDBCLSが連携(2025年度に統合予定)することにより、膨大な生命科学データに知識を紐付けた上でバイオデータサイエンスの発展に貢献できる研究基盤を構築し、国内外に広く提供する。また、バイオデータサイエンティスト養成のための人材育成コースの設置なども計画している。

日本のバイオデータを扱う科学技術振興機構(JST)、大阪大学、日本医療研究開発機構(AMED)、理化学研究所(RIKEN)、製品評価技術基盤機構(NITE)や海外の

バイオインフォマティクス研究機関である米国生物工学情報センター、欧州バイオインフォマティクス研究所などとも連携するとともに、国立遺伝学研究所が持つスーパーコンピューターと国内の研究機関が持つスーパーコンピューターとの連携によるデータ解析基盤も提供する。今後、上記のようにクローズドにまとまるのではなく、より拡張していく議論を諸外国の取組の実施内容との比較検討を踏まえて取組が拡大されることが期待される。



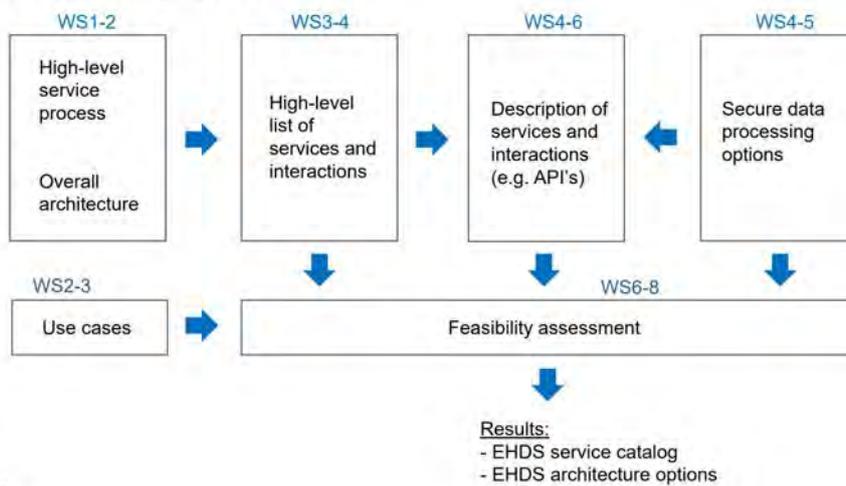
図表 8 BSI の取組

参考: 諸外国での利活用を想定したワークショップの例

欧州の European Health Data Space における実践を通じた議論の進め方について、ワークショップのプランと当該ワークショップで議論されたシナリオの事例 (TEHDAS - Towards the European Health Data Space Joint Action Report on EHDS architecture and infrastructure implementers' expectations / experiences より引用) においては、IT インフラも視野に入れた技術的にも踏み込んだ議論がなされている。



Workshop plan

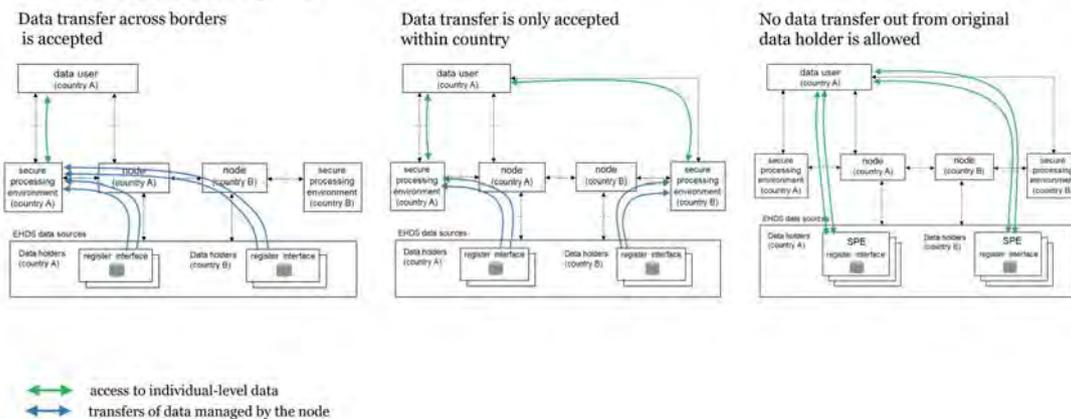


6

SITRA



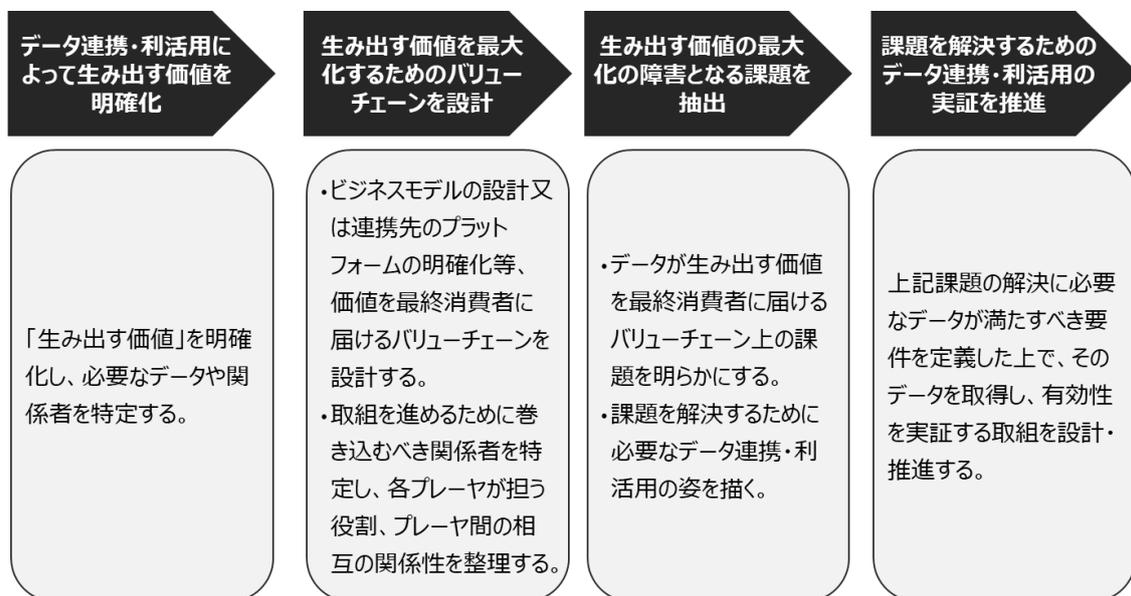
Policy scenarios under discussion in workshop 5



2.2. データ連携・利活用を成功に結び付けるメソッド

2.1 では、データ創出・利活用に携わるあらゆる関係者に対して、データ連携・利活用を考える際に前提として持つておくべき基礎概念を示した。

データ連携・利活用は、あらかじめ定義された一連のプロセスを実行すれば必ず上手くいくというものではなく、計画・実行・評価・改善のプロセスを繰り返すことが重要である。2.2 では、データ連携・利活用の効果的・効率的な推進に必要な一連の取組を示す。取組の目的に応じて、取捨選択して実践に活用いただきたい。



図表 9 データ連携・利活用を成功に導くために必要となる具体的な取組

2.2.1. データ連携・利活用によって生み出す価値を明確化

<目標>

取組の結果として「生み出す価値」を明確化し、それに必要なデータや関係者を特定する。

<取組内容>

「生み出す価値」の明確化のため、以下の取組を検討し、関係者の考えやニーズを集約する。

- データ利活用を希望する機関と「生み出す価値」の双方が明確である場合、

「オープンデータ官民ラウンドテーブル」等に準じラウンドテーブルミーティングを実施し、必要なデータや取り組むべき課題を特定する。

- データ利活用を検討している機関が明確であるが、「生み出す価値」が明確ではない場合、諸外国での実証実験的な取組事例を参考としつつ、アイデアソンやハッカソンを実施し、予備的なデータを収集・分析する。(データを起点としたビジネス機会の構築に至っていない状態であるため、予備的なデータを用いて「生み出す価値」を具体化することは効果的)。
- データ利活用を検討している機関及び「生み出す価値」の双方が明確ではない場合、関連分野の企業等を対象とした情報収集（ヒアリングやインタビュー）を実施し、「生み出す価値」の明確化を図る。これらの情報収集では不十分な場合には、諸外国での実証実験的な取組事例を踏まえた上で、アイデアソンやハッカソンを実施し、「生み出す価値」を明確化し、必要なデータや関係者を特定する。

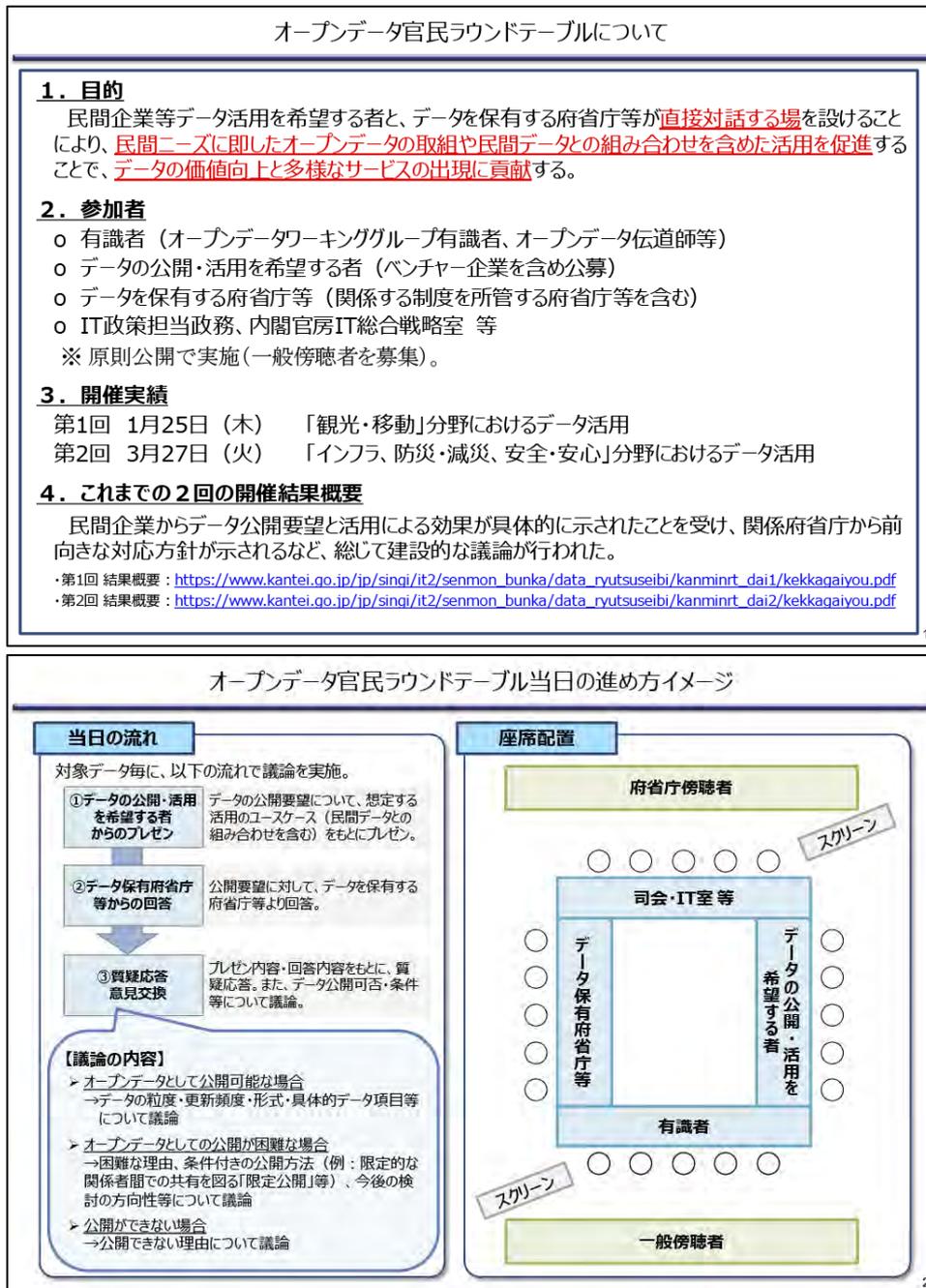
<留意点>

- DBCLS が定期的にバイオハッカソン²²を実施しており、それを参考に SIP 第 2 期バイオ・農業においても実施した例があるので必要に応じ参考にできる。なお、アイデアソンやハッカソンは参加者の技量や動機によって得られる成果が左右されるため、国際的な取組と矛盾しないように諸外国での実証実験的な取組事例を参考にしながら、我が国の実態に適合した対応を行うことが有効と考える。
- また、アイデアソンやハッカソンの開催にあたっては、成果の帰属等についての参加者の共通認識を形成することが重要である。これについては経済産業省「AI 人材育成のための企業間データ提供促進検討会」が作成した「AI・データサイエンス人材育成に向けたデータ提供に関する実務ガイドブック」に開催委託に関するモデル契約書や参加者の同意書例が提示されているが、案件毎に事情が異なるため参考としつつ法律の専門家の助言を経て対応することが望ましい。
- さらに、領域やデータの内容によっては、参加者が対象となるデータの解析に適切な能力を持たない場合も想定されるため²³、必要に応じ解析をサポート可能な体制を整備した上で実施することが望ましい。
- また、生み出す価値を議論するにあたり、利用する目的を充足する品質の確

22 <https://wiki.lifesciencedb.jp/mw/>

23 機械学習によるビッグデータ解析においては従来の統計学の範疇にとどまらない解析能力が必要となる場合がある。

保の観点も重要である。諸外国でもデータベースの中で出てくるデータの品質の議論とその向上に関する将来的な取組の方向性についても議論することが望ましい²⁴。



図表 10 第3回 EBPM 推進委員会資料よりラウンドテーブルミーティングについて²⁵

24 欧州 TEHDAS のデータ品質の推奨事項：<https://tehdas.eu/results/tehdas-develops-data-quality-recommendations/>

25 https://www.kantei.go.jp/jp/singi/it2/ebpm/dai3/siryoul_2.pdf より引用

2.2.2. 生み出す価値を最大化するためのバリューチェーンを設計

<目標>

ビジネスモデルや連携先の明確化等、価値を最終消費者に届けるバリューチェーンを設計する。また、取組を進めるために巻き込むべき関係者を特定し、各プレーヤが担う役割・役割、プレーヤ間の相互の関係性を整理する。

<取組内容>

以下の情報を集約し、バリューチェーンの設計を行う。

- 収集・生成・集積、処理、連携、分析、活用、販売について定義を行い、それぞれの具体的なタスクと担当する機関を明確化する。
- バリューチェーンの各段階で付加される価値と必要なコストを算出する。それを用いて競争優位の源泉となる活動(強み)を明確化する。

<留意点>

当該バリューチェーンを通じて社会実装する製品・サービスのサプライチェーン構造も把握し、実装後に予想されるインパクト分析を行う必要がある。データを解析し、知見を得るだけでは市場獲得は難しく、インパクト分析を行うことによって「生み出す価値」が及ぼす影響が可視化され、市場獲得のイメージが明確となる。

バリューチェーン設計を行うにあたり、2者間のデータのやり取り等であれば、比較的バリューチェーンが設計しやすい一方で、データのやりとりがマルチになる場合は、バリューチェーンが複雑化し、設計の難易度も高まる。共通認識を醸成しながら進めることが重要であり各プレーヤが担う役割・機能、相互の関係性、情報やお金の流れなどを、言語化・可視化することが必要である。また、バリューチェーンを形成するに際し、集めたデータのプロトコルを新たなニーズに合わせた品質やプロトコル等のアップデートが必要な場合があると考えられ、科学的な見地からその必要性を議論するアドバイザリーボードの設置も考えられる。

2.2.3. 生み出す価値の最大化の障害となる課題を抽出

データ連携・利活用においては、技術、規制、組織文化など、様々な課題があることが考えられ、また、その解決策も多岐にわたるため、事前に全てを網羅し個別に議論することは困難であるが、可能な限り想定した上で活動を行うことが重要である。

<目標>

- 「生み出す価値」の実際の創出及び最終消費者に届けるバリューチェーン上の課題を明らかにする。
- その課題を解決するために必要なデータ連携・利活用の姿を描く。

<取組内容>

特性要因図(図表 11 参照)を作成し、データ連携・利活用やバリューチェーンの構築を阻害している要因の全体像を把握する。その上で、取り組むべき課題を明確化し、解決策の仮説を構築する。

解決策の仮説の中で、データ連携・利活用の方法を工夫する事で解決できる課題をまずは取組対象とする。データ連携・利活用の工夫で解決することが適切ではない課題²⁶については、別のアプローチの検討が必要であるが、その際には必要な人的・経済的資源を別途確保する。なお、特定の参加機関にのみ利益がもたらされる場合は、競争領域となるため取組の対象外とする。

<留意点>

データ連携・利活用ありきで議論すると本来の目的を見失ってしまうおそれがあるので、全ての課題解決に対して必ずしもデータ連携・利活用が最適ではないことを念頭において検討する必要がある。まずは課題の解決についてデータ連携・利活用による解決を前提とせず解決策を模索した上で、それでもデータ利活用による解決が最適であるとの結論が得られた場合に取組の対象とすべきである。

例えば、企業のデータ利活用に対するニーズは明確であるものの最終消費者またはサプライチェーン先の当該企業の顧客企業のニーズが明確では無く、構

26 例えば、ゲノム・データを利活用しゲノム編集を用いて高付加価値の食品を作る取組において、社会受容性の獲得は重要な支援活動であるが、データ連携・利活用を工夫して直接的に解決するものではない。

築すべきバリューチェーンが明確にならない場合、まずは前者のゴールに向かって検討を進めつつ、後者については消費者ニーズまたはサプライチェーン先の当該企業の顧客企業のニーズを明確にするためにアンケート結果のデータの解析や、少数の最終消費者に対するデプスインタビューやグループインタビューによる把握等最適な方法を検討することが必要である。

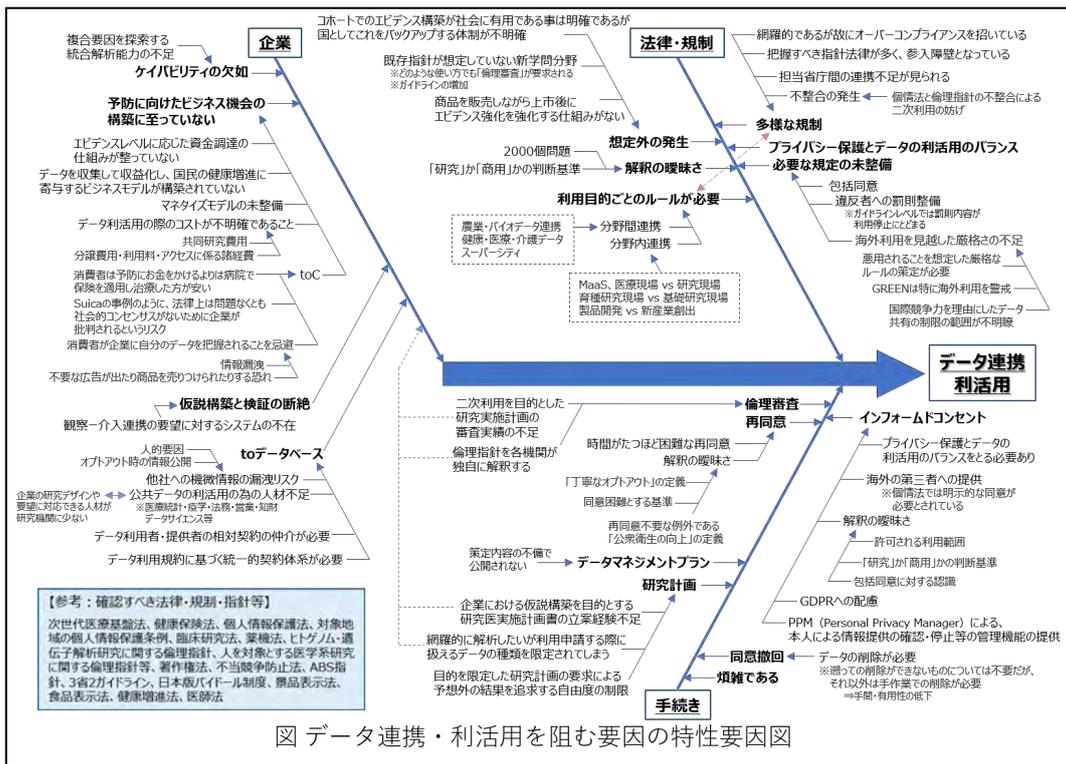


図 データ連携・利活用を阻む要因の特性要因図

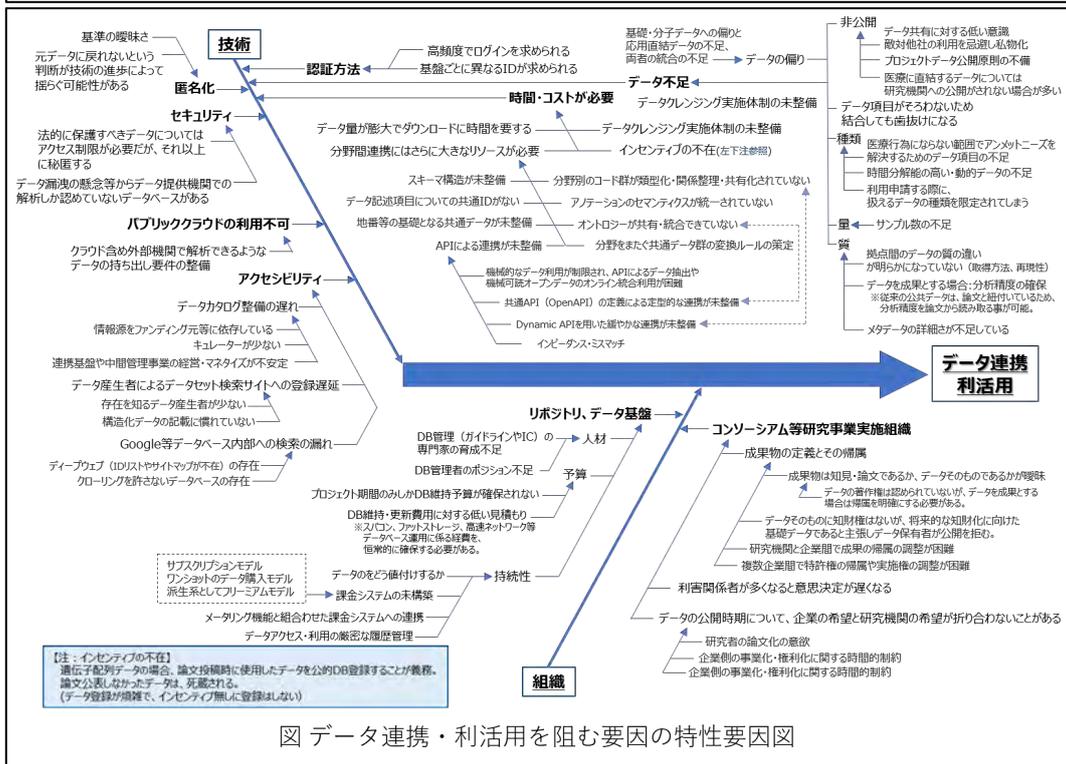


図 データ連携・利活用を阻む要因の特性要因図

図表 11 データ連携・利活用に関する特性要因図

2.2.4. 課題を解決するためのデータ連携・利活用の実証を推進

<目標>

上記課題の解決に必要なデータを取得し、有効性を実証する取組を設計・推進する。

<取組内容>

バリューチェーンの実現に向けた課題が、例えば「ユーザー企業や機関の獲得」であった場合、解決策として、制限公開可能なトライアルデータを使えるようにすることが一案である。そこで、それを可能とする2次利用条件の設定²⁷やデータホルダーとの事前合意、加えて、健康・医療寄りのデータであれば研究対象者の事前同意の取得といった要件を具体的に掲げ、対象となるデータの利活用について検討する。このように、課題を解決するために必要なデータの要件と取組を明確にし、それらを漏れなく含む実証計画を立案・推進することが重要である。

また、実証を効果的に推進するために、複数のプレーヤ間で協調しながら取組むことができる国の研究開発プロジェクトを活用することも一案である。国の研究開発プロジェクトは、通常、異なる立場のプレーヤが多数参画するため、データの取扱い、成果の権利帰属先の整理など、調整事項も生じる。3.2.～3.に示すデータマネジメントプランやシェアリングポリシーなどの策定を通じて、コンソーシアム内で共通認識を醸成しながら取組を進めていくことが期待される。また、国の研究開発プロジェクトの成果は公共性が高いことから、参画している企業等からのデータ共有も期待できる。

<留意点>

課題とその解決策は多岐にわたるため、事前に全てを網羅し個別に議論することは困難である。そこで、これまでに検討・策定された各種施策・ガイドラインなど先行事例を踏まえ、直面する可能性が高い課題とその要因、課題解決に資する取組等の把握に努めることが重要である。

²⁷ 健康医療データに関しては、どのような2次利用を想定するのか議論を行い対象となるユースケースも想定して制度設計を行うべきである。諸外国でも法律だけで紋切り型で仕組みを議論していないことに留意すべきである。TEHDASのDescription of steps in accessing individual-level data for national and EU researchers in a selection of centralised systems and decentralised systemsでの議論等を参照。

2.3. データ連携・利活用を効果的に進めるためのヒント

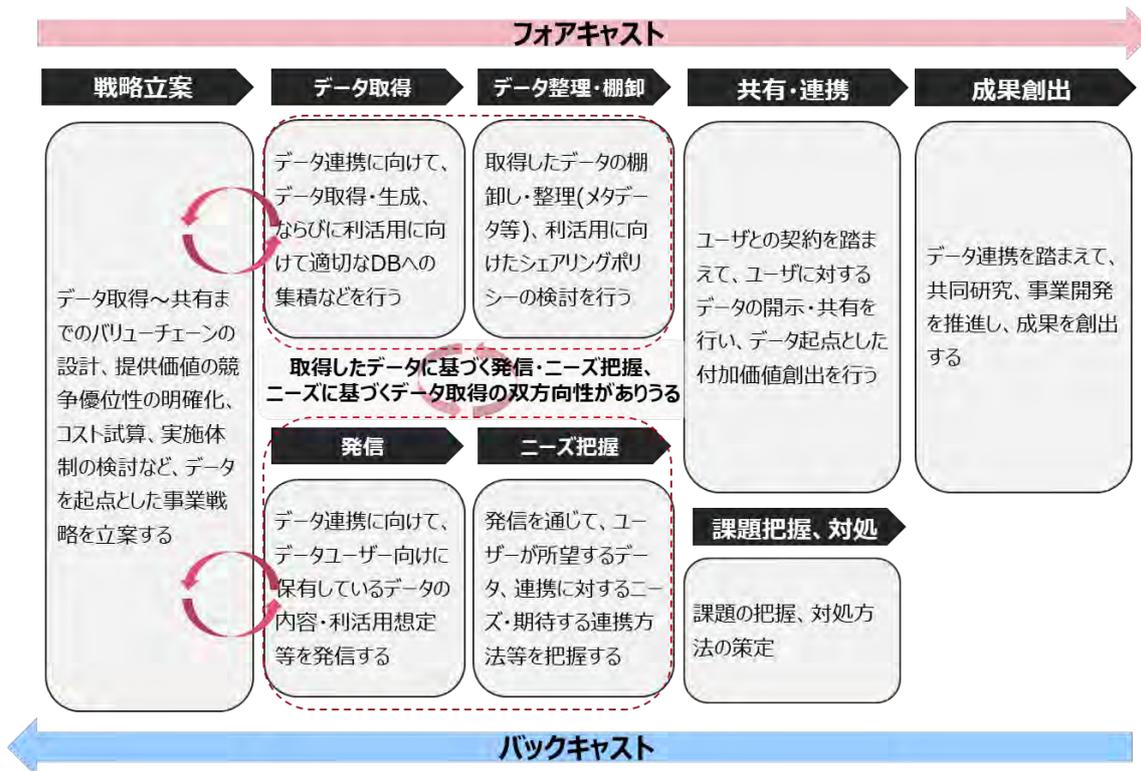
2.2 で示した一連の取組を行う上で、データ取得が優先されることで、利活用ニーズを踏まえて必要なデータを取得するバックキャスト視点の検討が後回しになってしまうことや、将来のデータ連携に向けたステークホルダとの調整がおざなりになってしまう等の課題が想定される。2.3 では、想定される課題に対して、フォアキャスト・バックキャストの視点、ステークホルダとのコミュニケーションの視点から、取組を一貫して意識すべき考え方を示す。

2.3.1. フォアキャスト・バックキャストの双方向のアプローチ

データ連携・利活用に取り組む際のアプローチは、データを取得した後にデータの価値や利活用用途を検討する「フォアキャスト型の取組」と、事前に目的や利活用用途を定めた上で必要なデータを取得する「バックキャスト型の取組」に分けて考えられる。

データ保有者が利用者候補に直接働きかけて、1対1で連携の可能性を探ることが主流であった時代は、保有者のニーズ把握が容易でないことから、フォアキャスト型の取組を中心にデータ連携・利活用が行われてきた。近年では、2.1.3 で記載した通り、複数のプレーヤ同士が、データ連携基盤やサービス基盤などを介して、N対Nで連携を行う形態が普及しつつある。こうした背景に伴って、データ連携がデータ保有者により大きな成果をもたらすことが期待され、予め利用者ニーズを見定めた上で、戦略的にデータ取得を行うバックキャスト型の取組の重要性が増しているとも考えられる。

上述のように、データ連携・利活用のあり方の変化に従って、フォアキャスト型からバックキャスト型への転換が起きているとも捉えられるが、データを取得する過程で利活用の方向性がより具体化することで、取得すべきデータが新たに見えてくる可能性もある。そのため、双方向のアプローチを意識しながら、フォアキャスト型、バックキャスト型のどちらのアプローチが適しているかを取組の目的に即して選択することが大切である。



図表 12 データ連携・利活用のアプローチ

2.3.2. データ連携の加速に繋がるコミュニケーション

一般的に、データ連携・利活用に取り組む際には、立場や属性が異なる様々なプレーヤと関わることになるため、関係者間での調整や共通認識の醸成が特に大切であり、適切にコミュニケーションを図ることが期待される。コミュニケーションが適切に図れていない場合、データ保有者・利用者間でのデータの内容や形式に対する認識のずれや、データの共有要件や成果の権利帰属関係の調整が上手くいかないなど、データ連携を阻害する事象が生じることが懸念される。したがって、連携の初期段階から関係者間でのコミュニケーションを図る機会を適切に設けることが大切である。

コミュニケーションを図る方法は、目的に応じて様々な方法が考えられるが、特定のテーマに対して参加者が意見を出し合うアイデアソンは、その自由度の高さからコミュニケーションの起点となり得る。アイデアソンを行うことによって、多様な立場の参加者による自由闊達な議論を通じて、様々な視点から意見を得ることが期待できる。例えば、データの利活用用途を見出すことを目的としたアイデアソンの場合は、題材となるデータと親和性がある分野の関係者を集

めて、各参加者から、興味・関心や取組を踏まえた利活用のアイデアを発散させる場を設けることも一案である。その他、データ連携を進める上で、データの利活用用途の探索以外に、関係者の懸念を予め抽出するために、想定される課題や対応策を事前に議論する場としてアイデアソンを活用することも有効である。

なお、こうしたコミュニケーションは、ステークホルダ間で継続的に行っていくことが重要であり、議論を活性化させる観点からは、コミュニケーションを図るための方法の検討やディスカッションの場の設計・開催等を大学 TL0・コンサルタントなどの外部機関に依頼することも一案である。また、場の開催に際して、参加者との調整や準備に労力が掛かることから、自組織以外のリソースを上手く活用しながら取組を進めていくことも有効である。

下記に、ディスカッションの場を設計・開催する際に留意すべきポイントを示す。

○ディスカッションの活性化を促す事前準備

コミュニケーションの場を設計する際に、単純に参加者を集めてテーマを設けるだけでは、ディスカッションの活性化を図ることは難しい。アイデアの創発を促し、参加者に積極的に意見を発信いただくために、ディスカッションの題材に関する事前の情報提供を行うことも効果的である。特に、データの利活用用途を見出すことを目的とした議論等、想像力が求められる場面では、参加者が事前にアイデアを思案する機会が提供できることも期待される。例えば、ディスカッションの題材となるデータセットに関する情報を予め共有しておくこと等も一案である。また、アイデアの創発を促すために、特定の産業における技術動向やデータ連携・利活用によって成果創出に至った事例等を参考情報として提示することも有効であると考えられる。

○ディスカッションを効果的に実施するためのファシリテーション

ディスカッションの参加者が立場や所属の異なる複数のプレーヤから構成される場合は、各参加者の志向や興味・関心に従って、多様な観点から意見が出されることになる。その場合、議論が発散する等、当初のディスカッションの目的から逸脱する懸念がある。対応策として、参加者への事前アンケート等を通じて、各参加者の興味・関心事や発信しようとしているアイデアを予め収集し、議論の流れを想定しておくことも一案である。また、ディスカッションの冒頭で、議論の目的やゴール、意見を出して欲しい観点を改めて明示することも効果的であると考えられる。ディスカッション当日は、議論がしやすい人数

(5名程度)の小グループに分かれ、各グループにファシリテータを配置することが期待される。イーゼルパッドや付箋などを活用することで、参加者から出た意見が可視化・整理され、議論の活性化が期待される。こうしたツールも活用しながら、議論の発散と収束のバランスを意識したファシリテーションを行うことも大切である。



図表 13 ディスカッションを効果的に実施するためのツール

<事例：SIP 第2期 バイオ・農業のアイデアソン>

複数のコンソーシアムから参加者を募り、当課題内でのデータ連携・利活用促進に向けたアイデアソンを実施した。ディスカッションの場は、目的や参加者の属性に応じて様々なやり方が考えられるため、あくまで場の作り方の参考事例として、実施結果を紹介したい。

○概要

特定のコンソーシアムの成果(データセット)に対して、下記2つの観点からディスカッションを行った。

- ① 利活用用途のアイデアソンを通じた潜在的ニーズの探索
- ② 国の研究開発プロジェクトで創出されたデータの産学官連携を踏まえたデータ連携・利活用の在り方

ディスカッションの題材と親和性の高い4つのコンソーシアムから、国立研究機関、民間研究機関・民間企業などに所属する約20名が参加し、対面・オンラインの形式で実施した。

○実施内容

2.3.2で述べた「ディスカッションの場を設計・開催する際に留意すべきポイント」に留意して実施した。利活用用途のアイデアソンを通じた潜在的ニーズの探索の議論を活性化させるべく、題材に関する概要(研究開発成果の概要、保有するデータの利活用実績等)を事前情報として参加者に共有するとともに、事前アンケートを通じて考えられるデータ連携・利活用のアイデアを予め収集した。事前アンケートでは、利活用のアイデアに加えて、参加者がSIP第2期 バイオ・農業の取組を通じて創出した研究開発成果、データ連携・利活用を目指す上で感じている課題等も収集し、参加者の取組内容や興味・関心等を踏まえて、当日ディスカッションを行うグループ分けや議論の流れの想定を行った。

アイデアソン当日は、各グループに事務局または題材の提供者がファシリテータとして参加した。また、各産業で生じている主たる技術動向と関連するデータ利活用事例を整理し、参考情報として提示した。こうした情報提供によって参加者のアイデア創発にきっかけを与え、自身が専門とする領域を超えて活発な意見交換が実施された。

産業分類	動向	事例
農業、林業	<p>◆農業DX化、FaaS</p> <p>農業者の高齢化や労働力不足が進む中、デジタル技術を活用して効率の高い営農を実行しつつ、消費者ニーズをデータで捉え、消費者が価値を実感できる形で農産物・食品を提供していく農業（FaaS: Farming as a Service）への変革を推進する動向が見られる。</p>	<p>◆農作物の生育データ活用による収穫予測</p> <p>ローカル5Gに接続された農機からリアルタイムに取得した農作物の高解像度画像を活用し、農作物の収穫予測に活用</p> <p>例えば、AIにより収穫適期のいちごを検出することで事前に収量を予測</p>
漁業	<p>◆暗黙知(経験知)の形式知化</p> <p>衛星からの観測データや解析データ、その他様々なデータを統合し、旧来の漁業従事者の「経験則」や「勘」をシステム化し、汎用化することで、漁業そのものの効率性を高める動向が見られる。</p>	<p>◆衛星データ活用による漁業効率化</p> <p>独自の「スマートセンサブイ」から収集した気温、気圧、水温、水圧、潮流、塩分濃度といったデータと「スマートカメラブイ」から収集した映像データの2つのデータと漁獲量のデータとの相関関係を導くことにより、漁師の勘をデータとして見える化し、経験の少ない漁師でも漁獲量を増やせるようなシステム構築に活用</p>
鉱業、採石業、砂利採取業	<p>◆IoT・衛星データを用いた作業高度化</p> <p>機器から取得したデータによる生産性向上、衛星データ活用による、これまで把握しきれなかった地理的条件(鉱区)、作業状況管理などに繋げる動向が見られる。</p>	<p>◆掘削機器の使用状況データ活用による機器故障予防</p> <p>掘削機器や鉱山内のセンサーから取得したリアルタイムデータを活用して、採掘場において遠隔地から採掘機器の故障の予防保全を行う。</p> <p>◆衛星データを活用した鉱業地帯の効率的な管理</p> <p>衛星データを用いて鉱業地域全体の地理的条件を観察することで、作業安全性確保、作業場所の明確化、違法工業地域を特定に活用</p>
建設業	<p>◆IoT・ヒトのバイタルデータを用いた作業高度化</p> <p>・作業現場におけるリスク回避のためのバイタルデータ活用による現場作業の高度化に関する動向が見受けられる。</p>	<p>◆作業員のバイタルデータ活用による建設現場の安全管理</p> <p>作業員の位置情報や健康情報をデジタル上で管理することで、現場作業の効率化を行うとともに、作業員の健康情報を踏まえた休憩タイミングの連絡等安全管理を実施</p>
電気・ガス・熱供給・水道業	<p>◆ライフライン使用データに基づく新たなサービス開発</p> <p>ライフラインの使用状況を計測するだけでなく、収集したデータを活用・解析新たなサービスを生み出す動向が見られる。</p>	<p>◆ライフラインの利用状況データ活用による高度な防犯対策</p> <p>電気・ガス・水道等の使用量をデジタルで計測する通信機能を備えたスマートメーターから取得した各住宅の30分毎の電力使用量計測結果を、ライフサイクルにあった広告・マーケティング戦略、空き家特定・高度な防犯対策へ活用</p>

図表 14 産業動向、データ連携・利活用事例イメージ

国の研究開発プロジェクトで創出されたデータの産学官連携を踏まえたデータ連携・利活用のあり方に関する意見交換でも、事前アンケートで収集したデータ連携に対する課題感を踏まえて、小グループでのディスカッションを行った。参加者の課題感が多岐に渡るが、発散と収束の時間帯を区切ってファシリテータが議論をマネジメントすることで、建設的な意見交換が実施された。

イベント終了時は、事後アンケートを実施し、アイデアソンの感想、参加者同士のマッチング希望を収集した。一部の参加者からの希望もあり、参加者全員で連絡先を共有することとなった。

イベント終了後に参加者からは、「異なる立場のプレーヤ同士でざっくばらんな意見交換ができ、大変有意義な場であった」というフィードバックや「アイデアソンで知り合った参加者と継続的に意見交換をしていきたい」「自分が所属する組織でもアイデアソンを企画したい」という声も聞かれた。

上記で示したアイデアソンを起点とし、提示された課題の棚卸し・明確化と合意を得る場として「個別の議論の場」、得られるデータを協調して成果に繋げる場・利用者のリテラシー向上の場として「ハッカソン・データソン」、得られる成果の高度化や利用者募集の場として「成果発信の場」といった施策の実施・活性化に繋げ、それらの施策から得られた合意内容やポイントを下記の書面等に集約・反映する事が必要となる。

- ・ 研究計画書における出口戦略
- ・ データシェアリングポリシー
- ・ データマネジメントプラン
- ・ 各種契約書

なお、施策を実施する際及び各種書面に含まれているべきポイント・留意点について、3.に記載する。