

データインフラに関する調査研究

報告書

2025年3月

株式会社 日本アプライドリサーチ研究所

目次

第1章 調査の目的および背景	3
1. 調査の目的	3
2. 調査の背景	4
第2章 調査内容および方法	14
1. 既に顕在化しているデータプラットフォームやメタデータに関するニーズ の検討 (項目①)	14
2. データプラットフォームやメタデータが整備されることによる新たな活用方法 に関する検討 (項目②)	14
3. 将来的に A I などの技術が進歩することにより生まれる新たな活用方法 に関する検討 (項目③)	15
4. 学協会等とのネットワーキング (項目④)	17
5. アンケート調査について	17
第3章 調査結果	19
1. 既に顕在化しているデータプラットフォームやメタデータに関するニーズ の検討 (項目①)	19
2. データプラットフォームやメタデータが整備されることによる新たな活用方法 に関する検討 (項目②)	37
3. 将来的に A I などの技術が進歩することにより生まれる新たな活用方法 に関する検討 (項目③)	43
4. 学協会等とのネットワーキング (項目④)	49
おわりに	50

第1章 調査の目的および背景

1. 調査の目的

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（以下、「**SciREX** 事業」という）は、エビデンスに基づく政策形成（以下、「**EBPM**」という）の実現を目指して、文部科学省が2011年度から開始した事業である。**SciREX** 事業では**EBPM**の実現に向けて科学と政策の共進化の重要性を掲げ取り組みを進めてきた。さらに、**SciREX** 事業終了後も科学技術イノベーション政策（以下、「**STI** 政策」という）における**EBPM**の推進が図られるような体制や政策コミュニティ、人材育成・活用の在り方、データ基盤整備等についての総括や提案を行っていくことも期待されている。この**SciREX** 事業全体のプログラムについては、文部科学省とともに、政策研究大学院大学 科学技術イノベーション政策研究センター（以下、「**SciREX**センター」という）が運営を担っている。

令和5年度には、データインフラに関するフィージビリティとして、複数の**STI** 政策関係などのデータのメタデータの入手可能性、ならびにメタデータを結合することで新たな分析が学術にも政策的にも可能になるかの検討のフレームワーク構築を目指した調査等を実施した。

今回の調査では、令和5年度における取組の継続・発展として、令和5年度の調査にて整理した「今後の検討課題」や研究会における意見などを踏まえて調査研究を進め、**SciREX** 事業の残りの期間における活動や、事業終了後の構想に役立てることを目的として、調査を実施した。

2. 調査の背景

(1) 昨年度調査にて整理された事項

昨年度調査（「データインフラに関するフィージビリティ調査」、以下同様）においては、以下のことが整理された。

1) 既に顕在化しているデータプラットフォームやメタデータに関するニーズ

①データが多岐に亘るために生じる問題点

STI関係のデータについては、非常に多岐に亘る。以下の表はSTI政策に係る学校教育関係のデータを調べ、それを表にまとめたものである。

表1 STI政策に係る学校教育関係のデータ

分類		考えられる指標の例	データのソース
資源	教育支出	・教育支出の対GDP比（全体、公費負担及び私費負担別）	・OECD「図表でみる教育 20XX」(Education at a Glance 20XX)
		・一般政府総支出に占める公財政教育支出の割合	・OECD「図表でみる教育 20XX」(Education at a Glance 20XX)
取組	初等中等教育	中学校・高等学校におけるキャリア教育の実施状況 ・全体計画の策定状況 ・実施内容、活動に充てる日数 ・教員に対する研修の実施状況 ・他の学校や諸機関との連携状況 ・キャリアパスポートの作成状況	・国立教育政策研究所「キャリア教育・進路指導に関する総合的実態調査」(平成24～28年度) → 「キャリア教育に関する総合的研究」(令和元年度)
		中学校における職場体験の実施状況 ・職場体験を実施した中学校の割合 ・職場体験を5日以上実施した中学校の割合	・国立教育政策研究所「職場体験・インターンシップの実施状況調査」
		高等学校におけるインターンシップの実施状況 ・インターンシップを実施した高等学校の割合（全体、普通科・職業に関する学科別） ・インターンシップを体験した高校生の数 ・体験日数別実施率	・国立教育政策研究所「職場体験・インターンシップの実施状況調査」
		小中学校における主体的・対話的で深い学びの視点による学習指導の改善に関する状況 ・課題の解決に向けて、自分で考え、自分から取り組むことができているか	・国立教育政策研究所「全国学力・学習状況調査」

分類	考えられる指標の例	データのソース
	高校におけるアクティブ・ラーニングの実施状況 ・参加型授業の導入状況 ・参加型授業の効果 ・参加型授業の悩み	・立教大学と日本教育研究イノベーションセンター 「高等学校における参加型学習に関する実態調査」
	・データサイエンス・A I を学ぶ高校生の数	・「統合イノベーション戦略 2020」や、「A I 戦略 2019」に目標値が記載、データ入手の可否は不明
	・学校の ICT 環境整備の現状 ・教員の ICT 活用指導力	・文部科学省「学校における教育の情報化の実態等に関する調査」
	ICT の利用状況 ・学校における ICT 機器利状況 ・国語・数学・理科の授業における ICT 利用状況 ・学校外における平日・休日のインターネット利用時間別生徒の割合 ・家庭における ICT 機器の利用状況 ・学校外における学習のための ICT 利用状況 ・余暇のための ICT 利用状況	・OECD 「生徒の学習到達度調査 (PISA)」
	・部活動の状況 (中学校、高等学校)	・スポーツ庁「運動部活動等に関する実態調査」
	・部活動の状況 (中学校)	・国立教育政策研究所「全国学力・学習状況調査」
入試方法	・推薦入試、AO 入試、一般入試の状況	・文部科学省「国公立大学入学者選抜実施状況」
高等教育	インターンシップの実施状況 ・単位認定を行うインターンシップと単位認定を行わないインターンシップの学校数・割合 ・参加学生数・参加率・実施期間 ・報酬の有無 ・外国人留学生の参加者数 ・海外インターンシップ実施校の数・参加学生数	・文部科学省「大学等におけるインターンシップ実施状況について」
	大学におけるアクティブ・ラーニングの実施状況 ・能動的学修 (アクティブ・ラーニング) を取り入れた授業を実際に行っている大学の割合	・文部科学省「大学における教育内容等の改革状況について」
	大学におけるキャリア教育の実施状況 ・キャリア教育を教育課程外で実施している大学の割合と取組内容	・文部科学省「大学における教育内容等の改革状況について」
	大学における ICT を活用した教育の実施状況 ・情報通信技術 (ICT) を活用した教育を実施する大学の割合	・文部科学省「大学における教育内容等の改革状況について」
	・データサイエンス・A I を学ぶ大学・高専生の数	・「A I 戦略 2019」に目標値が記載、データ入手の可否は不明

分類		考えられる指標の例	データのソース
		・起業家養成プログラム受講者数	・文部科学省「次世代アントレプレナー育成事業 (EDGE-NEXT)」関連 HP (文部科学省、各大学 HP)
	リカレント教育	・社会人学生の数 (高等教育機関への 25 歳以上の入学者の割合)	・文部科学省「学校基本調査」
		・大学・専門学校等における社会人受講者数 ・大学公開講座の受講者数	・文部科学省「開かれた大学づくりに関する調査研究」
		・数理・データサイエンス・AI 関係のリカレント教育を受けた社会人の数	・「統合イノベーション戦略 2020」や、「AI 戦略 2019」に目標値が記載、データ入手の可否は不明
	・MOOC (大規模公開オンライン講座) の学習者数 ・MOOC への参加機関数 ・MOOC の講座数	・各事業者の HP から集計? (要調査)	
成果	科学技術への興味・関心の醸成	児童・生徒たちの数学や理科に対する意識 ・(各科目について) 楽しいと思う児童・生徒たちの割合 ・(各科目について) 得意だと思う児童・生徒たちの割合 ・(各科目について) 学習すると日常生活に役立つと思う児童・生徒たちの割合 ・(各科目について) 将来自分が望む仕事に就くために、良い成績を取る必要があると思う児童・生徒たちの割合	・IEA「国際数学・理科教育動向調査 (TIMSS)」
		小中学校の国語・算数・数学の学力・学習状況 ・(各科目について) 勉強は好きか ・(各科目について) 勉強は大切だと思うか ・(各科目について) 授業の内容はよくわかるか ・(各科目について) 学習したことは社会に出てから役に立つと思うか	・国立教育政策研究所「全国学力・学習状況調査」
		・30 歳時に科学関連の職業に就いていることを期待している生徒の割合 (全体、男女別)	・OECD「生徒の学習到達度調査 (PISA)」
		・理工系、農学、医学、薬学各学部の入学者数、大学院進学率	・文部科学省「学校基本調査」
	初等中等教育	・数学・理科の学習到達度	・IEA「国際数学・理科教育動向調査 (TIMSS)」
		・読解力 ・数学的リテラシー ・科学的リテラシー	・OECD「生徒の学習到達度調査 (PISA)」
	デジタル関連	・「デジタル技術スキル」の国際比較	・IMD 世界デジタル競争力ランキング

[データが多岐に亘るために生じる問題点]

上記のように、学校教育について分析しようとするだけでも、多数のデータが必要となり、かつソースがバラバラであるため、以下の問題が生じる。

- ・データが分散しているため、必要なデータを探すのが大変となる。
- ・その前段階として、どのようなデータを調べたらいいかを考えたり、どのようなデータが存在するのかを調べたりするのも大変である。

②先行調査で指摘された問題・課題等

S T I 研究や EBPM におけるデータ収集について、先行調査においては、以下の問題・課題等が指摘されている。¹

表2 先行調査で指摘されたS T I 研究やE B P Mにおけるデータ収集の問題・課題

	先行調査で指摘された問題・課題等
政策関係文書	<ul style="list-style-type: none"> ○情報の発見しやすさの問題 ・各省庁が独自に HP にてアップしているため、体系的に探すのが大変。 ・中央省庁側も他省や別の組織の動きは把握しづらい。 ・審議会等の資料や議事録も探すのが大変。
国がシンクタンク等へ委託して行う調査	<ul style="list-style-type: none"> ○中身の問題 ・政策検討や研究者から見たモニタリング（診断）に向けては、足りないデータも見られる。 ○継続性の問題 ・省庁によっては、ほとんど定点観測しておらず、良いデータも時系列的な変化を追跡できない。 ○情報共有の問題 ・各省庁の HP にアップされるものとされないものがある。 ・調査結果が埋もれてしまいやすく、存在が認知されにくい。 ・発注側の担当が変わると類似の調査を繰り返すことが多くある。 ・省庁間でも調査結果が共有されないため、無駄が多くなる。
国の統計や委託調査の個票の開示	<ul style="list-style-type: none"> ○入手の容易性 ・（個票が欲しい場合）、入手が困難な場合が多い。 ○利用の利便性 ・（個票を使用したい場合）制約が多く、利用も大変。 →申請から使えるまでに1年かかる場合もある。 →データを見直してきれいにしても、他の人が使えない。 *返却（削除）が義務付けられている。 →博士課程の学生も使うことができない。 など

¹文部科学省令和4年度委託調査「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」に資する政策科学データの共用プラットフォーム構築に係る調査」報告書より

③現状の問題・課題を解決するために必要と考えられる対応策

前ページで指摘された問題・課題への対応策としては、以下のことが考えら、統合的なプラットフォーム構築のニーズを示している。

表3 指摘された問題・課題に対する対応策

	先行調査で指摘された問題・課題等	必要となる対応策
政策関係文書	<ul style="list-style-type: none"> ○情報の発見しやすさの問題 ・各省庁が独自にHPにてアップしているため、体系的に探すのが大変。 ・中央省庁側も他省や別の組織の動きは把握しづらい。 ・審議会等の資料や議事録も探すのが大変。 	<ul style="list-style-type: none"> ・組織横断的な検索を可能とするデータプラットフォームやメタデータの整備
国がシンクタンク等へ委託して行う調査	<ul style="list-style-type: none"> ○中身の問題 ・政策検討や研究者から見たモニタリング（診断）に向けては、足りないデータも見られる。 ○継続性の問題 ・省庁によっては、ほとんど定点観測しておらず、良いデータも時系列的な変化を追跡できない。 ○情報共有の問題 ・各省庁のHPにアップされるものとされないものがある。 ・調査結果が埋もれてしまいやすく、存在が認知されにくい。 ・発注側の担当が変わると類似の調査を繰り返すことが多くある。 ・省庁間でも調査結果が共有されないため、無駄が多くなる。 	<ul style="list-style-type: none"> ・組織横断的な検索を可能とするデータプラットフォームやメタデータの整備 ・STI研究や政策形成に必要なデータに関する検討と、それらを定点観測する体制の整備
国の統計や委託調査の個票の開示	<ul style="list-style-type: none"> ○入手の容易性 ・（個票が欲しい場合）、入手が困難な場合が多い。 ○利用の利便性 ・（個票を使用したい場合）制約が多く、利用も大変。 →申請から使えるまでに1年かかる場合もある。 →データを見直してきれいにしても、他の人が使えない。 *返却（削除）が義務付けられている。 →博士課程の学生も使うことができない。 など 	<ul style="list-style-type: none"> ・研究者へのデータ提供の改善

2) データプラットフォームやメタデータが整備されることによる新たな活用方法

①データプラットフォームのイメージ

昨年度調査においては、構築すべきデータプラットフォームのイメージを以下の図の様にまとめている。

〔統合的なデータプラットフォームができた場合に得られるデータ〕 * ネット上で図書館のようなものができる

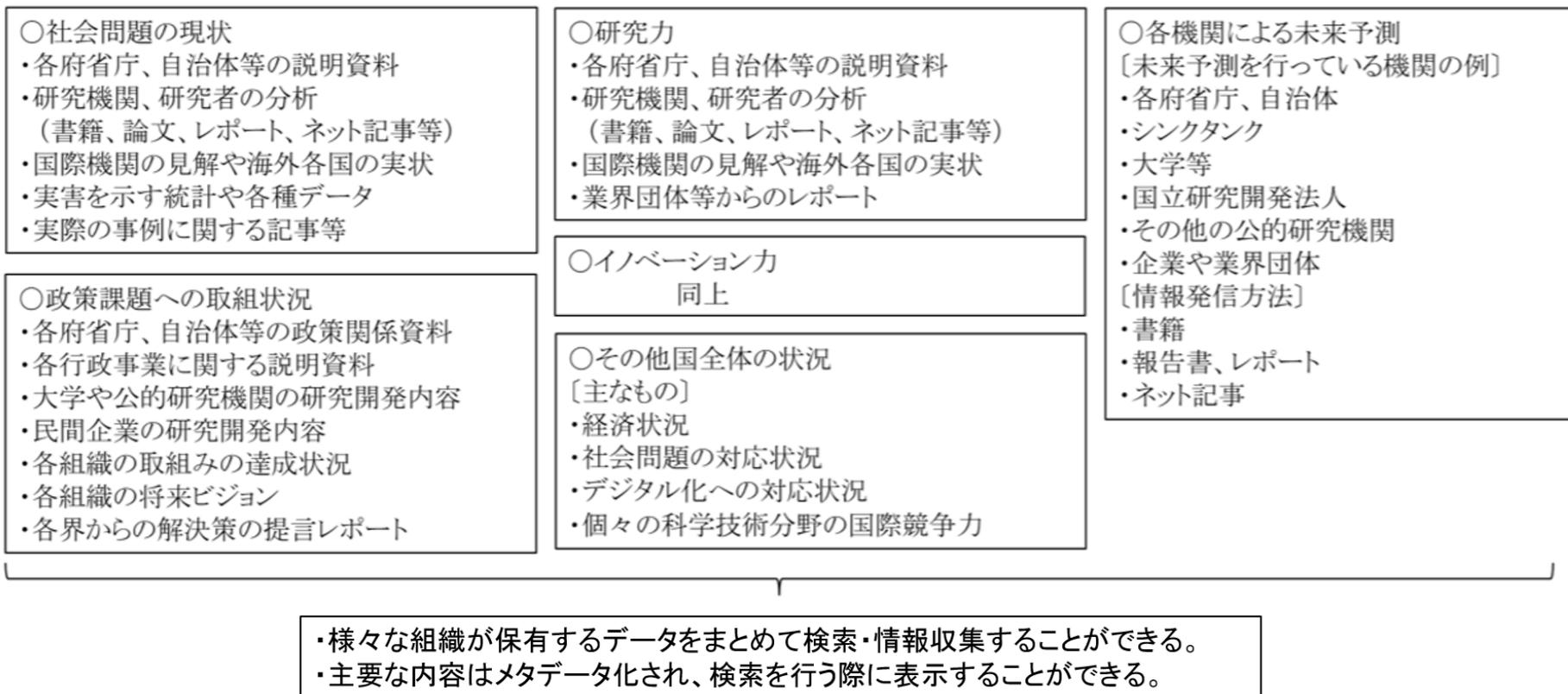


図1 データプラットフォームのイメージ

③データプラットフォームやメタデータが整備されることによる新たな活用方法

昨年度調査においては、データプラットフォームやメタデータが整備されることによる新たな活用方法について、その活効果を以下の図にまとめた。

〔考えられる効果のまとめ〕 * 昨年度調査より

※ネット上で図書館のようなものができる。

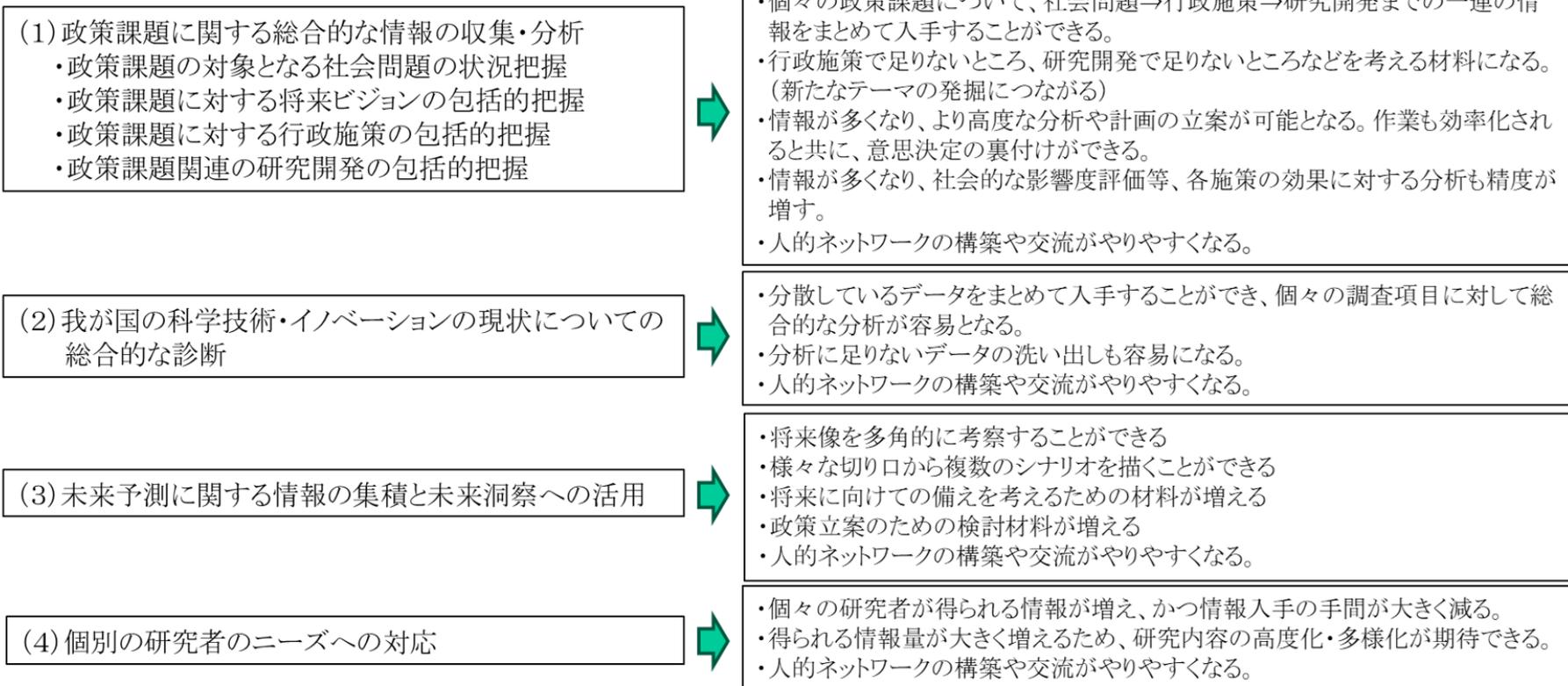


図2 データプラットフォーム構築により得られる効果

さらに研究会においては、具体的な活用方法として、以下のようなアイデアが出された。

表4 研究会で出されたデータプラットフォーム活用のアイデア

<p>研究会で出されたデータ活用のアイデア</p>	<ul style="list-style-type: none"> ・様々なデータを集められれば様々なことができる。例えば納税や疾病の状況を見ることができるとか。学術的には貴重な資源だし、EBPM としても有効。個人のデータを連結できれば、地域のレベルの変化と個人の変化を結び付けたりすることができる。 ・SNS の発言も状況に変化に応じた行動や考え方の変化も分析できる可能性もある。現状は個人のデータは公開していないが、様々なデータを個票と結びつけることができれば、人間に関する変化や行動の変化を分析できることが広がると考えられる。 ・食品ロスや薬価や器材等のサプライチェーンの安全保障などを調べているが、こうしたデータは非常に限られており、様々なデータに一元的にアクセスできれば、様々な分析が可能となる。 ・民間のデータは早いのでその活用は重要と考えている。モバイルデータや位置情報と組み合わせると、震災対策でも、誰がどこで孤立しているかなどがわかる。 ・ウェルビーイングに関するデータを集めているが、個人と社会・地域の変化とつなげられればいいと思う。 ・地域の中での総合計画のようなテキストのデータがあり、生成AI を使えばうまく整理できるのではないか。 ・数年前に京大の広井先生が地域の政策形成にシミュレーションを使って、それをデータがサポートするという事を言われていたが、今の技術だと政策形成にどういうことができるのかとか、あるいはウェルビーイング指標など、各所に散らばっているデータをつなげると、どういうものが見えてくるのか、を示してデータの整備の必要性を示すということが考えられる。 ・海外（イギリスなど）だと学生・院生個人の学習歴からキャリア、収入、納税、居住地等をデータで紐づけることが可能になっている。このように海外で動いていて日本でできていないことを議論する方法もあると思う。 ・未来予測情報に使うというのは現実的だと思う。外国の研究者が機械翻訳と併用すれば、外国の研究者との共同研究の道が開けると思う。インデックスデータがあるだけでも凄いと思う。進化の方向性としてこのような進め方が考えられる。 ・ユースケースを出すのも大事だが、データから生成されるシミュレーションを考えることもあると思う。 ・政策のようなマクロな部分と、ウェルビーイングのようなミクロな部分をどうつなげていくか、ということも考えられる。 ・日本の研究力やイノベーション力が落ちていると言われていた中で、大学の研究力に関するデータはいろいろとあるが、どうすれば研究パフォーマンスが上がるのか、研究サポート環境とか、研究力やイノベーション力を高めるための政策をどうするのか、についてのエビデンスをレガシーなデータではなく、違うデータで見ることができれば面白いと思う。 ・ちゃんとしたアカデミックデータのレジメがあるといいと思う。オープン。イノベーションや社会的イノベーションもきちんと理解されていないと思う。そういうプリミティブなところの知見もしっかりしていないと、行政官も世界で戦えない。そういうコンパクトな情報も必要だと思う。
---------------------------	--

3) データプラットフォームやメタデータの整備に向けて今後検討すべき課題

昨年度調査においては、今後検討すべき課題を以下のように整理した。

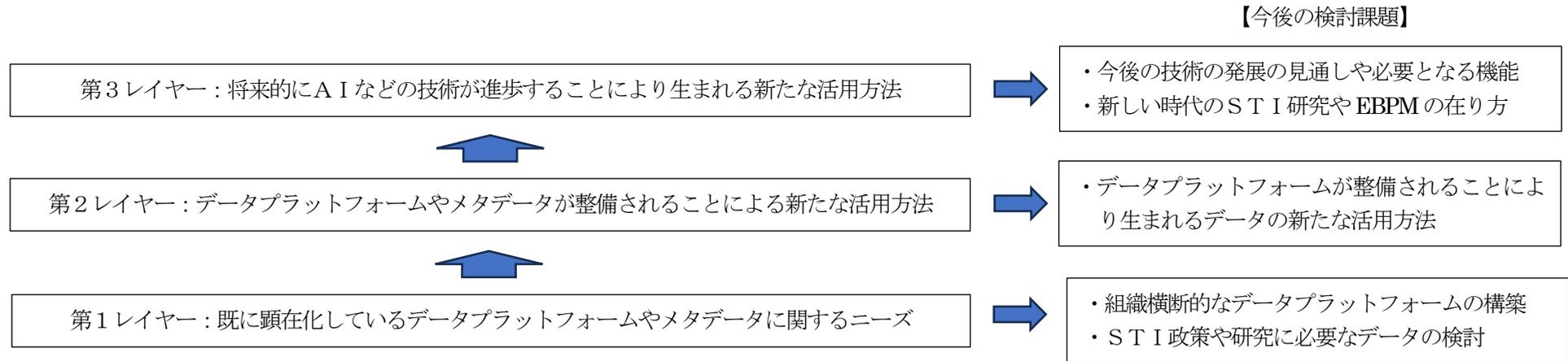


図3 レイヤーごとの検討課題

それぞれの検討課題において検討すべき内容を以下の表に整理した。

表5 今後検討すべき内容

	今後の検討課題	検討すべき内容
第3レイヤー	・今後の技術の発展の見通しや必要となる機能	・生成A Iなどの技術の今後の進歩や必要な技術について検討する。
	・新しい時代のS T I 研究やEBPM の在り方	・技術の進歩を踏まえ、新しい時代のS T I 研究やEBPM の在り方を探る。
第2レイヤー	・データプラットフォームが整備されることにより生まれるデータの新たな活用方法	<ul style="list-style-type: none"> ・第1レイヤーで検討するデータプラットフォームが整備されることを前提とし、そこから生まれるデータの新たな活用方法について検討する。 ・具体的なユースケースを多く出すこととする。
第1レイヤー	・組織横断的なデータプラットフォームの構築	<ul style="list-style-type: none"> ・多くの政策課題は府省庁横断的であり、大学等の研究機関や産業界も関係するため、組織横断的に広く情報を収集できるデータプラットフォームの構築が必要。 ・そのため、構築に向けての進め方を検討する。
	・S T I 政策や研究に必要なデータの検討	・現在、S T I 政策形成やS T I 研究に必要なデータが揃っているとは言い難く、必要となるデータについての検討を行う。

(2) 今年度の調査のポイント

昨年度調査で挙げた検討課題を踏まえ、今年度調査においては以下のことを検討することとなった。

表6 今年度の検討内容

	今後の検討課題	検討すべき内容
項目①：既に顕在化しているデータプラットフォームやメタデータに関するニーズの検討	<ul style="list-style-type: none"> 組織横断的なデータプラットフォームの構築（進め方） STI 政策や研究に必要なデータの検討 	<ul style="list-style-type: none"> 多くの政策課題は府省庁横断的であり、大学等の研究機関や産業界も関係するため、組織横断的に広く情報を収集できるデータプラットフォームの構築が必要。そのため、構築に向けての進め方を検討する。 現在、STI 政策形成や STI 研究に必要なデータが揃っているとは言い難く、必要となるデータについての検討を行う。
項目②：データプラットフォームやメタデータが整備されることによる新たな活用方法に関する検討	<ul style="list-style-type: none"> データプラットフォームが整備されることにより生まれるデータの新たな活用方法 	<ul style="list-style-type: none"> 項目①で検討するデータプラットフォームが整備されることを前提とし、そこから生まれるデータの新たな活用方法について検討する。 具体的なユースケースを多く出す。
項目③：将来的に AI などの技術が進歩することにより生まれる新たな活用方法に関する検討	<ul style="list-style-type: none"> 今後の技術の発展の見通しや必要となる機能 新しい時代の STI 研究や EBPM の在り方 	<ul style="list-style-type: none"> 生成 AI などの技術の今後の進歩や必要な技術について検討する。 技術の進歩を踏まえ、新しい時代の STI 研究や EBPM の在り方を探る。
項目④：学協会等とのネットワーキング	<ul style="list-style-type: none"> 項目①～③を進めるにあたっては、国内の関連する学協会等との連携・協力による取組が期待されるため、本学担当者とも協議しながら本件のテーマに関する主要アクターとのネットワーク形成を支援する。 	

第2章 調査内容および方法

1. 既に顕在化しているデータプラットフォームやメタデータに関するニーズの検討（項目①）

以下のように調査を実施した。

調査内容	調査方法
組織横断的なデータプラットフォームの構築（進め方）	・第3章の調査結果を踏まえて、今後の組織横断的なデータプラットフォームの構築の進め方を提案した。
S T I 政策や研究に必要なデータの検討	<ul style="list-style-type: none"> ・昨年度までの検討内容をベースとした。 ・研究・イノベーション学会の会員に対するアンケート調査を行い、ニーズを調査した。アンケートについては、「5. アンケート調査の実施」を参照。 ・さらに、以下の観点を追加した。 <ul style="list-style-type: none"> - 様々な社会情勢の変化や社会問題に関する事項 - I C Tに関する事項

2. データプラットフォームやメタデータが整備されることによる新たな活用方法に関する検討（項目②）

以下のように調査を実施した。

調査内容	調査方法
項目①で検討するデータプラットフォームが整備されることを前提とし、そこから生まれるデータの新たな活用方法についての検討	・研究・イノベーション学会の会員に対するアンケート調査を行い、アイデアを入手した。アンケートについては、「5. アンケート調査の実施」を参照。
具体的なユースケース	

3. 将来的に A I などの技術が進歩することにより生まれる新たな活用方法に関する検討（項目③）

以下のように調査を実施した。

調査内容	調査方法
今後の技術の発展の見通しや必要となる機能	<ul style="list-style-type: none"> ・将来的な A I 等の技術の発展の見通しや、それにより生まれる新たな活用方法については、I C T 系企業 2 社（3 名）、コンサル系企業 1 社、有識者 1 名へのヒアリング調査を実施した。
技術の進歩を踏まえた新しい時代の S T I 研究や EBPM の在り方	<ul style="list-style-type: none"> ・将来的な A I 等の技術の発展の見通しや、それにより生まれる新たな活用方法については、I C T 系企業 2 社（3 名）、コンサル系企業 1 社、有識者 1 名へのヒアリング調査を実施した。 ・研究・イノベーション学会の会員（研究会のメンバー）にもアンケート調査からも具体的事例の情報を得た。アンケートについては、「5. アンケート調査の実施」を参照。 ・有識者 4 名から成る検討会議を記載し、議論を行った。検討会議については、次ページを参照。

[検討会議について]

検討会議は以下のように進めた。

開催日時	2025年3月25日 19:00～21:00
参加者	<p>【委員】(五十音順)</p> <ul style="list-style-type: none"> ・大久保敏弘 慶應義塾大学経済学部教授 ・白川展之 新潟大学 人文社会科学系・工学部協創経営プログラム准教授 ・森健 株式会社野村総合研究所未来創発センター デジタル社会・経済研究室長 ・吉岡徹 一橋大学商学部准教授 <p>【オブザーバ】</p> <ul style="list-style-type: none"> ・根津純也 文部科学省科学技術・学術政策局 研究開発戦略課 政策科学推進室長/企画官 <p>【政策研究大学院大学】</p> <ul style="list-style-type: none"> ・林隆之 政策研究大学院大学 SciREX センター教授 ・野呂高樹 政策研究大学院大学 SciREX センター准教授 <p>【事務局】</p> <ul style="list-style-type: none"> ・大野幸雄 株式会社日本アプライドリサーチ研究所 主幹研究員 ・小沼良直 株式会社日本アプライドリサーチ研究所 特任主席研究員
議題	<ol style="list-style-type: none"> 1. 自己紹介 2. 趣旨や背景、参考情報の説明（事務局より） 3. 議論 <ol style="list-style-type: none"> (1) データ収集について (2) STI（科学技術イノベーション）研究やEBPM（エビデンスに基づく政策形成）において、AI等のICTに何を期待したいか？ <ul style="list-style-type: none"> ・今困っていること ・「こういうことができれば、さらに高度化できる」などの願望 ・最近起こりつつあること、構想されていること など (3) 今後のSTI研究やEBPMの在り方 <ul style="list-style-type: none"> ・将来に向けての展望や願望

4. 学協会等とのネットワーキング（項目④）

本調査においては、以下の場面において、研究・イノベーション学会とのネットワークを活用した。

アンケート調査	研究・イノベーション学会 課題研究型WG参加者 計143名に発送、うち回答者40名
検討会議のメンバー	研究・イノベーション学会の会員から2名、選出した。

5. アンケート調査について

（1）調査の目的

本政策研究大学院大学の SciREX センターでは、「データインフラに関する調査研究」を進めている。この調査では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（SciREX 事業）の一環として「調査、分析、研究に活用するデータを体系的かつ継続的に蓄積し、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」に資するデータ基盤を構築」することを目指しているが、その中で、以下のようなことを確認するためにアンケート調査を実施した。

- ・現在のSTI研究やEBPMにおけるデータの利活用における問題・課題
- ・将来的にAIなどのICT技術が進歩することにより生まれる新たな活用方法

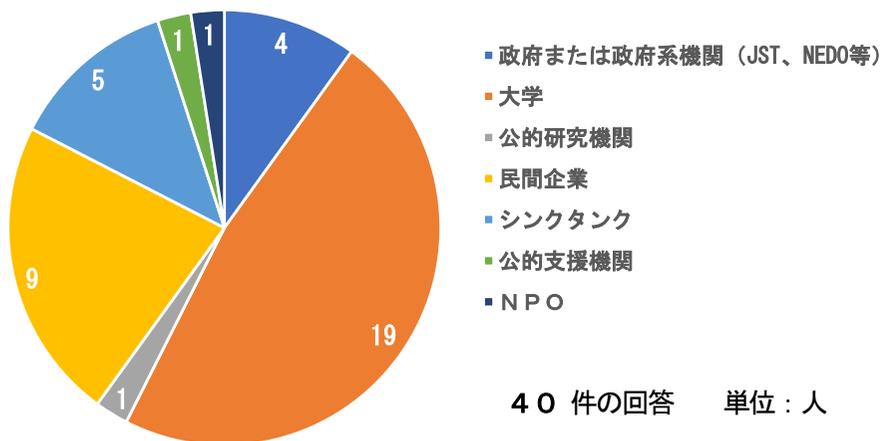
（2）調査対象および調査実施期間

アンケート調査対象および調査実施期間は以下のとおり。

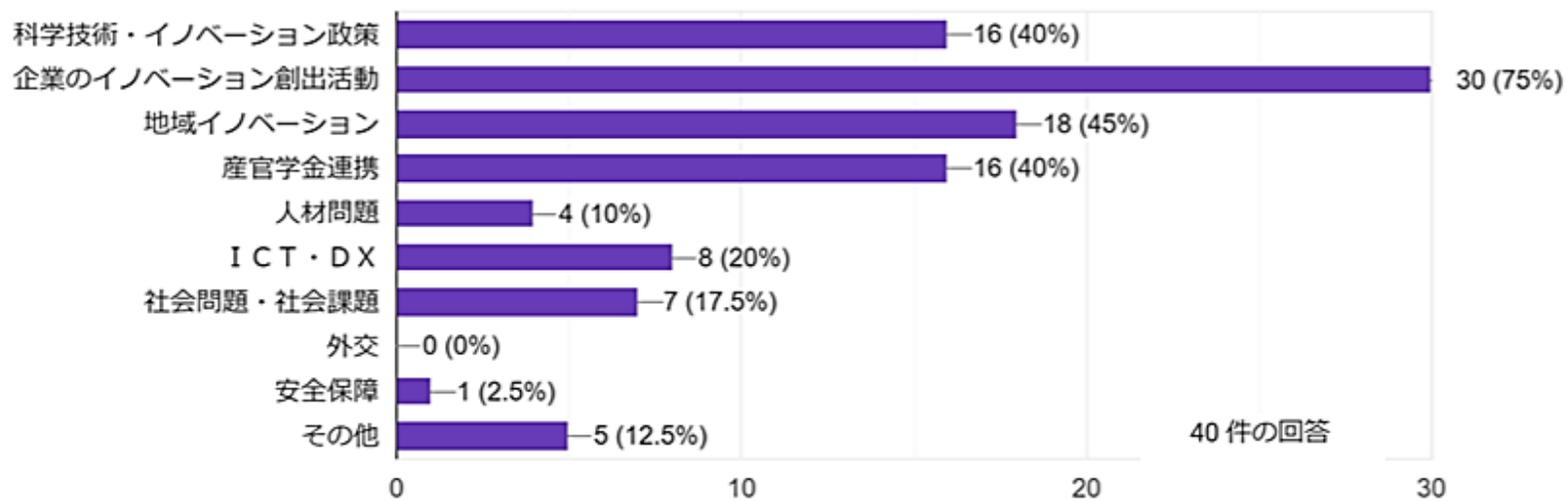
調査対象	研究・イノベーション学会 課題研究型WG参加者 計143名に発送、うち回答者40名（回答率28.0%）
調査実施期間	2024年12月25日～2025年1月25日

(3) 回答者の属性

○回答者所属組織の分類



○回答者の専門領域



第3章 調査結果

1. 既に顕在化しているデータプラットフォームやメタデータに関するニーズの検討（項目①）

（1）組織横断的なデータプラットフォームの構築（進め方）

多くの政策課題は府省庁横断的であり、大学等の研究機関や産業界も関係するため、組織横断的に広く情報を収集できるデータプラットフォームの構築が必要。そのため、構築に向けての進め方についての検討を行った。

本調査でまとめた新しい時代の S T I 研究や EBPM の在り方は以下のとおり。（第3章、P 4 8～4 9）

表7 新しい時代の S T I 研究や EBPM の在り方

		新しい時代の S T I 研究や EBPM の在り方
データの収集	データ収集の手間の削減	<ul style="list-style-type: none"> ・我が国には欧米や韓国に見られるような S T I 研究や E B P M に適する統合的なデータプラットフォームは存在しない、このため、前述のように研究者や行政官たちは、必要なデータを探すのに苦労している。また、統合的なデータプラットフォームができれば、利活用の幅が大きく広がることも、アンケート結果が示している。 ・この問題は先行調査や本調査を見ても切実であり、本来は我が国でも政府が主体となってデータプラットフォームを構築することが望ましい。 ・しかしながら、統合的なデータプラットフォームの構築には、膨大な費用と時間がかかることが想定され、実施主体や運用主体を決めるところから、多大な調整を必要とする。 ・こうした中で、生成 A I の進化に伴い、データの収集やメタデータ構築に向けては、生成 A I を核とするシステムの構築が将来的に有望な可能性として浮上してきた。 ・今後の検討の方向性としては、政府または関係機関にて生成 A I を核とするシステムの構築の実現可能性を継続して検討していくことが望ましいと考えられる。
	新たな情報や知見の発見	<ul style="list-style-type: none"> ・仮に統合的なデータプラットフォーム（生成 A I によるデータ収集機能でも同様の機能を有すると仮定する）が構築された場合、本調査のアンケート結果からも、新たな情報や知見の発見に加えて分析の精度向上や高度化が多く期待されている。 ・統合的なデータプラットフォームやそれに相当するシステムの構築は、単なるデータ収集の効率化にとどまらず、新たな価値につながる可能性を多く秘めている。
	ネットワークの構築	<ul style="list-style-type: none"> ・統合的なデータプラットフォームやそれに相当するシステムの構築は、今まで接点がなかった様々な組織の取組みを相互に知ることができるようになるため、新たなネットワーク作りや交流につながると考えられる。こうした新たなネットワーク作りや交流の推進につながるような仕組み作りは非常に重要だと考えられる。

		新しい時代の ST I 研究や EBPM の在り方
データの利用	分析の精度向上や高度化	<ul style="list-style-type: none"> ・統合的なデータプラットフォームやそれに相当するシステムの構築により、アンケート結果にもいくつかの例が紹介された様に、様々な分析の精度向上や高度化が期待できる。
	信頼性の向上	<ul style="list-style-type: none"> ・統合的なデータプラットフォームやそれに相当するシステムの構築により、様々なデータを入手することが可能となるため、多くのデータを比較することにより、分析の信頼性は増すと考えられる。 ・しかしながら、その一方で多くの偽情報や誤情報も入手してしまうことが考えられるため、データの信頼性の担保は大きな課題になると考えられ、偽情報や誤情報を判定する仕組みも同時に構築する必要があると考えられる。
	政策や戦略の立案	<ul style="list-style-type: none"> ・既に企業では、生成 AI を活用した戦略提案が実現されており、自治体においても政策形成や EBPM において AI を試行的に実施している例は見られる。 ・こうした動きは今後も広がると思われるが、その際に問題となるのが、AI・生成 AI の判断の妥当性であり、ベースとなるデータの信頼性が欠如していたり、あるいは必要なデータそのものがなかったりする場合には、AI・生成 AI 側の判断も怪しいものとなってしまふ。 ・このため、AI・生成 AI を活用する場合においても、人によるチェック機能は必要となり、妥当性を判断できる人材の育成も重要となる。
	AI エージェントの活用	<ul style="list-style-type: none"> ・多様な意見を吸収する目的での AI エージェントの活用も広がると思われる。人口の少ない地方自治体などで、元々多様な意見を吸収することが難しい場合や、民間企業において社員同士の考え方が似通ってしまい、社内で発想の多様性を確保することが難しいなどの場合には、AI エージェントが有効に機能する可能性は大いにある。 ・その一方で、人間にも個性があるため、AI エージェントが本当にその立場を代表しているか、ということも確認することも必要になると思われる。
	未来洞察・未来予測	<ul style="list-style-type: none"> ・未来洞察・未来予測においては、複数のシナリオを作成することが多くあり、その場合は洞察・予測の正確性よりも洞察や予測の多様性が必要となる。こうした複数のシナリオを設定する場合には、むしろ AI との親和性が高いという指摘が検討会議の中で出されている。 ・人間の知恵を超えた未来洞察・未来予測をどこまで AI・生成 AI が描くことができるのか、今後のさらなる進歩が望まれる。
全体を通して		<ul style="list-style-type: none"> ・AI・生成 AI は ST I 研究や EBPM の推進に向けて、様々な可能性を秘めていると考えられる。 ・しかしながら、同時に信頼性の担保が非常に重要となり、それを判断するためのチェック機能も必要となる。 ・こうしたことから、AI・生成 AI を使う人間の側にも高度な判断力が要求され、人間の側の能力開発も重要になると考えられる。

また、組織横断的なデータプラットフォームに必要な機能を以下の表にまとめた。(第3章、P44)

表8 今後のSTI研究やEBPMの進化に向けてAI等の技術に必要な機能

今後のSTI研究やEBPMの進化に向けてAI等の技術に必要な機能		
データプラットフォーム構築 やそれに伴うメタデータ作成 に向けてのデータ収集関係	データ収集機能	<ul style="list-style-type: none"> ・様々な公開情報から、生成AIがSTI研究やEBPMに必要なデータを検索して収集することができるようにする。 ・公式サイトのみならず、必要に応じてSNS等からも情報を入手できるようにする。 ・一つの生成AIが全てを処理する方法もあるが、政府の各府省庁が窓淵となる生成AIを設置し、生成AI同士が会話して情報を入手することも考えられる。
	メタデータ作成機能	<ul style="list-style-type: none"> ・データを収集する際に、同時にインデックスとなるようなメタデータも同時に作成できるようにする。
	データの信頼性チェック機能	<ul style="list-style-type: none"> ・データを収集する際に、偽情報や誤情報(数字の単位のミスなど)等を判別するなど、データの信頼性をチェックできる機能も必要となる。
データの利活用関係	問題・課題分析機能	<ul style="list-style-type: none"> ・様々なデータを用いて問題・課題の根本原因を分析できるようにする。 ・その際に、データ収集関係でも記載したように、データの信頼性をチェックできる機能も必要となる。
	シミュレーション機能	<ul style="list-style-type: none"> ・様々な問題・課題や解決策に対して、どのように変化していくかをシミュレーションできる機能を有することが望ましい。
	AIエージェント機能	<ul style="list-style-type: none"> ・問題・課題の解決に向けて多様な考え方を反映するために、AIエージェント機能を有することが望ましい。
	未来洞察・未来予測機能	<ul style="list-style-type: none"> ・未来洞察・未来予測は様々なシナリオが描けるが、上記のシミュレーション機能も使いながら、様々なシナリオを描く機能を有することが望ましい。
	ソリューション提案機能	<ul style="list-style-type: none"> ・上記の「問題・課題分析」、「シミュレーション」、「AIエージェント」、「未来洞察・未来予測」の結果も踏まえて、問題・課題に対応する政策や戦略等のソリューションを提案することができるようにする。
	説明資料作成機能	<ul style="list-style-type: none"> ・ソリューション提案の内容や、それに至るプロセスなどを生成AI等を用いて作成できる機能を有することが望ましい。

これらの調査結果を踏まえ、今後のデータプラットフォームの構築の進め方については、以下のような進め方を提案する。

①検討すべき内容と検討時期（全体）

データプラットフォームの構築に向けては、以下のように、データ・システムの両面からの検討が引き続き必要と考えられる。以下の表に進め方の案をまとめた。

表9 データプラットフォームの構築に向けての進め方

	検討すべき内容	パートナー	検討時期
データ側からのニーズ整理	<ul style="list-style-type: none"> データプラットフォームの必要性についてのコンセンサス作り 構築するシステムのイメージ作り 	<ul style="list-style-type: none"> データの中身を把握し、全体調整ができるシンクタンク 行政官（複数の省庁が参加することが望ましいが、最初は文部科学省のみとし、徐々に広げていくことが望ましい） ニーズを有する研究者数名 	令和7年度 <ul style="list-style-type: none"> 文部科学省のみ 令和6年度 <ul style="list-style-type: none"> 他府省も参加
システムの実現可能性の検討	<ul style="list-style-type: none"> A I ・生成A I を使ったシステム構築の実現可能性の検討 予算的に可能であれば、プロトタイプシステムの構築を行う。 	<ul style="list-style-type: none"> 要件定義ができるシンクタンク I C T系企業 A I ・生成A I を使ってシステム構築した経験を有するコンサル系企業 ニーズを有する研究者数名 	令和7年度 <ul style="list-style-type: none"> 机上検討 予算面の検討 令和8年度 <ul style="list-style-type: none"> プロトタイプシステムの構築

(2) S T I 政策や研究に必要なデータの検討

1) 昨年度調査で抽出した、S T I 政策や研究に必要なデータ

昨年度調査においては、以下のデータを報告書に記載した。

表 1 0 昨年度調査で抽出した、S T I 政策や研究に必要なデータ

データの利用目的	代表的な項目の例		
A 科学技術・イノベーションの状況の把握	研究力関係	人材の確保・育成	<ul style="list-style-type: none"> ○初等中等教育 <ul style="list-style-type: none"> ・科学技術への興味・関心 ・進路選択関係 ・学力関係 ○高等教育 <ul style="list-style-type: none"> ・進路選択、就職関係 ・産業界ニーズとのマッチング ○生涯教育 <ul style="list-style-type: none"> ・受講状況 ○教育全般 <ul style="list-style-type: none"> ・教育支出の状況 ・デジタル化への対応状況 など
		多様な人材の活躍の促進	<ul style="list-style-type: none"> ○若手研究者 <ul style="list-style-type: none"> ・若手研究者の登用状況 ・活躍促進に向けた環境整備の状況 ○女性研究者の活躍促進 <ul style="list-style-type: none"> ・女性研究者の登用状況 ・活躍促進に向けた環境整備の状況 ○外国人研究者の活躍促進 <ul style="list-style-type: none"> ・外国人研究者の登用状況 ・活躍促進に向けた環境整備の状況 ○人材の流動化、キャリアパスの多様性 <ul style="list-style-type: none"> ・研究者の流動性やキャリア関係

データの利用目的	代表的な項目の例		
		研究環境の整備	<ul style="list-style-type: none"> ○研究設備 <ul style="list-style-type: none"> ・最先端の設備の保有や共同利用の状況 ○研究資金 <ul style="list-style-type: none"> ・研究資金の額 ○研究領域 <ul style="list-style-type: none"> ・多様な研究領域へのチャレンジの状況 ○研究時間 <ul style="list-style-type: none"> ・研究に専念できる時間 ○国際活動の推進 <ul style="list-style-type: none"> ・国際共同研究や交流活動の状況
		生み出された成果	<ul style="list-style-type: none"> ○研究による成果 <ul style="list-style-type: none"> ・優れた論文の割合 ・論文数
イノベーション力関係		企業のイノベーション創出活動の状況	<ul style="list-style-type: none"> ○ビジョン・戦略 <ul style="list-style-type: none"> ・ビジョン・戦略の策定状況 ・事業構想の取組状況 ○人材・組織環境 <ul style="list-style-type: none"> ・人材の確保・育成状況 ・人材の多様性、流動化の状況 ・評価・処遇に関する状況 ・人材を活性化させる風土構築の状況 ○研究開発 <ul style="list-style-type: none"> ・研究開発投資の状況 ・外部連携・オープンイノベーションの状況 ○事業化・実用化 <ul style="list-style-type: none"> ・マーケティング ・新規事業の立ち上げ ・海外展開
		周辺環境の変化	<ul style="list-style-type: none"> ○競争環境の変化 <ul style="list-style-type: none"> ・国内外の競争環境の変化に関するデータ ・経済安全保障 ・競争環境の変化
		生み出された成果	○イノベーションの創出状況

データの利用目的	代表的な項目の例		
			<ul style="list-style-type: none"> ・プロダクト・イノベーションの創出状況 ・プロセス・イノベーションの創出状況 ○幸福度 ・幸福度に関する指標
	その他の国全体の状況	経済状況	<ul style="list-style-type: none"> ○国全体の経済状況 ・GDP ・所得 ・物価上昇率 ・失業率 ・人口の推移
		社会問題の状況	<ul style="list-style-type: none"> ○様々な社会問題の状況を示す情報 ・地球温暖化関連 ・少子高齢化関連 ・安全保障関連 ・食料問題 ・災害対策 ・インフラの劣化状況関連 ・犯罪対策 ・メンタルの問題 ・医療・介護関連 ・地域振興 ・貧困対策 ・資源の確保
		デジタル化への対応	<ul style="list-style-type: none"> ○デジタル化に向けたインフラの整備状況 ・インターネット接続数 ・端末の普及状況 ○デジタル化に向けた人材育成の状況 ・AI人材やデータサイエンティストの人数 ・学校教育におけるICT教育環境 ・リカレント教育の状況 ○国際比較 ・世界デジタル競争力ランキング
		個々の科学技術分野の国	○個々の科学技術分野の状況

データの利用目的	代表的な項目の例		
		国際競争力	<ul style="list-style-type: none"> ・日本が優位な分野、遅れている分野 ・最新の技術動向
B国などの政策の取組状況や効果の把握 (政策の例) <ul style="list-style-type: none"> ・科学技術・イノベーション基本計画 ・統合イノベーション戦略 ・未来投資戦略 ・AI戦略 ・各省庁のプログラム、プロジェクトなど 	国の政策の考え方に関する資料	<ul style="list-style-type: none"> ・政策文書 ・会議資料 ・議事録等 	
	戦略等	<ul style="list-style-type: none"> ・基本計画や各戦略の文書や検討資料 ・各計画や戦略の指標 ・各計画や戦略の実施状況や達成状況 	
	各省庁のプログラム、プロジェクト	<ul style="list-style-type: none"> ・各プログラム、プロジェクトの説明資料 ・各プログラム、プロジェクトの実施状況 ・各プログラム、プロジェクトの報告書 ・各プログラム、プロジェクトの評価結果 	

2) アンケート調査による研究者側のニーズの洗い出し

本調査においては、第2章で記載したように研究・イノベーション学会の会員に対するアンケート調査を実施し、研究者側のニーズを吸収した。以下の表は、そのアンケート結果である。

表11 昨年度調査で抽出した、STI政策や研究に必要なデータ

質問内容：ご自身の研究領域や事業領域と照らして、定点観測（毎年のように継続して行うデータ収集）すべきとお考えのデータをご記入ください。

*現在は国などが定点観測していないデータについて

		出された内容
政策関係	全般	<ul style="list-style-type: none"> 研究開発税制などの政策を実際に活用した企業や組織の名称・住所情報・法人番号。 日本含む主要国・地域の各種政策レポート動向、主要業界団体のレポート、ISO/IEC/ITUといった国際的な標準化機関の標準化対象に対応した各種の標準仕様データ、大学等の公的研究機関の知財状況（これらは一部で行われているが集計結果のみで元データが公開されていない場合が大部分） 日経と東大が行っている日次物価指数、リアルタイム生産データ、リアルタイム物流データなど、機動的な経済政策の立案に必要なデータ。秘密保持の確保の観点から秘密計算などの技術が必要となろう。 実際には非常に難しいと思いますが、政策やその実行についてのリスク、副作用などの監視結果のデータ。 日本政府、地方公共団体の研究開発関連の事項別表があれば、イノベーション政策分析に有効と思います。また民間企業におけるクラウド活用の実態、情報化投資のモジュールレベルでの詳細内容と推移があれば、日本産業におけるICT装備率の網羅的・経時的な発展度合いと問題点を明確化できると思います。 各種政策に紐づく組織単位などの実行状況を補足したリアルな定点観測データ
研究関係	全般	<ul style="list-style-type: none"> 日本および海外主要国（含む新興国）の科学・技術分野別、産業別の論文執筆者数の経年推移。 日本の研究力低下に関し、アカデミアにおける雇い止めや非正規雇用、頭脳流出がどの程度大学の生産性に影響を与えているのか把握し、アカデミアの雇用構造を最適化する必要がある。日本の研究者のキャリアを追跡することのできる任意ではないシステムがあると良いと思われる。 研究者の規範・考え方の変化 研究テーマの持続性観測（どれくらいの期間で開発成果を得ているか、もしくは諦めたか）
イノベーション関係	企業活動関係	<ul style="list-style-type: none"> 科学技術領域別に調査された企業間のM&Aデータ
	スタートアップ関係	<ul style="list-style-type: none"> 創業10年未満くらいのスタートアップにおける人材流動性（どのようなキャリアの人材がスタートアップに流入しているかが把握できるようなデータ） スタートアップ企業の状況（例えば、人員、研究開発投資、産学連携、特許取得状況等）をより網羅的に定点観測してほしい。（経済産業省の大学発ベンチャー企業のデータを拡充するイメージ）

		出された内容
		<ul style="list-style-type: none"> ベンチャー企業の活動（研究開発費、販売額、投資額、従業員数など） スタートアップの動向（起業分野、企業数、退出数、資金調達など）
	地域関係	<ul style="list-style-type: none"> （難しいかと思いますが）地域の企業や高度人材の出入り＊進出撤退や転出入 地域活動に関するデータ 地域イノベーション創出に係る取り組みの実態 地域に於ける環境（経済・人口・自然環境変化等）。電力生産設備別コスト比較
	様々な組織にまたがる内容	<ul style="list-style-type: none"> 国内における、研究者、研究開発関係従事者等といった科学技術・イノベーション活動に関わるさまざまな人材に関する、年齢階級別ほか各種属性別の人数 企業や大学及び研究機関におけるオープンイノベーションの実施件数と、そのオープンイノベーションの取組を通して具体的に新たな経済的価値等を創出した件数や売上規模等の定点観測を希望します。 社会実装における人数・資金・質の指標。VC・PE（プライベートエクイティ）の投資金額・分布など。 日本および海外主要国（含む新興国）の科学・技術分野別、産業別の①特許発明者数の経年推移。 特許出願技術動向調査の報告に使われるデータが、継続定点観測されたら良いと思います。 技術が企業や消費活動、生活活動、心理的活動、地域活動、社会活動をどのように変化させているのかを観測できると良いと思います。
その他	その他の利用	<ul style="list-style-type: none"> 個人の健康データとスポーツ・食事の関係、地域の特徴、これを集積・傾向分析し、健康寿命延伸の対策として活用する。 都市部の微小気候データ（分単位での定点観測） → 都市部特有の気候変動を把握するため、都市の気温、湿度、風速を微細なスケールでモニタリング。特に、降雨情報は、現在のレーダーでは波形が減衰することから、正確に観測できていない事例もあるため。 排出量削減データ（累計） ネット上の情報やソフトウェア（ゲームや音楽な度も含む）のアーカイブや利用状況も含む基本情報（国会図書館の納本制度にあたるもの）
	補足コメント	<ul style="list-style-type: none"> （課題や仮説があってから、それを検証するためにデータを探す、という形が多いため、新たに収集すべきデータは思いつきません。アンケートや売り上げなどの数値だけでなく、文書の分析にAIが有力になる中で、AIにのみアクセスできさせる機微情報などがあるのもいいかと思いました。また、マイナアプリを用い減税などと組み合わせた信頼性の高いデータ収集を始めて欲しいと思っています） 研究領域とは異なりますが、データの継続的な収集により研究が革新される分野の個人に紐づかないデータであれば定点観測して公開していくべきだと思います。例えば環境に関するデータなど。 定点観測は、調査対象者などに負担がかかる場合が多いので、よほど重要性の高いものに絞るべき。また、毎年のデータ収集にこだわらず、例えば5年に一度のデータ収集でも有意義なデータはあると思われる。

3) S T I 政策や研究に必要なデータの検討

前述の昨年度調査の項目や、本調査で実施したアンケート調査の結果に加え、以下の観点から机上検討を行い、必要なデータの検討を行った。
(検討にあたり、追加した観点)

- ・ 様々な社会情勢の変化や社会問題が研究開発やイノベーションに与える影響が大きな影響を与えることから、これらに関する内容を強化する。
- ・ I C Tの発展が社会全体やS T I 研究・E B P Mに及ぼす影響も大きいことから、I C Tに関する内容も増やす。

①データの利用目的

データの利用目大きく分けて以下の3つが考えられる。

表12 データの利用目的

データの利用目的	補足説明
A 政策や国民生活・経済活動に影響を及ぼす外部環境の情勢の把握と分析、将来予測	<ul style="list-style-type: none"> ・ 国際情勢の変化（紛争や国際関係の変化など） ・ 経済情勢の変化（国際競争、サプライチェーン、グローバルバリューチェーンの変化など） ・ I C Tなど技術の進化がもたらす変化による社会変革 ・ 様々な社会問題による影響 ・ パンデミック等の突発的脅威などによる影響 ・ 上記を含めた将来予測（社会像、リスク分析など）
B 科学技術・イノベーションの状況の把握および分析	<ul style="list-style-type: none"> ・ 国全体の研究力、イノベーション力、科学技術の成果、イノベーションの創出状況などを把握できるようにする。（健康診断の様なもの） ・ *民間企業の市場動向や地域イノベーションを含む ・ これらのデータは、国の科学技術やイノベーションの状況をモニタリングするためのもので、データの範囲は広範囲に亘る。 ・ 最近の状況とトレンドを把握する必要があることから、定点観測することが求められる。 ・ 海外各国との比較も必要となる。
C 国内外の政策や各組織（産業界、大学、研究機関、シンクタンク等）の取組状況や効果の把握	<ul style="list-style-type: none"> ・ 国や関連する組織は、我が国の様々な問題・課題を解決し、よりよい方向に導くための政策を展開しているが、それらの取組状況や効果を把握できるようにする。（治療や健康増進の取組や効果の見える化のようなもの） ・ これらのデータは政策の継続判断、方向性の再検討、類似の問題・課題が新たに発生する際の参考情報となる。

前述の A、B、C の関係を図に示すと以下のようになる。

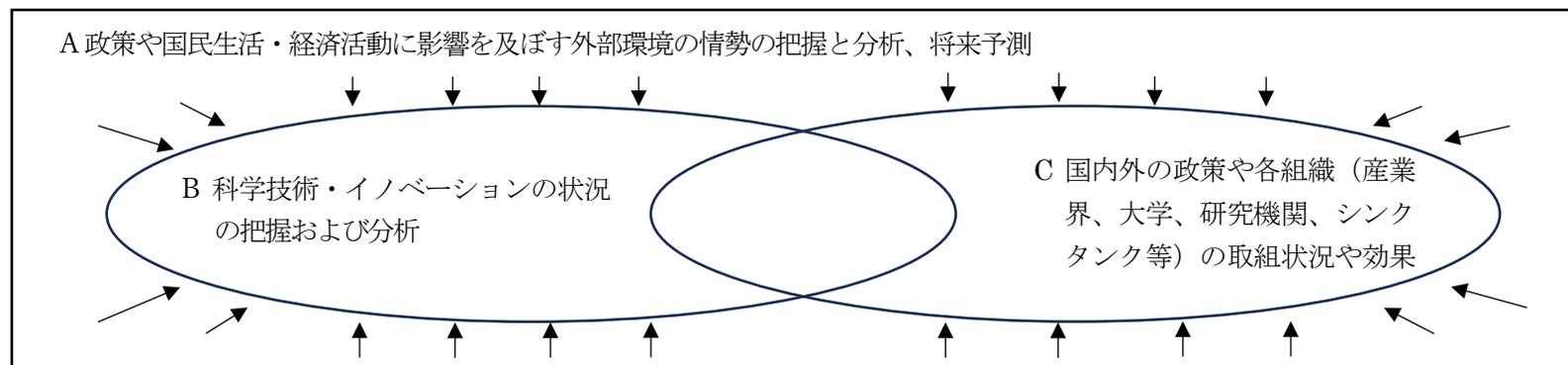


図4 A、B、Cのデータの関係性

②データの種類

データの種類としては、それぞれの目的ごとに以下のものが考えられる。

表13 考えられるデータの種類

データの利用目的	考えられるデータの種類
A 政策や国民生活・経済活動に影響を及ぼす外部環境の情勢の把握と分析、将来予測	<ul style="list-style-type: none"> ・国、シンクタンクなどの調査機関、個々の研究者が収集するデータが考えられる。 ①個々の研究者が集める情報や分析、将来予測 ②シンクタンクが受託調査や独自の調査などを通じて集める情報や分析、将来予測 ③国の統計や分析、将来予測
B 科学技術・イノベーションの状況の把握および分析	<ul style="list-style-type: none"> ・国、シンクタンクなどの調査機関、個々の研究者が収集するデータが考えられる。 ①個々の研究者が集めるデータ ②シンクタンクが受託調査などを通じて集めるデータ ③国の統計
C 国内外の政策や各組織（産業界、大学、研究機関、シンクタンク等）の取組状況や効果の把握	<ul style="list-style-type: none"> ・各国の政策の考え方に関する資料（政策文書、会議資料、議事録等）、 ・各国による各戦略のレビュー、指標データおよび各国・各省庁のプログラム、プロジェクトの報告書 ・各組織（産業界、大学、研究機関、シンクタンク等）の事業の取組みや効果等に関する情報 ・各組織（産業界、大学、研究機関、シンクタンク等）の研究開発の取組みや効果等に関する情報 ・各組織（産業界、大学、研究機関、シンクタンク等）間の交流や連携の取組みや効果等に関する情報

③データのレベル

データのレベルとしては、以下のようなものがある。

表14 データのレベル

データのレベル	主な例
文書レベル	<ul style="list-style-type: none"> ・国の政策の考え方に関する資料（政策文書、会議資料、議事録等） ・シンクタンクが受託調査などを通じて作成する報告書 ・シンクタンクや個々の研究者が作成する論文やレポートなど
グラフや表レベル	<ul style="list-style-type: none"> ・上記の文書に掲載されているグラフや表。
EXCELデータレベル	<ul style="list-style-type: none"> ・国の統計などでは、グラフや表の元となる EXCEL データが公開されている場合が多い。
個票レベルのデータ ＊国やシンクタンク、個々の研究者がアンケート調査を行う場合の回答の個票であり、個々の企業、大学、個人が回答した内容そのもの	<ul style="list-style-type: none"> ・個々の研究者が複数のデータをマージして新たな分析を行いたい場合などに、他者が行った調査の個票データを必要とする場合がある。 (例) <ul style="list-style-type: none"> －企業のイノベーションを研究する者が、国の調査で収集した個別企業の情報を求める －研究者のキャリアパスを研究する者が、研究対象の個人の情報を求める ・国が実施する調査の場合、大学等の研究者が申請すれば個票データを入手できる場合もあるが、情報保護の観点や調査実施時の回答者への約束などから、開示されない場合も多い。

④具体的なデータの例

以下に考えられる具体的なデータの例を挙げる。以下のものはあくまでも代表的な例であり、実際には多くの様々なデータから構成される。なお、前術の研究・イノベーション学会の会員へのアンケートからは、レベルを揃えた代表的なもののみ抽出した。

表15 STI研究やEBPMに必要な代表的なデータの例

データの利用目的	代表的なデータの例	
A 政策や国民生活・経済活動に影響を及ぼす外部環境の情勢の把握と分析、将来予測ICT	国際情勢の変化	<ul style="list-style-type: none"> ・紛争や国際関係の変化による影響
	経済安全保障	<ul style="list-style-type: none"> ・軍事的脅威 ・偽情報などの情報戦
	経済状況	<ul style="list-style-type: none"> ・GDP ・所得 ・物価上昇率 ・失業率 ・人口の推移 ・貿易額
	経済情勢の変化、経済安全保障	<ul style="list-style-type: none"> ・国際競争力関係の情報 ・サプライチェーン、グローバルバリューチェーンの変化
	個々の科学技術分野の国際競争力	<ul style="list-style-type: none"> ・個々の科学技術分野の状況 ・日本が優位な分野、遅れている分野
ICTを巡る状況	<ul style="list-style-type: none"> ○ ICT産業関係 <ul style="list-style-type: none"> ・ ICT関係の投資額 ・ ICT関係の生産額 ・ ICT関係産業の売上高 ・ ICTによる経済波及効果 ○ ICTの利用状況 <ul style="list-style-type: none"> ・ 情報通信機器の世帯保有率 ・ インターネット利用率 ・ IoT・AI等のシステム・サービスの導入・利用状況 ・ 行政手続きのオンライン利用率 	

データの利用目的	代表的なデータの例		
	様々な社会問題の状況	(例) ・地球温暖化関連 ・少子高齢化関連 ・食料問題 ・災害対策 ・インフラの劣化状況関連 ・犯罪対策 ・メンタルの問題 ・医療・介護関連 ・貧困対策 など	
	パンデミック等の突発的脅威などによる影響	・状況把握のための情報 ・対策および効果に関する情報 など	
	上記を含めた将来予測	・未来社会予測、未来洞察 ・将来に向けたリスク分析 など	
B 科学技術・イノベーションの状況の把握および分析	研究力関係	人材の確保・育成	○初等中等教育 ・科学技術への興味・関心 ・進路選択関係 ・学力関係 ○高等教育 ・進路選択、就職関係 ・産業界ニーズとのマッチング ○生涯教育 ・受講状況 ○教育全般 ・教育支出の状況 ・デジタル化への対応状況 など
		多様な人材の活躍の促進	○若手研究者 ・若手研究者の登用状況 ・活躍促進に向けた環境整備の状況 ○女性研究者の活躍促進 ・女性研究者の登用状況 ・活躍促進に向けた環境整備の状況 ○外国人研究者の活躍促進

データの利用目的	代表的なデータの例		
			<ul style="list-style-type: none"> ・外国人研究者の登用状況 ・活躍促進に向けた環境整備の状況 ○人材の流動化、キャリアパスの多様性 ・研究者の流動性やキャリア関係
	研究環境の整備		<ul style="list-style-type: none"> ○研究設備 ・最先端の設備の保有や共同利用の状況 ○研究資金 ・研究資金の額 ○研究領域 ・多様な研究領域へのチャレンジの状況 ○研究時間 ・研究に専念できる時間 ○国際活動の推進 ・国際共同研究や交流活動の状況
	研究テーマ		<ul style="list-style-type: none"> ・研究テーマの多様性 ・研究テーマの持続性
	生み出された成果		<ul style="list-style-type: none"> ○研究による成果 ・優れた論文の割合 ・論文数
イノベーション力関係	企業のイノベーション創出活動の状況		<ul style="list-style-type: none"> ○ビジョン・戦略 ・ビジョン・戦略の策定状況 ・事業構想の取組状況 ○人材・組織環境 ・人材の確保・育成状況 ・人材の多様性、流動化の状況 ・評価・処遇に関する状況 ・人材を活性化させる風土構築の状況 ○研究開発 ・研究開発投資の状況 ・外部連携・オープンイノベーションの状況 ○ICT関係 ・ICT関連の投資額 ・ICT関係の研究者・技術者の数

データの利用目的	代表的なデータの例		
			<ul style="list-style-type: none"> ・ A I 関係の究者・技術者の数 ○事業化・実用化 ・マーケティング ・新規事業の立ち上げ ・海外展開 ○M&A ・M&Aの数
		スタートアップ企業	<ul style="list-style-type: none"> ・開業率 ・資金調達状況 ・研究開発投資 ・販売額 ・ユニコーン企業の数 ・大学発ベンチャーの数 ・成果や出口戦略の達成状況
		地域イノベーション	<ul style="list-style-type: none"> ・地域活動に関するデータ ・地域イノベーション創出に係る取り組みの実態 ・地域における環境 ・経済・人口・自然環境変化等
		周辺環境の変化	<ul style="list-style-type: none"> ○競争環境の変化 ・国内外の競争環境の変化に関するデータ ・経済安全保障 ・競争環境の変化
		国際競争力ランキング	<ul style="list-style-type: none"> ○国際競争力のランキング ○デジタル化関係の国際競争力ランキング ○人材に関する国際競争力ランキング
		生み出された成果	<ul style="list-style-type: none"> ○イノベーションの創出状況 ・プロダクト・イノベーションの創出状況 ・プロセス・イノベーションの創出状況 ・ICTによる経済波及効果 ○幸福度 ・幸福度に関する指標 ○知的財産 ・特許件数、特許収入等

データの利用目的	代表的なデータの例	
<p>C 国内外の政策や各組織（産業界、大学、研究機関、シンクタンク等）の取組状況や効果の把握（政策の例）</p> <ul style="list-style-type: none"> • STI 政策、外交政策、安全保障政策、経済政策など（各組織の取組の例） • 事業（政策関連、各組織独自の取組み） • 研究開発（同上） • 組織間の交流や連携（同上） 	国の政策の考え方に関する資料	<ul style="list-style-type: none"> • 政策文書 • 会議資料 • 議事録等
	戦略等	<ul style="list-style-type: none"> • 基本計画や各戦略の文書や検討資料 • 各計画や戦略の指標 • 各計画や戦略の実施状況や達成状況
	各省庁のプログラム、プロジェクト	<ul style="list-style-type: none"> • 各プログラム、プロジェクトの説明資料 • 各プログラム、プロジェクトの実施状況 • 各プログラム、プロジェクトの報告書 • 各プログラム、プロジェクトの評価結果
	各組織の事業	<ul style="list-style-type: none"> • 各組織（産業界、大学、研究機関、シンクタンク等）の事業の取組みや効果等に関する情報
	各組織の研究開発	<ul style="list-style-type: none"> • 各組織（産業界、大学、研究機関、シンクタンク等）の研究開発の取組みや効果等に関する情報
各組織間の交流や連携	<ul style="list-style-type: none"> • 各組織（産業界、大学、研究機関、シンクタンク等）間の交流や連携の取組みや効果等に関する情報 	

2. データプラットフォームやメタデータが整備されることによる新たな活用方法に関する検討（項目②）

(1) データに関する研究者たちの問題意識

データの収集や利用に関する問題としては、先行調査の検討結果をP7に示したが、本調査においても前術の研究・イノベーション学会の会員へのアンケート調査にて、研究者の問題意識を調査した。（回答40名）

表16 データに関する問題意識（研究・イノベーション学会の会員へのアンケートより）

		意見等
データの収集	データに対する必要性の認識不足	<ul style="list-style-type: none"> ・社会全体としてまだ、（研究はまだしも）政策形成等においてその根拠となるデータや情報が必要不可欠であるという認識が十分に共有されておらず、（活用もそうですが、その前段として）整備・蓄積するということがなかなか進まないことに課題があるように思われる。
	データの所在のわかりにくさ	<ul style="list-style-type: none"> ・どこに、どういったデータが整備されているのか分かりにくい。 ・STI関係のデータは、全般的に量的にも質的にも不十分であるが、それ以上に、既にデータが存在しているにも関わらず、そのことが知られておらず活用が進んでいない場合があることが問題だと感じる。 ・統合的なデータのポータル（ワンストップで各種のデータのリストや利用情報が分かるサイトが存在しない）、継続性とデータの分類やデータ間の連携（データベース間の紐づけ、分類やそれらのための各種データコードの整備）、社会階層SSM調査のように継続的な実施主体が存在しないことが多く場当たりのデータ収集になる）、データの公開性（元データが利用可能な形で公開されていないことが多い） ・縦割りの整理のため、探すのに苦労する。 ・省庁が独自にHPにアップしているので方向性をもって探すのが困難、入手困難なデータがある。 ・集中的なデータプラットフォームが存在せず、各組織にてデータを分散的に保管し利用しにくい状況にあること。どこにどのようなデータがあるのかわかりにくいこと。保管データが書類をPDF化されたものが中心で、構造化されたデータとなっていない事例が多いこと。セキュリティ概念が曖昧でオープン化されていないこと。これらによって、総合値の獲得を困難にしていること。 ・多くの政策文書等について、位置づけ・整理のためのメタデータに基づく系統索引が使いやすく整備されていると助かります。 ・欲しいデータの検索が困難な場合がある。データ間での整合性など。 ・政府文書が体系化されておらず、相互の関連や最新バージョンの把握が難しい。 ・どのようなデータが揃っているか、分野ごとの一覧表があれば検索しやすいと思う。（既に用意されているかも知れません） ・データが散在しているため、検索・情報収集に時間がかかる。また、定点観測においては、項目におうじ

		意見等
		た横断的な検索ができない。
	データの継続性	<ul style="list-style-type: none"> 統計データの連続性が担保されていないケースが結構ある。 現場が必要とするデータ収集となっておらず、政策に必要な統計や一過性のデータとなっている。
	データの中身	<ul style="list-style-type: none"> 省庁や自治公設試験機関を横断するデータは少なく、機関毎に収集されている。さらに、分析精度の高いデータほどリアルタイムのデータではなく、対象年次が古いものも多く、課題整理や戦略立案、意思決定に使いづらく、自身で補足収集するには難がある。 市民や市民団体と呼ばれる人や組織の活動に関するデータが必要。 企業の取引関係が分かるようなデータがあると助かります（現在は購入しないといけない）。 政府のデータポータルが充実してきたので良い方向性だと思うが、一方で、地域の産業に関するデータの整備は十分ではない。 「地域」を扱うための適切な指標、データが不足している。 ①加工や分析されていない、統計データの入手が煩雑な場合があり、研究等に活用するのに付不便を感じます。②法律や調査によって、企業規模の区分けが異なる点に不便を感じます。例えば、中小企業基本法の中小企業の範囲と全国イノベーション調査の企業規模の区分けが異なる点等です。 政策文書の作成の過程（例：審議会や委員会）で出された意見や議論が具体的に記録に残っていない場合がある。
	その他	<ul style="list-style-type: none"> エビデンスをどのように構築するかが鍵。 国内において詳細データ整備を志向することも重要だとは思いますが、詳細ではなくても海外主要国と同じ基準で比較可能な基本的データをしっかり整備、利用しやすくする方がより重要ではないか、と個人的には考えている。 データの客観性、信頼性 統計データ（大学の研究開発費）に関し、人件費が教育向けと研究向けに分割されていないこと。 統計データ（大学の研究開発費）において、私学の支出額が大きいこと。 STI や EBPM の分野に関するデータは、直接その分野に関わる仕事をしていなければ、もしくは、資金や人材など資源豊富な大企業でなければ敷居が高い（アクセスしづらい）ものになっており、これは他分野のデータ活用にも一般的に同様に見られる問題だと思います。近年指摘されているように現代におけるクラウドデータは石油であり、データの収集力や保有量、アクセス性等における世の中の均衡を保つことは、経済的、政治的などあらゆる面での公平性を保つために不可欠だと思います。データを世の中の一部だけが使える状態は、結果的に一部の企業や機関への権力の集中をもたらし、均衡を崩してしまうのではと考えます。現在米国の政治界や経済界で見られる権力の集中は少し怖いものがあります。
データ利用	データの公開・共有	<ul style="list-style-type: none"> 情報の共有・利活用の整備が進んでいない。（特に自身の研究分野・医療・ヘルス케어での問題意識です。） 分析結果の共有（生データだけでなく、それを加工・分析した結果も可能な限り共有して欲しい） 公開場所、ファイル形式などがオーソライズされてない。紙ベースの報告書を単に電子化しただけのもの

		意見等
		も多く、活用が難しい。
	政府統計の利用	<ul style="list-style-type: none"> ・政府統計の利用期間が限定され、論文の改定等があり、利用期間を超えてしまうと、再利用が難しくなる。継続利用ができる仕組みがあると有益である。 ・政府等が収集した個票データを利用できないことが多い。類似のデータが複数の機関に分散して収集、保管されていることがあり、全体を把握することが難しい。 ・政府データを外部に貸与する際、もともと調査時点で「研究目的で使います」旨の告知をしていないにも関わらず貸与が進んでいるように見受けられます。回答企業としてはそのようなことを想定せずに部外秘データを政府に提出しているということも十分あり得る（行政が政策立案のために使用するから回答したが、研究者が使ったりするなら回答イヤだ、というような）ため、回答者との合意形成をどうするかについて、丁寧な議論が必要なのではと思います。 ・データの利用プロセスを分かりやすくしてほしい。 ・ローデータ（個票）を活用できれば、自分の研究でも学生の教育面でとても役に立つが、多くの場合でローデータの提供は不可か可能な場合でも手続きにとっても時間と手間がかかる（総務省統計局の情報等）。 ・企業におけるイノベーションに関連した分析を遂行する際、すでに行われた過去の調査の「個票データ」が使えるとありがたいが、なかなか難しいこと。（例）欧州のように、イノベーション調査等の個票など、活用させていただけるとありがたいです。 ・大学等に所属する研究者が政府統計の個票データを利用したい場合、各統計において開示される範囲や、開示を申請する際の担当部署等の連絡先などが、少なくとも researchmap に登録している研究者に対しては、あらかじめ一覧表のような形式で公開されていると利活用する上で利便性が高まります。 ・（個票が欲しい場合）、入手が困難な場合が多い。 ・（個票を使用したい場合）制約が多く、利用も大変 <ul style="list-style-type: none"> →申請から使えるまでに1年かかる場合もある →データを見直してきれいにしても、他の人が使えない →返却（削除）が義務付けられている →博士課程の学生も使うことができないなど
	政策との関係	<ul style="list-style-type: none"> ・現在省庁が実施しているのは EBPM ではない。各省庁にとって都合の良い政策目標が先に掲げられ、これに都合の良いデータを切り貼りしてストーリーを作り上げる PBEM（ポリシーベースドエビデンスメイキング）である。これが罷り通っているのは、政策形成部署にデータ分析の専門家を加えず、2-3年で担当部署を転々とする官僚が上長の指示に従ったストーリー作成をしているからである。この構造が変わらない限り日本で EBPM が実現することはないと考える。
	その他	<ul style="list-style-type: none"> ・分析に用いるのに適切なデータかどうか、迷うことがある。 ・データの利活用には多大な手間がかかるが、データ提供者もデータ利用希望者のどちらも、そのことをあまり認識しないことがしばしばあり、データ利活用が進まない要因になっていると感じる。

上記の内容からは、研究者たちは統合されたデータプラットフォームの必要性を強く感じていることが伺える。

(2) データプラットフォームやメタデータが整備されることによる新たな活用方法のアイデア

昨年度調査においては、データプラットフォームやメタデータが整備されることによる新たな活用方法として検討してした内容は、P10、11に記載したが、本調査においては、前術の研究・イノベーション学会の会員へのアンケート調査にて、ユースケースを含めてアイデアを入手した。

表17 データプラットフォームやメタデータが整備されることによる新たな活用方法 に関するアンケート結果 (具体的なユースケースを含む)

		新たな活用方法 (具体的なユースケースを含む)
データの収集	データ収集の手間の削減	<ul style="list-style-type: none"> 複数の組織が保有するデータを一括で利用申請できる窓口があると手間が省けてありがたい。 個票データの場合、名簿情報がないとデータの横断的な利用が難しい。ただし、名簿情報の利用はかなり制限されるので、もらえないことが多い。このため、一括申請を請け負う窓口があれば、この問題も解消できる。 資料の作成などに時間の短縮が可能 ソフトを使った網羅的なデータ収集が可能になり、研究コストが大幅に下げられる。
	新たな情報や知見の発見	<ul style="list-style-type: none"> データプラットフォームやメタデータを通じて、研究者等が、それまで知らなかったデータの存在を知ることができる可能性があり、そのような点で意義は大きいと思われる。
	その他	<ul style="list-style-type: none"> 複数の組織が保有するデータを一括で利用申請できる窓口があると手間が省けてありがたい。 個票データの場合、名簿情報がないとデータの横断的な利用が難しい。ただし、名簿情報の利用はかなり制限されるので、もらえないことが多い。このため、一括申請を請け負う窓口があれば、この問題も解消できる。 データやそのフォーマットの標準化が併せて実施されると有用であろう。使えないデータをプラットフォーム化してもその効果は限定的かもしれない。 現在大学や公的機関が持っている定義や形式もバラバラなデータを寄せ集める「プラットフォーム」を作成したところで、分析できるのは過去の限られた事象のみであると思われる。それよりも、今後個々のデータを同一定義で収集できるようなシステムを開発して、必要に応じて過去データをそのデータベースに遡及入力する方がはるかに利活用の幅は広がるのではないだろうか。 有効活用の可能性は高まると思われるが。インデックスの整理、情報密度のバラツキが予測される。
データの利用	分析の精度向上や高度化	<ul style="list-style-type: none"> 既に蓄積されているデータや情報に基づくという前提ではあるが、確認したり分析したりする対象に関して (部分ではなく) より全体に近い情報を把握することができる可能性が高まる。 イノベーション政策のバックボーンとなる各種エビデンスを定量的に示すことができる。 メタデータからバックキャストして何が分析できるかを考えることができる。 異なるデータの新たな組み合わせが可能になれば、これまでにない分析が進展する可能性がある。 一定の項目の組織別データの比較など様々なクロス集計等も可能となり、多様な視座による分析が可能になると考える。 研究課程において環境条件変化が起きても対応出来、問題点の解決にすぐに役立つ そのようなものを公に整備していくのであれば、組織や個人など様々なレベルでの研究がより活発化するのではないでしょう。また、それによっていろいろな研究の再現も容易になるのであれば、研究の価値の連続性が高まるのではない

		新たな活用方法（具体的なユースケースを含む）
		しょうか。例えばフィンランドでの最先端の研究を日本のフリーランスの研究者が活用し新しい技術を開発するなど。
	信頼性の向上	<ul style="list-style-type: none"> ・学術研究や報告書作成のためのデータソース、データへの信頼性の付与。 ・ファクトチェックに寄与。 ・ある課題を裏付けるデータの統一感が出てくる。
	戦略構築	<ul style="list-style-type: none"> ・情報の相互補完による横断的戦略の構築 ・国全体等での研究の取組み、進捗の俯瞰、政府の政策の体系的整備を行うことで、中長期の研究戦略・長期開発計画の立案に活用できます。
	未来予測	<ul style="list-style-type: none"> ・各機関における未来予測への活用
	具体的な活用例	<ul style="list-style-type: none"> ・医療・ヘルスケア分野での予防や早期発見・対応に活用出来る。 ・経済・経営学的な調査データ等と社会的なデータを包括的に調べ、地域の全体像を多角的に表現する。現状はそうしたデータがバラバラに存在するため、包括的にとらえることが難しい。 ・スタートアップ企業の情報とその企業が立地する地域や関連する大学等の機関の情報を連携して、スタートアップ企業が中心となるエコシステムを分析するということが考えられる。 ・類似の目的で類似分野の活動をしている組織の組織種類別の状態の共通点・相違点等の分析が可能 → たとえば官と民の違い、分野間や産業間の違い等が明確になれば其々の生産性向上やマッチング改善等に役立つと推測する。 ・地方自治体が個別に保有されていたデータプラットフォームを活用し、たとえば、ウェルビーイングに関する自治体の特徴を他府県自治体と比較することや推進すべき自治体戦略計画の政策立案に役立てている。 ・ファクトデータや論文・レポートのプラットフォームが整備されると、論文のレビューが効率的・効果的に実施できる。 ・様々な組織や地域単位、目的に応じた課題整理や戦略立案、意思決定等に活用できる。 ・地域経済のより多様な分析ができると思う。また現在、能登の創造的復興について検討しているが、ロードマッピング手法とデータが連携できれば（モデルを作ったうえで）、より緊張感のある未来構想も可能になっていくと思う ・地域の実態（人口問題等）を起点とした政策比較など。 ・（１）サプライチェーンの構成企業群におけるイノベーション創出の実態や連鎖（？）の様相、（２）イノベーションを引き起こす文脈（われわれの考え方、価値観、モノゴトの認知の仕方など）の変化と、イノベーションの関連性についての実証分析、等 ・各個票に対し、整合性のある多変量データが得られれば、国や個別企業についての総合的な技術力評価指標を生成することが可能になると考えます。 <p>【参考：昨年度調査の研究会で出されたアイデア（P 1 1の内容を再掲）】</p> <ul style="list-style-type: none"> ・様々なデータを集められれば様々なことができる。例えば納税や疾病の状況を見ることができるとか。学術的には貴重な資源だし、EBPM としても有効。個人のデータを連結できれば、地域のレベルの変化と個人の変化を結び付けたりすることができる。 ・SNS の発言も状況に変化に応じた行動や考え方の変化も分析できる可能性もある。現状は個人のデータは公開していないが、様々なデータを個票と結びつけることができれば、人間に関する変化や行動の変化を分析できることが広がると考えられる。

新たな活用方法（具体的なユースケースを含む）	
	<ul style="list-style-type: none"> ・食品ロスや薬価や器材等のサプライチェーンの安全保障などを調べているが、こうしたデータは非常に限られており、様々なデータに一元的にアクセスできれば、様々な分析が可能となる。 ・民間のデータは早いのでその活用は重要と考えている。モバイルデータや位置情報と組み合わせると、震災対策でも、誰がどこで孤立しているかなどがわかる。 ・ウェルビーイングに関するデータを集めているが、個人と社会・地域の変化とつなげられれば良いと思う。 ・地域の中での総合計画のようなテキストのデータがあり、生成AIを使えばうまく整理できるのではないかな。 ・数年前に京大の広井先生が地域の政策形成にシミュレーションを使って、それをデータがサポートするということを言われていたが、今の技術だと政策形成にどういうことができるかとか、あるいはウェルビーイング指標など、各所に散らばっているデータをつなげると、どういうものが見えてくるのか、を示してデータの整備の必要性を示すということが考えられる。 ・海外（イギリスなど）だと学生・院生個人の学習歴からキャリア、収入、納税、居住地等をデータで紐づけることが可能になっている。このように海外で動いていて日本でできていないことを議論する方法もあると思う。 ・未来予測情報に使うというのは現実的だと思う。外国の研究者が機械翻訳と併用すれば、外国の研究者との共同研究の道が開けると思う。インデックスデータがあるだけでも凄いと思う。進化の方向性としてこのような進め方が考えられる。 ・ユースケースを出すのも大事だが、データから生成されるシミュレーションを考えることもあると思う。 ・政策のようなマクロな部分と、ウェルビーイングのようなミクロな部分をどうつなげていくか、ということも考えられる。 ・日本の研究力やイノベーション力が落ちていると言われている中で、大学の研究力に関するデータはいろいろとあるが、どうすれば研究パフォーマンスが上がるのか、研究サポート環境とか、研究力やイノベーション力を高めるための政策をどうするのか、についてのエビデンスをレガシーなデータではなく、違うデータで見ることができれば面白いと思う。 ・ちゃんとしたアカデミックデータのレジメがあるといいと思う。オープン。イノベーションや社会的イノベーションもきちんと理解されていないと思う。そういうプリミティブなところの知見もしっかりしていないと、行政官も世界で戦えない。そういうコンパクトな情報も必要だと思う。
その他	<ul style="list-style-type: none"> ・データプラットフォームは論理的には好ましいことではありますが、それぞれの収集や整理の仕方は、収集者の個人的好み、時には利害によって決められています。その背景とそれを補正する方法論を含めた情報公開なしに単に集めて後はそれぞれの研究者が使うということになると正しい政策論とならない可能性もあります。その点に関する検討が必要です。 ・政策への活用、研究への活用が可能となる。また、こういったプラットフォームが使える前提で学生への教育も行える。 ・生成AIの活用による参照情報の効率的な検索や二次利用 ・データのマッシュアップが可能になる

3. 将来的に AI などの技術が進歩することにより生まれる新たな活用方法に関する検討（項目③）

将来的なAI等の技術の発展の見通しや、それにより生まれる新たな活用方法については、ICT系企業2社（3名）、コンサル系企業1社、有識者1名へのヒアリング調査を実施した。また、研究・イノベーション学会の会員（研究会のメンバー）にもアンケート調査からも具体的事例の情報を得ることができた。それらから得られた示唆をベースに検討を行った。

（1）今後の技術の発展の見通しや必要となる機能

1) 今後の技術の発展の見通し

ヒアリングやアンケートからは、今後の技術の発展の見通しとして以下の様な示唆を得ることができた。

表18 AI等の今後の技術の発展の見通し

AI等の今後の技術の発展の見通し（ヒアリングやアンケートから得られた示唆より）	
データプラットフォーム構築やそれに伴うメタデータ作成に向けてのデータ収集機能関係	<ul style="list-style-type: none"> 現時点でもICT企業では、ChatGPTのような汎用生成AIではなく、顧客向けに独自に生成AIを使ったシステムを組み上げて提供しており、本調査がターゲットとするシステムを独自に構築することは考えられる。 生成AIによるメタデータ作成は現時点でもかなりのことができるし、将来的には十分可能と考えられる。データだけを入れておけばメタデータの作成は自動的にAIがするという仕組みになると考えられる。各機関に生成AIという窓口があり、生成AI同士が会話して融通し合うという方法も考えられる。 将来的には生成AIが進化すれば、既存の検索エンジンも不要となり、生成AI自体が検索エンジンの役割を果たすことも考えられる。その場合は、メタデータ自体を作成しなくとも、いきなり生成AIを使って必要なデータを都度集めることも考えられる。
データの利活用関係	<ul style="list-style-type: none"> 現時点でもコンサル系企業においては、顧客向けの戦略立案の提案において生成AIを活用して、課題抽出やソリューション提案、プレゼン資料作成までを行っており、将来艇にはさらに高度な機能を有することは確実である。 政策形成においても、データ収集、それらを用いた分析、解決策の提案などの一連の業務を生成AIで行うことが普及し、内容的にも高度なアウトプットが出せるようになると考えられる。 分析におけるシミュレーションの機能も進化すると考えられる。 複数のシナリオが存在する未来洞察・未来予測についても、生成AIの活用が拡大すると考えられる。 戦略や政策形成において、多様な意見を反映してよりの確かな答を導くために、様々な立場を代弁する複数のAIエージェントを活用が進むことも考えられる。（既に企業等で使用されつつある）

2) 必要となる機能

今後のSTI研究やEBPMの進化に向けては、AI等の技術に以下の機能が必要になると考えられる。

表19 今後のSTI研究やEBPMの進化に向けてAI等の技術に必要となる機能

今後のSTI研究やEBPMの進化に向けてAI等の技術に必要となる機能		
データプラットフォーム構築 やそれに伴うメタデータ作成 に向けてのデータ収集関係	データ収集機能	<ul style="list-style-type: none"> ・様々な公開情報から、生成AIがSTI研究やEBPMに必要なデータを検索して収集することができるようにする。 ・公式サイトのみならず、必要に応じてSNS等からも情報を入手できるようにする。 ・一つの生成AIが全てを処理する方法もあるが、政府の各府省庁が窓淵となる生成AIを設置し、生成AI同士が会話して情報を入手することも考えられる。
	メタデータ作成機能	<ul style="list-style-type: none"> ・データを収集する際に、同時にインデックスとなるようなメタデータも同時に作成できるようにする。
	データの信頼性チェック機能	<ul style="list-style-type: none"> ・データを収集する際に、偽情報や誤情報（数字の単位ミスなど）等を判別するなど、データの信頼性をチェックできる機能も必要となる。
データの利活用関係	問題・課題分析機能	<ul style="list-style-type: none"> ・様々なデータを用いて問題・課題の根本原因を分析できるようにする。 ・その際に、データ収集関係でも記載したように、データの信頼性をチェックできる機能も必要となる。
	シミュレーション機能	<ul style="list-style-type: none"> ・様々な問題・課題や解決策に対して、どのように変化していくかをシミュレーションできる機能を有することが望ましい。
	AIエージェント機能	<ul style="list-style-type: none"> ・問題・課題の解決に向けて多様な考え方を反映するために、AIエージェント機能を有することが望ましい。
	未来洞察・未来予測機能	<ul style="list-style-type: none"> ・未来洞察・未来予測は様々なシナリオが描けるが、上記のシミュレーション機能も使いながら、様々なシナリオを描く機能を有することが望ましい。
	ソリューション提案機能	<ul style="list-style-type: none"> ・上記の「問題・課題分析」、「シミュレーション」、「AIエージェント」、「未来洞察・未来予測」の結果も踏まえて、問題・課題に対応する政策や戦略等のソリューションを提案することができるようにする。
	説明資料作成機能	<ul style="list-style-type: none"> ・ソリューション提案の内容や、それに至るプロセスなどを生成AI等を用いて作成できる機能を有することが望ましい。

(2) 技術の進歩を踏まえた新しい時代の S T I 研究や EBPM の在り方

1) A I を用いた E B P M

①現時点での取組事例

ヒアリングやアンケート結果からは、A I を用いた E B P M の事例として、以下のものが挙げられていた。なお、E B P M そのものではないが、プロセスに共通性がある企業の戦略立案での事例も含めて記載することとした。

表 2 0 A I を用いた E B P M および企業の戦略立案の事例

		A I を用いた E B P M および企業の戦略立案の事例
A I を用いた E B P M の事例	ヒアリングより	<ul style="list-style-type: none"> ・相模原市では、市長の答弁などを、生成 A I を使って作ることにトライしている。条例や規則などを学習させて適切な回答を作成するといった活用をしている。 ・政策の費用対効果の定量化にも取り組んでいる。 ・生成 A I も特化型が出てくると思っており、汎用的な答えを出す A I はもちろん、例えば法律に詳しい A I とか、自治体政策に詳しい A I、それらを窓口的に受ける A I など出てきている。 ・内閣府の場合は研究戦略で、国の研究投資としてはどの分野にどの程度するとイノベーションが進むかなどに取り組んでいる。 ・北海道石狩市・帯広市でアクションプランを作成した。2050 年を目標にして、石狩市の人口変化などによるモデルを作ってそれを動かしているいろいろなシナリオについて議論した。条件として脱炭素は達成できても人口が大きく減少するというシナリオもあるし、データセンターを誘致して風力発電を合わせて地域を盛り上げながら脱炭素を図るというシナリオもある。 ・震災後の復興計画の立案する際、どのような順に投資をすることで復興できるか、といったようなシナリオを考えている。 ・A I エージェントの使い方の例：103 万円の壁に対して各党を代表するエージェント、それに対して保守的な人、賛成派、反対派などが議論を交わすことによって政策の在り方を深めていくというものである。論点に対して各立場からの意見を出させてそこから議論を反駁していくということで、それぞれの立場から意見を述べて議論を深めていく。
	アンケートより	<ul style="list-style-type: none"> ・いくつかの国々・機関等では試行的に進んでいるようであるが、先験的に設定して分類を行うのではなく、対象とするデータに基づく分類分けを行う（たとえば、提案された研究プロジェクトの内容に即して、そのためにその採択評価のためのパネルを設定する）といったことがあるようである。

		A I を用いたE B P Mおよび企業の戦略立案の事例
		<ul style="list-style-type: none"> ・開発経済学や環境経済学の研究で、GPS データの意時点比較で変化があった（道路の舗装なや森林の開発など）箇所をA I が検出し、それを活用したデータを用いた研究が登場している。 ・洪水予測と避難計画（国土交通省）：ビッグデータと AI を活用して水位や雨量を予測。リスクの高い地域の住民に迅速な避難情報を提供。また、ダムの運用に関する農水省と国交省の連絡体制の改善等にも繋がっている。 ・中国の大気汚染モニタリングシステム（Blue Sky Program）：AI を使った大気汚染モニタリングを実施し、環境政策を強化。 ・行政文書を ChatGPT に作成させている例があると聞いております。
A I を用いた企業の戦略立案の事例	ヒアリングより	<ul style="list-style-type: none"> ・E Vの戦略について、東京一極集中なのか地方分散なのかなど、社会情勢に大きく影響されるため、社会モデルを作ってそれと一緒に企業の戦略を考えるということに取り組んでいる。強みとしては、人口規模ごとに自治体の標準モデルがあるので、自治体の特徴に合わせた戦略を考えることができる。 ・鉄道会社はディベロッパーでもあるので、沿線の人口構成の将来像を描く。 ・（顧客への戦略提案）セールスの戦略を決めていくというようなシーンにおいては、顧客の事業が BtoBとした場合、実際にお客様になる企業の財務諸表や有価証券報告書、中期経営計画などのデータを生成A I で分析して、顧客の持っている課題を抽出して、それに対して自社の製品やソリューションがどう当てはまるかマッチングして、リコメンドして、対顧客への提案書に仕立てていく。PowerPoint の資料も作成している。 ・（A I エージェントの活用）C F Oは自社の経営状況を見てコメントする。実際はどのオプションを取ったらいいいのかというシミュレーションは裏で予測のモデルが動いていて、それを随時A I エージェントが起動してその結果を評価し、社外取締役や競合他社の状況なども踏まえてコメントをする。いろいろな立場のA I エージェントを置いて、その中で経営の意思決定を高度化していく。
	アンケートより	<ul style="list-style-type: none"> ・政府等の政策形成でなく、民間企業の経営方針形成についてになりますが、弊社でも過去の経営データ等のAI で解析して経営課題の抽出、財務リスク分析等を行っております。競合の通信機器メーカーでも同様の取り組みをしております。

②今後の進化の方向性

今後の進化の方向性としては、以下のように進むことが考えられる。

表 2 1 AI による ST I 研究や EB PM の進化の見通し

今後の ST I 研究や EB PM の進化に向けて AI 等の技術に必要となる機能		
データプラットフォーム構築 やそれに伴うメタデータ作成 に向けてのデータ収集関係	データ収集機能	<ul style="list-style-type: none"> ・様々な公開情報から、生成 AI が ST I 研究や EB PM に必要なデータを検索して収集することができるようになる。 ・公式サイトのみならず、必要に応じて SNS 等からも情報を入手できるようになる。 ・一つの生成 AI が全てを処理する方法もあるが、政府の各府省庁が窓淵となる生成 AI を設置し、生成 AI 同士が会話して情報を入手することも考えられる。
	メタデータ作成機能	<ul style="list-style-type: none"> ・データを収集する際に、同時にインデックスとなるようなメタデータも同時に作成できるようになる。
	データの信頼性チェック機能	<ul style="list-style-type: none"> ・データを収集する際に、偽情報や誤情報（数字の単位のミスなど）等を判別するなど、データの信頼性をチェックできる機能が付加される。
データの利活用関係	問題・課題分析機能	<ul style="list-style-type: none"> ・様々なデータを用いて問題・課題の根本原因を分析できるようになる。 ・その際に、データ収集関係でも記載したように、データの信頼性をチェックできる機能も付加される。
	シミュレーション機能	<ul style="list-style-type: none"> ・様々な問題・課題や解決策に対して、どのように変化していくかをシミュレーションできるようになる。
	AI エージェント機能	<ul style="list-style-type: none"> ・問題・課題の解決に向けて多様な考え方を反映するために、AI エージェントが広く普及する。
	未来洞察・未来予測機能	<ul style="list-style-type: none"> ・未来洞察・未来予測は様々なシナリオが描けるが、上記のシミュレーション機能も使いながら、様々なシナリオを生成 AI が描くようになる。
	ソリューション提案機能	<ul style="list-style-type: none"> ・上記の「問題・課題分析」、「シミュレーション」、「AI エージェント」、「未来洞察・未来予測」の結果も踏まえて、問題・課題に対応する政策や戦略等のソリューションを提案する機能が広く普及する。
	説明資料作成機能	<ul style="list-style-type: none"> ・ソリューション提案の内容や、それに至るプロセスなどを、生成 AI 等を用いて作成する機能が広く普及する。

2) 技術の進歩を踏まえた新しい時代の S T I 研究や EBPM の在り方

A I 等の技術が進歩する中で、新しい時代の S T I 研究や EBPM の在り方を以下の表にまとめた。

表 2 2 新しい時代の S T I 研究や EBPM の在り方

		新しい時代の S T I 研究や EBPM の在り方
データの収集	データ収集の手間の削減	<ul style="list-style-type: none"> ・我が国には欧米や韓国に見られるような S T I 研究や E B P M に適する統合的なデータプラットフォームは存在しない、このため、前述のように研究者や行政官たちは、必要なデータを探すのに苦労している。また、統合的なデータプラットフォームができれば、利活用の幅が大きく広がることも、アンケート結果が示している。 ・この問題は先行調査や本調査を見ても切実であり、本来は我が国でも政府が主体となってデータプラットフォームを構築することが望ましい。 ・しかしながら、統合的なデータプラットフォームの構築には、膨大な費用と時間がかかることが想定され、実施主体や運用主体を決めるところから、多大な調整を必要とする。 ・こうした中で、生成 A I の進化に伴い、データの収集やメタデータ構築に向けては、生成 A I を核とするシステムの構築が将来的に有望な可能性として浮上してきた。 ・今後の検討の方向性としては、政府または関係機関にて生成 A I を核とするシステムの構築の実現可能性を継続して検討していくことが望ましいと考えられる。
	新たな情報や知見の発見	<ul style="list-style-type: none"> ・仮に統合的なデータプラットフォーム（生成 A I によるデータ収集機能でも同様の機能を有すると仮定する）が構築された場合、本調査のアンケート結果からも、新たな情報や知見の発見に加えて分析の精度向上や高度化が多く期待されている。 ・統合的なデータプラットフォームやそれに相当するシステムの構築は、単なるデータ収集の効率化にとどまらず、新たな価値につながる可能性を多く秘めている。
	ネットワークの構築	<ul style="list-style-type: none"> ・統合的なデータプラットフォームやそれに相当するシステムの構築は、今まで接点がなかった様々な組織の取組みを相互に知ることができるようになるため、新たなネットワーク作りや交流につながると考えられる。こうした新たなネットワーク作りや交流の推進につながるような仕組み作りは非常に重要だと考えられる。
データの利用	分析の精度向上や高度化	<ul style="list-style-type: none"> ・統合的なデータプラットフォームやそれに相当するシステムの構築により、アンケート結果にもいくつかの例が紹介された様に、様々な分析の精度向上や高度化が期待できる。
	信頼性の向上	<ul style="list-style-type: none"> ・統合的なデータプラットフォームやそれに相当するシステムの構築により、様々なデータを入手することが可能となるため、多くのデータを比較することにより、分析の信頼性は増すと考えられる。 ・しかしながら、その一方で多くの偽情報や誤情報も入手してしまうことが考えられるため、データの信頼性の担保は大きな課題になると考えられ、偽情報や誤情報を判定する仕組みも同時に構築する必要があると考えられる。
	政策や戦略の立案	<ul style="list-style-type: none"> ・既に企業では、生成 A I を活用した戦略提案が実現されており、自治体においても政策形成や E B P M において A I を試行的に実施している例は見られる。

		新しい時代の S T I 研究や EBPM の在り方
		<ul style="list-style-type: none"> こうした動きは今後も広がると思われるが、その際に問題となるのが、A I ・生成 A I の判断の妥当性であり、ベースとなるデータの信頼性が欠如していたり、あるいは必要なデータそのものがなかったりする場合には、A I ・生成 A I 側の判断も怪しいものになってしまう。 このため、A I ・生成 A I を活用する場合においても、人によるチェック機能は必要となり、妥当性を判断できる人材の育成も重要となる。
	A I エージェントの活用	<ul style="list-style-type: none"> 多様な意見を吸収する目的での A I エージェントの活用も広がると思われる。人口の少ない地方自治体などで、元々多様な意見を吸収することが難しい場合や、民間企業において社員同士の考え方が似通ってしまい、社内で発想の多様性を確保することが難しいなどの場合には、A I エージェントが有効に機能する可能性は大いにある。 その一方で、人間にも個性があるため、A I エージェントが本当にその立場を代表しているか、ということも確認することも必要になると思われる。
	未来洞察・未来予測	<ul style="list-style-type: none"> 未来洞察・未来予測においては、複数のシナリオを作成することが多くあり、その場合は洞察・予測の正確性よりも洞察や予測の多様性が必要となる。こうした複数のシナリオを設定する場合には、むしろ A I との親和性が高いという指摘が検討会議の中で出されている。 人間の知恵を超えた未来洞察・未来予測をどこまで A I ・生成 A I が描くことができるのか、今後のさらなる進歩が望まれる。
全体を通して		<ul style="list-style-type: none"> A I ・生成 A I は S T I 研究や EBPM の推進に向けて、様々な可能性を秘めていると考えられる。 しかしながら、同時に信頼性の担保が非常に重要となり、それを判断するためのチェック機能も必要となる。 こうしたことから、A I ・生成 A I を使う人間の側にも高度な判断力が要求され、人間の側の能力開発も重要になると考えられる。

4. 学協会等とのネットワーキング（項目④）

本調査においては、以下の場面において、研究・イノベーション学会とのネットワークを活用した。

表 2 3 本調査において活用した研究・イノベーション学会とのネットワーク

アンケート調査	研究・イノベーション学会 課題研究型WG参加者 計143名に発送、うち回答者40名
検討会議のメンバー	研究・イノベーション学会の会員から2名、選出した。

ジャンルの、本調査が対象としている S T I 研究や EBPM は、研究・イノベーション学会が扱っている領域そのものであり、研究・イノベーション学会との繋がり、今後も有益となる。

おわりに

我が国には欧米や韓国に見られるようなSTI研究やEBPMに適する統合的なデータプラットフォームは存在しない、このため、前述のように研究者や行政官たちは、必要なデータを探すのに苦労している。また、統合的なデータプラットフォームができれば、利活用の幅が大きく広がることも、アンケート結果が示している。

この問題は先行調査や本調査を見ても切実であり、本来は我が国でも政府が主体となってデータプラットフォームを構築することが望ましいと言えるが、統合的なデータプラットフォームの構築には、膨大な費用と時間がかかることが想定され、実施主体や運用主体を決めるところから、多大な調整を必要とする。

こうした中で、近年の生成AIの進化に伴い、データの収集やメタデータ構築に向けては、生成AIを核とするシステムの構築が将来的に有望な可能性として浮上してきた。今後の検討の方向性としては、政府または関係機関にて生成AIを核とするシステムの構築の実現可能性を継続して検討していくことが望ましいと考えられる。

統合的なデータプラットフォームやそれに相当するシステムの構築により、アンケート結果にもいくつかの例が紹介された様に、様々な分析の精度向上や高度化が期待でき、AI・生成AIはSTI研究やEBPMの推進に向けて、様々な可能性を秘めていると考えられる。

しかしながら、同時に信頼性の担保が非常に重要となり、それを判断するためのチェック機能も必要となる。こうしたことから、AI・生成AIを使う人間の側にも高度な判断力が要求され、人間の側の能力開発も重要であると考えられる。

本調査は、昨年度の調査を受けて、さらにAI等の技術の進歩を調べつつ、統合的なデータプラットフォームの利用の可能性をより広く探ったものであるが、実際のデータプラットフォームの構築に向けては、検討すべき項目が多く残されている。

本調査が、今後の統合的なデータプラットフォームの構築に向けて役立つものとなれば幸いである。

最後に、本調査においてヒアリング、アンケート、検討会議にご協力いただいた方々に心より感謝を申し上げたい。