

当財団では、2024年度より新たな中期計画を策定し、地域社会の発展に貢献する「中部圏の総合的・中立的シンクタンク」の実現に向け、研究力の強化、発信力の向上、ステークホルダーとの連携を柱として、事業を推進しています。

本レポートは、研究力強化の一環として2024年度に創設した「研究の場」、すなわち大学院生をインターンとして迎え入れて共同研究を行い、将来の研究者として指導・育成する事業として作成した成果報告書の概要となります。詳細は成果報告書としてホームページ^(※1)に掲載しておりますのでご一読ください。

事業評価票の評価システム作成と支援システムの必要機能についての検討

名古屋工業大学大学院工学研究科工学専攻 橋本 慧海

[ポイント]

本レポートでは、地方自治体における事業評価票作成支援システムの実現に向けて、大規模言語モデル（以下、「LLM」）を用いた事業評価票評価システムを試作し、その有用性と支援システムに必要な機能を検討した。実際の事業評価票に適用した結果、評価票の改善支援に一定の効果が確認できた。

①事業評価票に共通する課題（図2、3）

中部圏内の11自治体44件の事業評価票を対象に評価を行い、スコア傾向を分析した。その結果、論理的な整合性を示すスコアは自治体・事業間で大きな差がない一方、記述の具体性を示すスコアには差が生じることが分かった。

②実際のシステム利用例（図4）

本システムを実際の評価票に適用し、指摘された問題点と改善案を反映することで、評価票が改善されてスコアが向上していく過程を示した。

③支援システムに求められる機能（5.）

スコア傾向と自治体職員へのインタビューを踏まえ、事業のライフサイクルに応じた支援機能の切り替えや、作成支援を通じた職員の学習・育成を促す機能などが必要であることを示した。

1. 研究背景

行政においてエビデンスに基づく政策形成（以下、「EBPM」）が進められており、多くの自治体では事業評価票を作成している。しかし、杉谷（2024）や杉谷（2022）では、自治体の人手不足や担当職員ごとの能力の違いといったさまざま

な理由で、事業評価票の質にばらつきが生じることが指摘されている。

一定水準以上の事業評価票の作成を実現するために、作成支援システムの導入が望まれる。そこで、本レポートでは図1に示すような事業評価票作成支援システムの開発を最終目標とし、その一環として、LLMを用いた事業評価票の評価機能

(※1) <https://www.criser.jp/research/chiikikeiei.html>

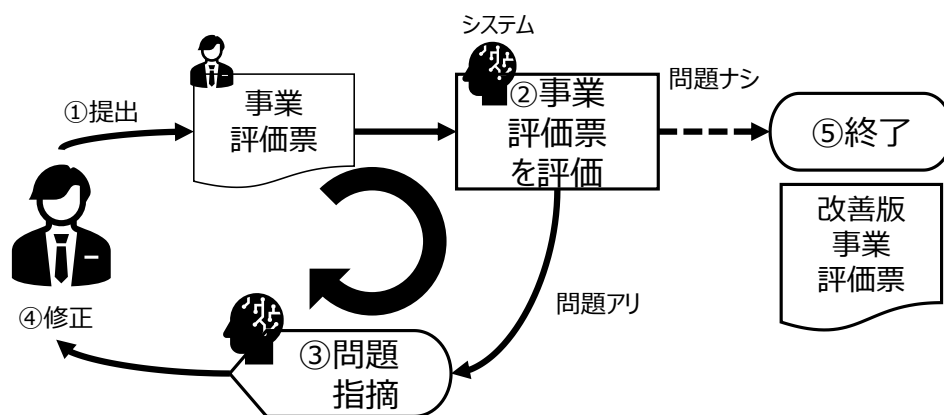


図1 事業評価票作成支援システムの概要図

①システムはユーザから作業中の事業評価票を受け取る、②ガイドラインに沿って評価、③問題がある場合はそれを指摘、④職員に改善を促す。

の開発を行う。

2. 評価機能の開発

適切な支援には作成された評価票の問題点を適切に指摘することが重要である。そのため、本レポートでは評価機能の開発を目指す。本レポートでの評価方法は「事業評価票が適切に記述されているか」を基本として評価する。

具体的には、内閣府が示す「行政評価のためのガイドライン」^(※2)をベースに筆者が作成した評価項目に対して、個々の事業評価票の記述がどの程度適合しているかを5段階（0～4）で評価する。

また、評価票の評価を単一の尺度で一元的に行うことは難しいため、記載の有無・具体性を測る記述スコアと、事業論理の整合性を測る論理スコアの二つの観点から評価を行った。評価のための基準を表1に示す。

3. 分析方法

本レポートでは、中部圏内の自治体11市から各市4事業を選定し、計44件の事業評価票を対象に評価機能をテストした。対象市の選定基準として、事業評価票を公開しており、かつ評価票全体の文字数が近い市を対象とした。また、事業ごとの特

表1 記述スコアと論理スコアの基準 0が最も悪い評価、4が最も良い評価である。

スコア	記述スコア	論理スコア
0	評価項目に対する直接記述がまったくない。	全体の論理が破綻しており、矛盾や不整合が至るところにみられる。
1	評価項目は書かれてはいるが抽象的で、意図や理念にとどまり、測定や対象が分からない。	事業論理の主要部分に明らかな不整合が残り、事業の流れがつかない。
2	評価項目の一部は具体だが、定義や対象の切り分けがあいまいである。	事業論理におおむね筋は通るが、部分的にあいまいさや軽度の不整合が混在している。
3	評価項目は具体的で、対象・基準や目標の情報がそろっている。	事業論理の流れは一貫しており、わずかな表記ゆれ程度の揺らぎしかない。
4	評価項目は完全に具体化され、測定方法や母集団・集計ルールまで明示されている。	事業論理は完全に首尾一貫しており、数値・時系列・役割・スコープがすべて整合している。

(※2) 自治体では、有効性や効率性といった観点から施策を評価しており、「事業評価」や「行政評価」と呼ばれる。また、その評価をまとめたものを「事業評価票」として公表している。

(※3) レビューシートを政策立案や予算要求という将来に向けての「意思決定」の一環として積極的に活用し、基礎的なEBPMを実践するための考え方・具体的な方法をまとめるための要綱。内閣官房行政改革推進本部事務局（2025）を参照。

徴を測るために、各市の事業内容が近い事業（災害対策、福祉、人づくり・街づくり、その他の四類型）を収集し、事業類型別にも集計した。

本稿では、二つの観点から分析を行った。一つ目は、評価スコアに基づく事業評価票の傾向を分析した。LLMによる評価結果をもとに、全体のスコア、事業類型別、自治体別の傾向の3点から分析する。二つ目は、評価スコアの傾向に基づいて支援機能を検討した。評価スコアの傾向から支援システムに求められる必要機能を明らかにする。

4. 評価結果の傾向と考察

図2・3に、11自治体の事業評価票に評価機能を適用し、得られたスコアを示す。

図2は全44事業のスコア平均値と4領域に分けた評価項目（事業目標、アウトカム、指標設定、ロジックモデル）の各スコア平均値を示している。全44事業の平均値は、記述スコア1.71、論理スコア1.92であり、論理スコアの方が高い傾向にあった。評価項目別で記述スコアをみると、「事業目標」のスコアは2.20と最も高い一方、「指標設定」

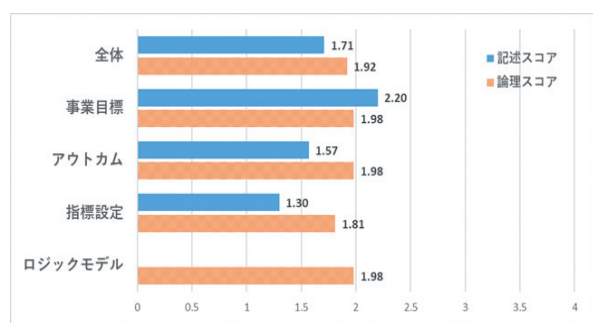


図2 全体のスコアと項目別のスコア

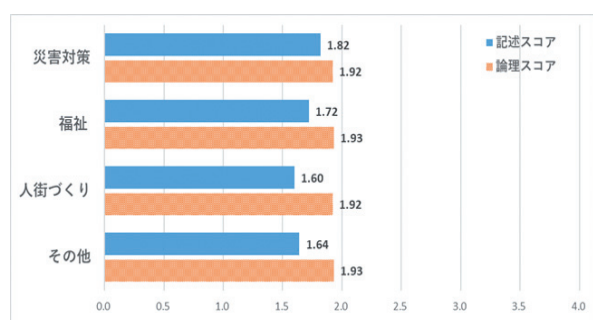


図3 事業類型別スコア

のスコアは1.30と最も低かった。

図3は事業類型別のスコア平均値を示している。事業類型別にみると、「災害対策」関連事業の記述スコアは1.82であるのに対して、「人づくり・街づくり」のスコアは1.60とやや低い値を示している。

以上から、論理スコアは多くの事業や項目において、一定以上のスコアを得ていることが分かった。しかし、記述スコアにおいては、「人づくり・街づくり」などの抽象的な概念を取り扱う事業ではスコアが低く出やすく、その中でも指標設定に関する評価項目のスコアが低い傾向にあった。

ここで、ある自治体の実際の事業評価票を用いて、1つの評価項目のスコアを1点分上げるために、筆者自身が作業を行った例を図4に示した。図4では、評価項目「票内のアウトカム指標を採用した理由・根拠が示されているか」の記述スコアを対象としている。

5. 事業評価票作成支援システムについて

分析した結果、多くの自治体で論理スコアが一定の水準だった一方、記述スコアは伸び悩みがみられた。これは、「アンケート調査の設計」や「アウトカムがみえにくい事業に対する指標設計」、「指標選定の根拠の明示」といった、高度なEBPM実践に関わる項目でスコアが低くなった自治体が多かったためである。今回の結果を踏まえると、支援システムにおいて重点的に支援すべき領域は、次の三点に整理できる。

- ①指標・評価設計の質を高める支援
- ②ロジックモデルの一貫性を高める支援
- ③「書きぶり」のばらつきをならす支援

6. まとめ

本レポートでは、EBPMの潮流を踏まえ、中部圏11自治体・全44事業評価票を対象として、LLMを用いた評価機能の試作と検証を行い、ス

評価項目：票内のアウトカム指標を採用した理由・根拠が明示されているか

事業の成果や効果を示す指標名(説明)	令和4年度	令和5年度	令和6年度	令和7年度	最終目標値	達成率(%)
① スポーツ推進委員 単位 目標 90 実績 82.1 83.5 81.7 - 最終目標年度 令和10年度	90	90	90	90	90	90.8
式・説明 定例研修会における出席率。研修会出席率＝延べ出席者数／(委員数×年10回開催)	82.1	83.5	81.7	-	-	-
各種大会・つどいの 単位 目標 6,000.00 実績 4,140 4,601 4,747 最終目標年度 令和10年度	6,000.00	6,000.00	6,000.00	6,000.00	6,000.00	79.1
式・説明 参加者数	4,140	4,601	4,747	-	-	-
式・説明 年度目標は前年度実績をもとに設定	-	-	-	-	-	-
アスリート先生派遣 単位 目標 30 実績 30 30 33 最終目標年度 令和10年度	30	30	30	33	35	114.3
式・説明 事業実施回数	27	30	40	-	-	-
式・説明 年度目標は前年度実績をもとに設定	-	-	-	-	-	-

システム出力

記述スコア：2

指摘：＜略＞「なぜその指標を選んだか」「目標値をどのように決めたか」の説明は限定的で、採用理由の体系的な根拠や＜略＞が示されていない。したがって、指標の採用根拠の明示は不十分であり、改善により説明の補強が可能である。

改善提案：指標設定の理由を文書で明記する（根拠＝前年度実績）、目標設定方法を数式で明示する（例：前年×1.05）、指標の測定頻度とデータ源を明記する（例：年次集計）

改善1回目

事業の成果や効果を示す指標名(説明)	令和4年度	令和5年度	令和6年度	令和7年度	最終目標値	達成率(%)
① スポーツ推進委員 単位 目標 90 実績 82.1 83.5 81.7 - 最終目標年度 令和10年度	90	90	90	90	90	90.8
式・説明 定例研修会における出席率。研修会出席率＝延べ出席者数／(委員数×年10回開催)	82.1	83.5	81.7	-	-	-
各種大会・つどいの 単位 目標 6,000.00 実績 4,140 4,601 4,747 最終目標年度 令和10年度	6,000.00	6,000.00	6,000.00	6,000.00	6,000.00	79.1
式・説明 参加者数が増えることで、市民の豊かなスポーツライフに繋がる。年度目標はこれまでの実績をもとに設定。	4,140	4,601	4,747	-	-	-
アスリート先生派遣 単位 目標 30 実績 30 30 33 最終目標年度 令和10年度	30	30	30	33	35	114.3
式・説明 事業実施回数	27	30	40	-	-	-
式・説明 アスリートとの接点を増やすことは、市民のスポーツへの関心向上に繋がる。年度目標はこれまでの実績をもとに設定。	-	-	-	-	-	-

システム出力

記述スコア：2

指摘：＜略＞。「参加者数が増えることで豊かなスポーツライフに繋がる」「興味を持つようになるため」といった文言は因果関係の想定を述べるにとどまり、期待する具体的な成果＜略＞は不十分である。＜略＞

改善提案：各指標と期待する成果の因果を定量化する（例：参加者増→健康指標5%改善）、指標の測定頻度とデータ源を記載する（例：年1回の集計を施設集計で実施）、目標設定の根拠を示す（例：過去3年の実績平均を基準に年次目標を設定）

改善2回目

事業の成果や効果を示す指標名(説明)	令和4年度	令和5年度	令和6年度	令和7年度	最終目標値	達成率(%)
① スポーツ推進委員 単位 目標 90 実績 82.1 83.5 81.7 - 最終目標年度 令和10年度	90	90	90	90	90	90.8
式・説明 【採用理由】スポーツ推進委員の研修受講は、指導・運営能力の向上を通じて事業の質を高める。【算定】研修会出席率＝延べ出席者数／(委員数×年10回開催)。【目標】直近実績から90%を目標とする	82.1	83.5	81.7	-	-	-
各種大会・つどいの 単位 目標 6,000.00 実績 4,140 4,601 4,747 最終目標年度 令和10年度	6,000.00	6,000.00	6,000.00	6,000.00	6,000.00	79.1
式・説明 【採用理由】大会・つどい等への参加が増えることは、市民のスポーツ実施の裾野拡大を示すため。【算定】各事業の参加者名簿に基づき、年度の延べ参加者数を集計。【目標】直近3年実績を基準に、改善目標として6,000人を設定する	4,140	4,601	4,747	-	-	-
アスリート先生派遣 単位 目標 30 実績 30 30 33 最終目標年度 令和10年度	30	30	30	33	35	114.3
式・説明 【採用理由】アスリートとの接点を増やすことは、市民（特に子ども等）のスポーツへの関心・参加意欲の向上に繋がるため。【算定】今年度の実施回数(回)。【目標】前年実績と前年目標の平均を用いて設定する。	27	30	40	-	-	-

記述スコア：3

図4 実際の評価票を用いて評価を行い、改善する様子（筆者が作業）

上部：原票、指標②、③についての指摘が中心。中央：原票への改善提案を基に手を加えた票（1度目）、依然スコアは2である。下部：何度かの改善を通して、最終的なスコアが3の票。下部票ではアウトカム設計への詳細な記述を増やした。しかし、このような記述を行うには、共通の票構造から変更する必要があり、数語を追加する程度では、十分にスコアアップが望めない可能性も示唆されている。

コアの結果分析を行った。

事業類型別・自治体別・評価項目別の分析を通じて、多くの自治体で事業目標や指標の基本的な整理・記述は一定程度達成されている一方で、アンケートを用いた効果測定的设计や、アウトカムがみえにくい事業における指標の工夫・根拠の明示といった、高度なEBPM実践に関わる部分が

共通の弱点となっていることが明らかになった。

また、論理スコアは項目別で大きな差はなかったものの、記述スコアには項目別で差がみられることが分かった。これらの結果を踏まえ、事業評価票作成支援システムが備えるべき機能の方向性も整理した。

参考文献

杉谷和哉（2022）『政策にエビデンスは必要なのか：EBPMと政治のあいだ』ミネルヴァ書房

杉谷和哉（2024）『日本の政策はなぜ機能しないのか？ EBPMの導入と課題』光文社

内閣官房行政改革推進本部事務局（2025）「行政事業レビューシート作成ガイドブック」2025年3月31日

https://www.gyokaku.go.jp/review/img/R07sakusei-guidebook_ver1-2.pdf