

《令和7年度》

○ 少年院における矯正教育等の効果検証及び見直し（法務省）

大規模言語モデル（LLM）の活用等によって、少年院における矯正教育等の効果を分析するとともに、矯正教育等の効果把握に当たり、アンケートや他の手法も用いて検証し、矯正教育等の効果を多角的に把握

○ 全国学力・学習状況調査結果の長期的分析等（文部科学省）

全国学力・学習状況調査結果という大規模なデータに着目し、データを正確に扱うために求められるデータクレンジングのノウハウを整理し、その結果を踏まえて経年的にデータを分析するとともに、特徴的な傾向を示す学校・市町村等を把握

○ 宮城県仙台市における健康増進事業の効果検証

仙台市が実施する健康増進事業（歩数等の運動記録、食事記録、健康意識のアンケート等）について、同市における健康行動や健康課題等を把握するとともに、健康増進事業による医療費・介護給付費の抑制効果額の推計を算出

《令和8年度実施（予定）案件》

○一日合同行政相談所における広報・普及啓発の効果検証に関する調査研究

※内容検討中

○地域の特性を踏まえた市民サービスの推進に関する調査研究

※内容検討中

- 再犯防止のためのエビデンスに基づく知見は蓄積途上
- 既存のアンケートや効果検証の手法は、限られた時点の情報であり、変化を経時的に十分に反映できず、効果検証に一定の制約。

- AIの活用を含めた矯正教育の効果把握・検証手法の高度化
(1. 統計的比較に代わる分析、2. 時系列変化の把握、3. 新しい知見等に基づくアンケートの改善)
- 分析結果から法令の見直し等に資する検討材料の提供。

1. 統計的比較に代わる分析

【分析課題】

取組対象が全員/取組対象が特性ある集団に限定されるなどの理由により、比較する集団を分けることが困難で、統計的比較ができない。
(面接記録などテキストデータは存在)

【分析の手法と結果】

- ・AI (LLM (大規模言語化モデル)) を用いて、再入院者の面接記録に含まれる自由記述から、再犯に関連する要因や要因同士の関連表現を抽出・図示。
 - ・再犯に着目した出現頻度の高い要因と要因間の関係(経路)を整理。
- ➔ **自由記述中の頻出要素や「再犯」を終点とした関係図の始点となる要因を特定できた。**

【分析から得られた示唆】

要因間の関係を可視化し、性別等の属性ごとの要因関係図・経路の差異を明確化。**矯正教育等の内容を検討する手がかりを得た。**

<留意点>

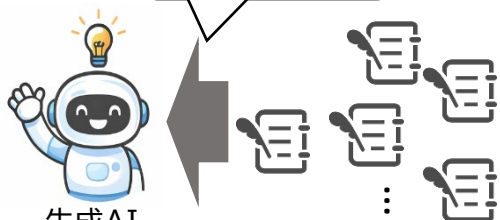
- ・少年本人の認識に基づく分析であり、統計的な因果推論とは性質が異なる。
- ・自由記述に表れる内容は少年本人の認識する要因にすぎず、そこから導き出される要因は本質的な再犯関係要因とは断定できない。

【分析イメージ図】

- ① LLMに、テキストを読み込ませ、「再犯」に至る要因を分析

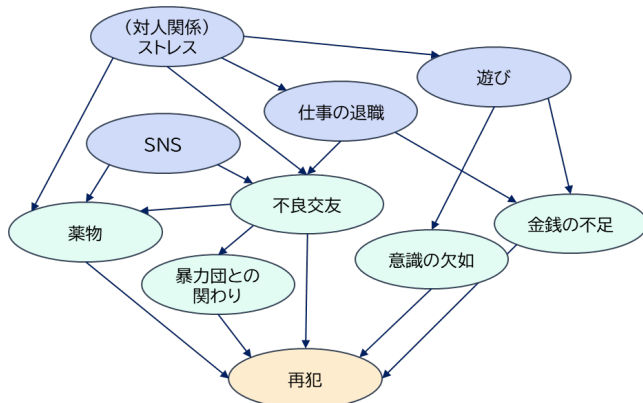
再入院時面接記録196件
出院時アンケート自由記述1,019件

AIに読み込ませ、分析

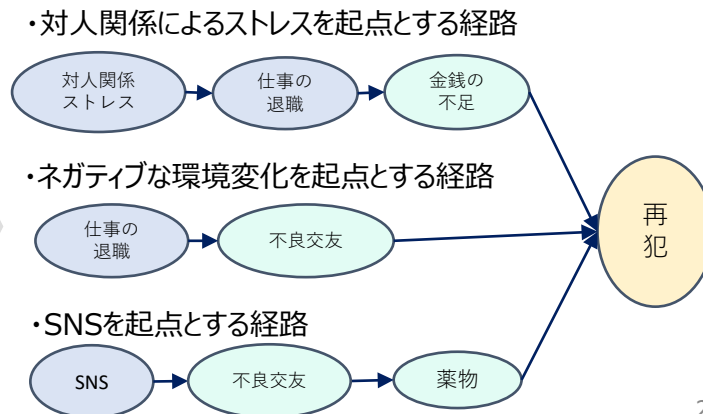


テキストデータ

- ② LLMに、終点として「再犯」に至る要因を抽出させ、図示化



- ③ 出現頻度の高い要素及び経路を導出



2. 時系列変化の把握

【分析課題】

アンケートで限られた時点の情報は把握できているが、少年や取組の時系列変化の把握が困難

【分析の手法と結果】

日記のテキストデータをAIを活用して整理・分析

- ①心理的指標を類型化し、定義
- ②専門家による心情変化の判断をAIに学習させ、AIで判定
- ③AIの判定結果を基に、時系列での心情の変化をグラフ化
- ➡ 時系列で良好な変化のあったケースを整理できた
- ➡ 良好な変化を示したケースでは、入院初期から、職員のコメントに対して少年から日記中で応答がみられるという特徴があった。

【分析から得られた示唆】

- ・テキストデータを活用し、心情変化の推移を類型化し、**心情変化の背景や要因分析の可能性**を得た
- ・職員との日記を介した対話が少年の成長と関連する可能性を踏まえ、再犯防止に**有効な矯正教育の検討材料とする。**

<留意点>

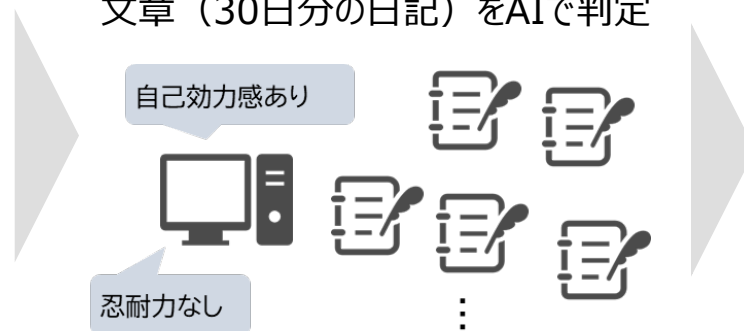
少数のサンプル（9名・238人日）を対象とした探索的分析である点に留意が必要。

【分析イメージ図】

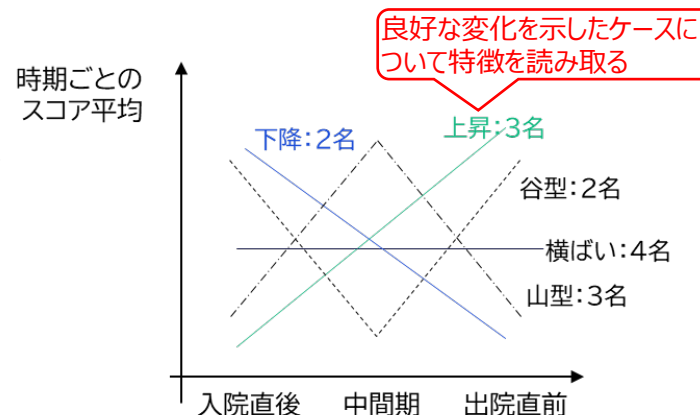
①心情の変化の変数を定義

変数一覧	
対人関係	忍耐力
自己効力感	変化を受け入れる準備
反省	大人への信頼感
主体性	援助要請の姿勢
目標を持つ	—

②専門家の判定をAIに学習させ、多数の文章（30日分の日記）をAIで判定



③時系列での心情の変化をグラフ化



3. 新しい知見等に基づくアンケートの改善

現在実施しているアンケートについても、特定の回答に偏るといった天井効果の緩和など改善の余地があり、以下の点について、改善の可能性を検証した。

- ・既存の出院時アンケートに犯罪心理学等の知見を用いた設問を追加して回答の天井効果の解消
- ・職員への信頼の程度や少年同士の関係性等を尋ねるアンケート項目を新設し、施設風土の定量的把握
- ➡ 回答方式（5件法→10件法とするなど）の変更よりも、**設問文の工夫や回答のスコア化などの設問工夫の方が、天井効果の緩和や回答の多様化に有効。**
- ➡ 留意点として、少年院への入所は年齢や非行種別等により入所施設が割り振られるため、施設ごとの「回答傾向の差」が施設の取組により醸成された「風土の違い」を意味するとは限らず、施設風土の把握は困難であることがわかった。

本調査研究で得られた**テキストデータに対するAI分析のノウハウ**は、他省庁でも以下の取組をする場合に、**応用可能**と考えられる。

- ①要因間の関係を抽出し、優先度の高い対策案を検討、②テキストデータから時系列変化を抽出し、注目すべき事例と要因を把握

実証的共同研究の概要 事例2:全国学力・学習状況調査結果の長期的分析等

背景

- 全国学力・学習状況調査（以下「全国学調」という。）のうち毎年度悉皆で行っている調査は、令和6年度まで、経年での変化を把握することを前提とした調査設計でなかった。
- 経年での**長期的な分析**に当たっては、年度間の**調査項目の整理**や**データクレンジングの上、データを経年的に接続する必要**。

データクレンジング

1.調査項目をつなげる

年度間で異なる設問文や選択肢の接続を検討・整理、調査項目にIDを採番

＜クレンジング前後のイメージ＞

年度	設問文
2021年	学習の中で コンピューター などのICT機器を使うのは勉強の役に立つと思いますか
2022年	学習の中で PC・タブレット などのICT機器を使うのは勉強の役に立つと思いますか

設問文に軽微な変更があるが、同一の調査項目とみなし、**ID**を採番

変数ID	新変数名
SS_S000000_000	学習の中で PC・タブレット などのICT機器を使うのは勉強の役に立つと思いますか

2.調査対象をつなげる

- 全ての年度のデータに共通する学校コード（接続キー）を振る
 - 児童生徒単位のデータを学校・市町村単位に集計
- 各年度のデータを接続し、経年化（同一の学校を経年で追跡可能）

＜クレンジング・経年データ作成のイメージ＞
【単年度 児童単位データ】

	学校名	国語正答率	算数正答率
児童a	A小学校	62	70
児童b	A小学校	65	71
児童c	B小学校	57	67
児童d	B小学校	60	70
児童e	C小学校	65	80

全年度に共通する**学校コード**を採番する

	学校コード	学校名	国語正答率	算数正答率
児童a	0001	A小学校	62	70
児童b	0001	A小学校	65	71
児童c	0002	B小学校	57	67
児童d	0002	B小学校	60	70
児童e	0003	C小学校	65	80

児童生徒単位を
学校単位に集計
↓
各年度のデータを
接続
(経年データ化)

【経年データ】

学校コード	学校名	調査年	国語正答率	算数正答率
0001	A小学校	2015	63.5	70.5
0001	A小学校	2016	65.8	72.3
0001	A小学校	2017	70.3	73.3
...	...	...	...	...
0002	B小学校	2015	58.5	68.5
0002	B小学校	2016	60.8	73.3
0002	B小学校	2017	63.3	72.2
...	...	...	...	...
0003	C小学校	2015	66.2	80.0
0003	C小学校	2016	67.8	78.3
0003	C小学校	2017	69.4	73.2
...	...	...	...	...

※架空のデータ

調査目的

- 以下のノウハウ及び留意点を蓄積し、政府全体に共有することで各府省等におけるEBPMの推進につなげること。
- 長期的な分析を可能とする、①**データクレンジングに関するポイント**、②**個票データを経年的に分析する手法**など

長期的な分析

【分析結果】

学力向上に有効な取組の把握（固定効果モデル※1を用いて、教科調査×質問調査の分析）

■ 学校運営の改善や教職員の資質向上につながる取組が見られる学校では、学力の向上が見られるのではないかな。

■ 学校内の学力格差の是正に有効な取組があるのではないかな。

- その他、コロナ禍前後でのSES（社会経済的地位）と学力の関係を分析したが、明確な傾向は見られなかった。

※1 経年データを用いる分析。経年でアウトカムの変化が見られる学校の特徴となる要素（取組等）を把握することが可能
※2 分析の際には、正答率を標準化したスコア（標準化スコア）を用いることで集団内での相対的な位置をアウトカムとした。

【今後の更なるデータ活用に向けて】

- **SES（社会経済的地位）が低いものの、高い成果を挙げている学校・市町村を抽出**
→抽出された学校・市町村での学力向上に有効な取組を把握し、横展開することが考えられる。
- **外部データを活用した分析方針の検討**
→他の調査結果と接続することで、より有効な分析が可能になると考えられる。

- 本調査研究で得られたデータクレンジングに関するノウハウを「**データクレンジング及び経年データ作成に関する手引き**」に整理

今後、「主・対・深の取組」、ICTの使用頻度を高める取組等の横展開が想定される。

実証的共同研究の概要 事例3:宮城県仙台市における健康増進事業の効果検証

背景・課題

- ❑ 仙台市では、メタボリックシンドロームの該当者割合が他政令指定都市と比較しても高い水準。
- ❑ 仙台市では、R6年度にスマホアプリにより行動を記録し、歩数増加を促す事業を実施したが、十分な定量的評価による効果把握ができていない。
- ❑ 医療費・介護給付費の効果検証事例は限定的

調査目的

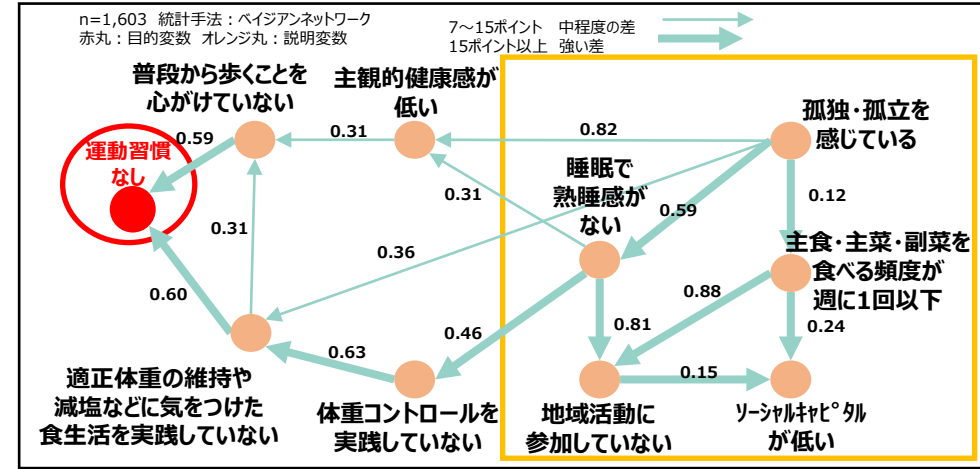
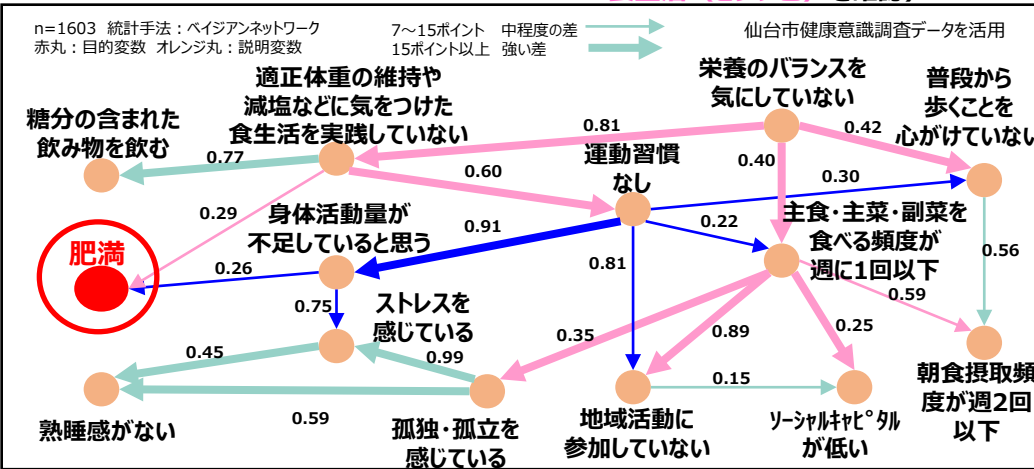
- ❑ 仙台市の健康意識調査結果を基に、健康課題と対策案の検討
- ❑ 仙台市の健康アプリデータを分析し、事業の効果向上のための知見を整理し、試行的に、医療費・介護給付費抑制効果を算出

健康増進のための行動変容を促す効果が見込まれる
対策の知見を共有し、市町村レベルでの健康増進事業
の効果向上につなげる

1.仙台市の健康課題の分析

【分析の手法】

- ① 健康関連指標※について、仙台市と10政令指定都市（札幌市、横浜市など）を比較し、課題を整理 ※ 平均余命、特定健診受診率など21項目
- ②-1 肥満に着目し、要因を可視化（要因として身体活動（青線）、食生活（ピンク色）を確認）
- ②-2 「運動習慣なし」に着目し、要因を可視化（要因としてストレス・地域参加（黄）を確認）



- ③ 身体活動、食生活、ストレス・地域参加に基づき、健康行動を類型化し、ターゲット層を明確化
- (1)健康意識はあるが行動が伴っていない層、(2)健康行動は良好だがストレス課題がある層、(3)心理・社会面の課題を抱える層、(4)生活習慣全体に課題を抱える層

【分析結果】

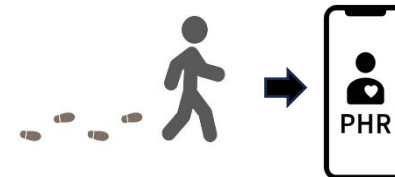
- ・健康課題（肥満など）は、身体活動（身体活動量の認識、運動習慣など）や食生活（朝食や主食・主菜・副菜の摂取状況、栄養バランスへの配慮など）に加え、地域活動の不参加、ストレス、孤立・孤独も重なって発生していることを確認
- ・健康行動の改善が必要なターゲット層ごとの支援策例も整理（歩数や目標達成の可視化による習慣化支援、リラクゼーション・休養に関する情報提供、ストレスチェックや簡易的なセルフケア支援、複数の生活習慣を改善するための包括的な支援など）

2. 仙台市健康増進事業の効果分析

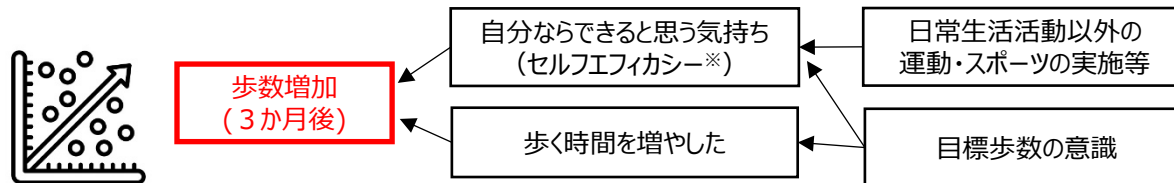
【分析の手法】

① 健康アプリのデータについて、属性等を整理の上、行動結果などを分析

※PHR：Personal Health Recordの頭文字をとった略語。個人の健診結果や予防接種歴、処方されている薬等の保健医療情報を指す。



② 「3か月後の歩数増減」と、健康アプリの利用者アンケート結果を活用し、重回帰分析を実施



※セルフエフィカシー：ある行動を自分がうまく実行できるという見込みや自信。本調査では、運動実施の自信の程度を用いた。

【分析結果】

- ・アプリデータから、**身体活動量が多い健康関心層の参加が多い**（事業参加者平均歩数8,124歩、仙台市健康意識調査の平均歩数5,243歩）。
- ・一方、**歩数データの継続的な取得状況**（歩数データのアップロード率）**は減少傾向**（開始時64.3%⇒3か月後49.9%）を確認。
- ・重回帰分析の結果、**目標設定や運動実施や継続の自信の高さが、歩数増加にプラスの関連を示すことを確認。**
- ・上記分析を踏まえたアウトカムや指標を整理し、ロジックモデル（案）を提示。

そのほか、医療費・介護給付費抑制効果の試行的にシミュレーションを実施し、医療費抑制効果：約101万円/年、介護給付費抑制効果：約123万円/年と試算（注）。

（注）仙台市と類似事業を実施している自治体の医療等データを活用

- ① 事業者保有のデータベースを活用し、有識者の知見も踏まえ、事業参加による医療費等への影響を分析



- ② ①の分析を踏まえ、シミュレーション手順を整理し、仙台市PHR事業の参加者データをあてはめ、年間医療費削減額を試算



- ③ 簡易に試算する手順を作成



【仙台市PHR事業の改善・他自治体での活用等について】

- ・**健康無関心層へのリーチが十分でなかった可能性があるため、口コミなどを通じた効果的なアプローチを検討予定。**
- ・「参加初期段階」は目標設定・達成状況の可視化など、「定着段階」は歩数推移・健康状態の変化のフィードバックなどにより継続を促すといった、**段階に応じて健康のための行動変容を促す事業設計が重要であり、ロジックモデル（案）も参考に施策のあり方を見直し予定。**

- ✓ 他の市町村における健康増進事業において、本分析から得た要因間のつながりや行動変容につながる要因の知見を踏まえ、より効果の高い事業実施が可能。
- ✓ 効果の高い健康増進施策の実践の横展開