

受託者：  
株式会社富士通総研

# 公教育分野に応用可能な 非構造化データ活用事例に関する調査研究の 結果報告

# 目次

## ■ 調査研究の設計・概要

1. 調査研究の目的と調査対象
2. 3つの調査カテゴリ

## ■ 調査研究の結果

3. 非構造化データと構造化データの利点
4. 非構造化データ分析方法と、期待される効果の整理軸の検討
5. 教育分野での事例の一部紹介
6. 教育分野以外での事例の一部紹介
7. 事例マッピングの試行－実現の難易度と利活用の影響度
8. 教育分野のデータ利活用の考察
9. 利活用に際して課題となりうる論点と、今後に向けて

# 調査研究の目的と調査対象

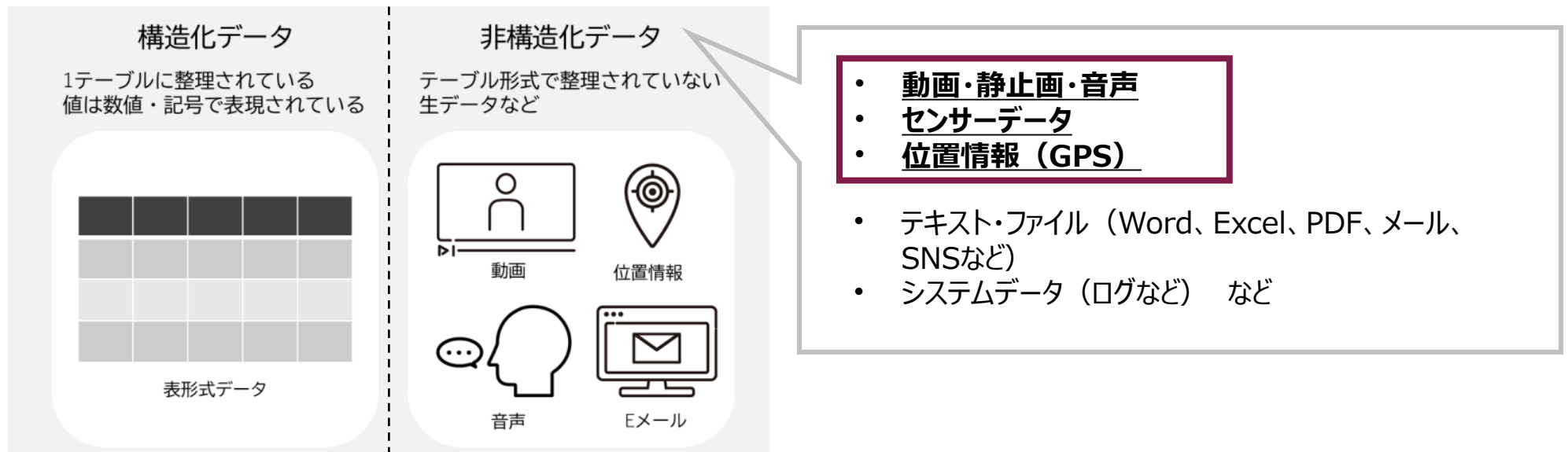
受託者：  
株式会社富士通総研

## 調査目的

- 教育分野に先んじて、**介護や、医療、障害者支援、防犯等の分野においては、センシングやAI等を用いて非構造化データが活用され、実用化が進みつつある状況にある。**またこれらの分野では、音声等の非構造化データの活用により、これまで円滑に意思疎通等ができなかった**状況が改善しているシーン**も見られる。
- そこで、**教育・非教育分野において対人支援等で実装されている非構造化データの活用**の状況について探索的に調査し、小・中・高等学校の公教育分野での応用可能性についてヒントを得る。

## 調査対象とした、非構造化データとは？

- 非構造化データは、音声・静止画・テキストなど、そのままでは表形式に変換できない生データのことを指し、主に以下のようなデータ種別（形式）がある。
- これらのデータ種別のうち、教育分野においては、**赤枠の動画・静止画・音声、センサーデータ、位置情報**について、利活用が進んでいない。
- そのため、本事業では、主に赤枠内のデータ種別を非構造化データの対象として調査を実施する。



出典：総務省統計局「高等学校における「情報II」のためのデータサイエンス・データ解析入門」より抜粋

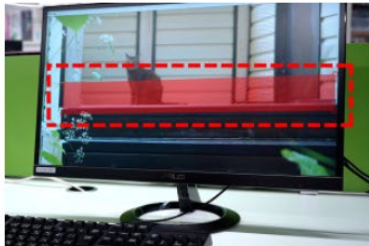
# 他分野で活用されている事例

受託者：  
株式会社富士通総研

## 骨格情報を捉える動画データで、事故やケガの未然防止を迅速に

- 専用アプリ（AIのディープラーニング技術を使った姿勢推定AIエンジン「VisionPose」）をインストールした専用のカメラで、予め指定した特定エリア内における特定の身体（骨格情報）の侵入や接触を検知できる。
- 動画データにより骨格情報をタイムリーに捉えることで、その場での警報発信等を行い、事故やケガ等の予防や回避につなげることが可能。
- 骨格情報を把握できる技術で、体の部位単位の細かい条件設定が可能。そのため、同技術によって乳幼児の見守りや、病院・介護施設の事故防止の事故防止などにも活用可能。
- また、検知結果については、アラート通知をメール等で送信させることが可能なため、警備に係る人件費などのコストカットにも寄与しうる。

1. 検出エリアを設定



カメラ映像範囲内で、アラートを通知するためのエリアを、矩形表示で設定を行います。

2. 人体の検知



人体を検知します。猫などの小動物などには反応せず、誤検知を減らします。

3. アラート通知



検出エリアに人体の部位が入ったらアラートを表示します。

手や足だけの侵入など、体の部位単位で設定が可能です。



上記は、指定したエリアに両足が入ったタイミングで、アラートが表示されている様子です。

乳幼児の見守り



自宅や保育園などの児童福祉施設の見守りに、徘徊による不慮な事故などを未然に検知します。

病院・介護施設の事故防止



高齢者や認知症患者の病室、住居での転倒や無断外出などの見守りに。危険行動を検出して介護負担を軽減など。



## 1. 教育分野における非構造化データの活用状況整理

- 国内外の事例について、以下の視点で机上調査、ヒアリング調査を実施。
- 教育分野におけるデスクリサーチ結果は37件（国内18件、海外19件）、ヒアリング結果は5件となった。
  1. データ取得の目的
  2. データ取得対象
  3. 取得方法（含 コスト）
  4. データ分析方法（含 コスト）
  5. 分析結果の還元方法
  6. データの取得・活用の効果/課題
  7. 取組もたらす影響、応用可能性 / 等

## 2. 教育を除く対人支援等の幅広い分野での非構造化データの活用状況整理

- 教育分野と同様の視点に加え、教育分野への応用可能性も含めた観点で、国内外の事例について机上調査、ヒアリング調査を実施。
- 机上調査結果は48件（国内26件、海外22件）、ヒアリング結果は5件（国内で視線や体動、声帯の周波数等を対象データとするもの）となった。
- 事例収集ができた分野は介護、医療、児童福祉、防犯、マーケティング、健康増進 等多岐にわたる。

## 3. 有識者ヒアリング

- 下記3名にヒアリングを実施し調査設計等について、意見を得た。
  - 神野元基（株式会社COMPASS ファウンダー、学校法人東明館学園 副校長/公教育でのデータ利活用）
  - 菅野 裕介（東京大学/コンピュータービジョン）
  - 松田 壮一郎（筑波大学/応用行動分析学 等）

# 非構造化データと構造化データの利点

受託者：  
株式会社富士通総研

- 構造化と非構造化の各データにはそれぞれ**利点と課題・留意点**があり、**活用目的に応じた選択**が必要であることが伺えた。
- ただし、調査時点では非構造化データは、**現在開発途中のものも多い点に留意**が必要。

|         | 利点   | 現状の課題や留意点   |
|---------|--|---|
| 構造化データ  | <ul style="list-style-type: none"><li>● 機械学習アルゴリズムで簡単にデータを利用可能</li><li>● <u>データに対するユーザの理解が容易</u></li><li>● <u>活用可能なツール・サービスが豊富</u></li><li>● 個人情報等、プライバシーに関する情報を含まない方法で取得しやすい</li><li>● データの精度を考慮する必要はほぼない</li></ul> | <ul style="list-style-type: none"><li>● <u>事前定義が必要なため早期収集が困難</u></li><li>● 利用目的が事前定義されることによる利用の限定</li><li>● 一つの情報に対して一つの処理パターンを定義する必要があり、<u>表情などのばらつき（幅）のある情報を定義することが困難</u></li></ul>                               |
| 非構造化データ | <ul style="list-style-type: none"><li>● 事前定義が不要なため<u>早期に収集可能</u></li><li>● 事前定義が不要なため幅広い目的で利用可能</li><li>● 表情などのばらつき（幅）のある情報についても一つのパターンとして取り扱うことができるため、<u>様々なデータ分析が可能</u></li></ul>                                  | <ul style="list-style-type: none"><li>● <u>利用するにはデータサイエンスの専門知識が必要</u></li><li>● 活用・分析には専門的なツールが必要</li><li>● 活用可能なツール・サービスは少ない</li><li>● <u>個人情報等、プライバシーに関する情報が含まれることが多い</u></li><li>● データの精度を確保するための考慮が必要</li></ul> |

# 非構造化データ分析方法と、期待される効果の整理軸の検討

受託者：  
株式会社富士通総研

- 今回の調査研究では、非構造化データ分析には①個別データを直接分析するものと、②全体データを把握したうえで、全体からずれがあるデータを抽出するものがあるという整理を試行した。
- 各分析の方法ごとに、教育分野で期待される効果が異なり、分析方法によって必要なデータ種別、対象なども異なることが予想される。このため、1:2つの分析方法(個別/全体)に加え2:主な期待効果、③データ種別、④データ対象で整理した。

| 分析方法              | 教育分野における期待効果<br>(非構造化データ活用で可能になりうることの例) |
|-------------------|---|
| ①個別データに対する分析(※)   | 特性(認知、興味関心)に応じた教育・指導                    |
|                   | 学習意欲・学習状況の把握・維持改善                       |
|                   | 授業・コミュニケーション活動の改善                       |
|                   | 主体性・協働性の評価・改善                           |
|                   | スポーツ、芸術活動等の改善                           |
|                   | 障害等を持つ児童・生徒等のコミュニケーション維持・改善             |
| ②全体データに対する特定要素の検出 | 児童・生徒同士の関係性等の維持・改善                      |
|                   | 心身の健康維持管理                               |
|                   | 事故・ケガ等の発生の事前把握・解消                       |
|                   | 見守り(不審者)                                |

※個別データに対する分析時には、他のデータ(データ取得を望まない人等)に対する配慮(映り込み防止やマスキング処理等)が必要

# 教育分野での事例の一部紹介

受託者：  
株式会社富士通総研

| 分類                | 利活用の効果                      | 活用例  | データの種別        | データ対象                  | 国内外 | 実用化の状況    |
|-------------------|-----------------------------|--|---------------|------------------------|-----|-----------|
| ①個別データに対する分析      | 特性（認知、興味関心）に応じた教育・指導        | 家庭教師の授業中（オンライン）の映像の声や表情や脈拍のデータから、生徒の状態（感情の変化、集中度など）を把握することにより、教員の授業の改善に活かせるとともに、学生一人一人の状況をより深く理解した個別指導の実践が可能。さらにデータは保護者にも共有できる。  | 動画、音声         | 表情、顔の向き、視線、音声、生体情報（脈拍） | 国内  | 実用化済      |
|                   |                             | 学習者のPPG信号（光学式心拍センサー）と顔の表情を分析することで、学習中の学習者の感情や認知の状態を暗黙的に推測することにより、学習中の学習者の感情や認知状態を講師へフィードバックでき、MOOCにおける講師と学習者の間の限られたインタラクションを向上することで低い修了率、低い熱中度、および個別最適な学習の改善を図る。   | 動画、センサー       | 表情、心拍数                 | 海外  | 未実用化      |
|                   | 学習意欲・学習状況の把握・維持改善           | 読書中の学生の高度な画像処理アルゴリズムと目の生理学的3Dモデルを用いて、空間における目の位置と注視点を高精度に推定することにより、学習行動、認知負荷、熱中度についての洞察を得ることができ、教材や学習状況を効果的に設計、評価、改善するのに役立つ。さらに、非典型的な読書パターンを明らかにするために使用でき、ディスレクシア（読字障害）に対する理解のキッカケとなり、失読症などの学習障害の客観的な診断に活用する可能性がある。 | 動画、センサー       | 視線                     | 海外  | 実用化済      |
|                   |                             | 教室にカメラとマイクを設置することで、授業中の発話や挙手などの行動記録を取得し、AIで分析することで、教師や児童生徒の行動、様子に関する客観的なデータを取得することができる。客観的なデータを活用することで、授業研究の効率化、授業の改善、児童生徒の学力の向上を目指す。  | 音声、動画         | 授業時の行動、発話といった活動情報      | 国内  | 実用化済      |
|                   | 障害等を持つ児童・生徒等のコミュニケーション維持・改善 | 熟練教師と教育実習生にウェアラブルカメラを装着してもらうことで、授業中の視線の映像を取得し、更にインタビューにおいて視線の対象やその意図について確認する。  | 動画            | 視線                     | 国内  | 未実用化      |
|                   |                             | 重度・重複障害があり、リハビリテーションを行っている人の腕に、ウェアラブル生体情報センシングデバイスを着けて生活し、活動量・心拍数・ストレスレベル・睡眠などのデータを保護者のスマートフォンへ送る。並行して、専用のカメラを使って療育活動の様子を撮影し、映像編集アプリケーションによって心拍数を表示。これにより、重度・重複障害があり、リハビリテーションを行っている人の心身の状態をタイムリーに把握することができる。      | センサー          | 活動量(心拍数、歩数)、睡眠状態       | 国内  | 未実用化      |
| ②全体データに対する特定要素の検出 | 事故・ケガ等の発生の事前把握・解消           | ソーシャルスキルトレーニングにおいて、VRを活用することで、これまでよりも日常生活に近い形で何度も練習できるため、学習効果を高めることができる。VR体験中の視線を解析することで、トレーニングで行う動作の意味を理解しているかなど、利用者の特性理解に活用できる。  | センサー          | VR体験中の児童の視線            | 国内  | 実用化済      |
|                   |                             | ProctorUは、高度な顔認識、物体認識、平面検出、音声テキスト化、目の動き検出、音声検出などの技術を使い、オンラインでの試験の際に、本人確認から試験中の監視などの試験監督をするソフトウェアであり、すでに実用化されている。   | 動画、音声         | 顔、声、視線                 | 海外  | 実用化済      |
|                   | 見守り（不審者）                    | ランドセルに「親カメ」（カメラ型見守り端末）をつけ、位置情報や写真送信により、登下校時の児童の見守りができるとともに、危険を検知し通知できる。  | 静止画、位置情報、センサー | 地図上の位置、視界情報            | 国内  | 8<br>実用化済 |



# 教育分野以外での事例の一部紹介

受託者：  
株式会社富士通総研

| 分類                | 利活用の効果                     | 活用例   | データの種別     | データ対象         | 国内外 | 実用化の状況 |
|-------------------|----------------------------|---|------------|---------------|-----|--------|
| ①個別データに対する分析      | 特性（認知、興味関心）に応じた教育・指導       | 実車やドライビングシミュレータで、ドライバーの視線を対象に、センサーを活用して視線の計測を行う。対象の計測機器を用いることで、視線が停留・注視している時間が定量的に評価可能となる。ドライバーがどのように気を配りながら運転しているか、メーターやカーナビの視認性、操作性の評価など、視覚情報を分析、評価することで、より安全で快適な自動車づくりに活用する。           | センサー（視線情報） | 視線（方向、箇所）     | 国内  | 実用化済   |
|                   | 学習意欲・学習状況の把握・維持改善          | 利用者の視線を対象に、可視光カメラを使って視線の動きを記録し、顔認識AIで視線情報を検出する。視線情報によって、利用者の興味関心・警戒・探索行動・操作要求を分析することができる。（遠隔視線推定等）  | 静止画        | 視線            | 国内  | 実用化済   |
|                   | 授業・コミュニケーション活動の改善          | 顧客の音声を対象に、最初の数秒の会話で音声分析AIが顧客を自動認証し、対象の顧客データを提供する。また、音声から顧客の感情を分析し、対話内容からオペレータに必要な情報をタイムリーに提供することで、顧客へのサービス提供の迅速化、患者対応品質の向上、医療関連文書管理の負担を軽減を可能とする。  | 音声         | 顧客の音声         | 海外  | 実用化済   |
|                   | スポーツ、芸術活動等の改善              | IoTベースのアスリートパフォーマンス評価ソリューション。アスリートを対象に、ウェアラブルデバイスを身に付けさせ、コーチがセンサーデータに基づいてアスリートのパフォーマンス（心拍数、体位、加速度）を評価することで、運動や力の運動が適切に行われているかどうかを判断することができる。トレーナーがトレーニング計画をパーソナライズし、アスリートの結果を改善することに寄与する。 | センサー、位置情報  | アスリートの生体情報、動作 | 海外  | 実用化済   |
|                   | 障害等を持つ児童生徒等のコミュニケーション維持・改善 | 聴覚障がい者と健聴者を対象に、モバイル型ロボット電話（ロボホン）を使って、センサーやカメラで聴覚障がい者が行う手話の動きをAIで認識・分析する。手話の意味する単語を日本語で発話し、また、発話した内容を外部デバイス（スマートフォンなど）に表示するといった一連の流れができるアプリケーションを開発している。                                   | カメラ、センサー   | 手話の動き         | 国内  | 共同開発中  |
| ②全体データに対する特定要素の検出 | 心身の健康維持管理                  | 利用者の音声（声帯の周波数等）を対象に、マイクで収集したデータを音声解析プログラムを用いて解析し、元気圧（メンタル状態の数値）と活動値（過去2週間のメンタル負荷の積み重ね）に数値化する。「心の状態」を可視化することで、ストレスの把握や心の病の予防等に活用する。  | 音声         | 声帯の周波数等       | 国内  | 実用化済   |
|                   |                            | 利用者の日々の顔写真を対象に、スマートフォンに搭載されているカメラ等を用いてデータを収集し、顔写真からAIが「喜び」や「苦しみ」などの感情を分析する。モチベーションを可視化することで、従業員一人ひとりの変化を把握、感情の落ち込みに対して迅速にケアを行う。メンタルヘルスの不調を未然に防ぐ。  | 静止画        | 人の顔写真         | 国内  | 実用化済   |
|                   |                            | スマートフォンのカメラアプリケーションで撮影した動画・顔画像を対象に、動画認識AIを搭載したアプリケーションにより、利用者の顔の表情、声からの感情、スピーチを分析し、ボディマス指数（BMI）、血圧、脂肪の割合を測定する。視覚情報に基づいた健康状態が把握できる。  | 静止画・動画     | 利用者の動画・顔画像    | 海外  | 実用化済   |
|                   | 事故・ケガ等の発生の事前把握・解消          | 歩行者を対象に、単眼カメラを使って歩行者の動きをデータ収集し、行動予測AIを使って人の行動を高精度に予測する。人と車両との接触リスクを把握し、事故の発生を未然に防ぐ。行動予測AIは、コンピュータービジョンに行動心理学の要素を加え、これまで困難とされてきた歩行者の行動を予測する。自動運転への搭載が期待され、各国で実証実験が行われている。                  | 動画         | 歩行者の行動        | 海外  | 実証実験中  |
|                   | 見守り（不審者）                   | 人物を対象に、監視カメラを用いて画像データを収集し、事前に200億枚以上の人物画像を学習したAIにより、外乱光や植栽のある環境や夜間の暗い環境においても、高精度に人物を検知する。例えば、侵入禁止エリアへの侵入やフェンスの乗り越え等について、人物のみを検知し、リアルタイムにPCやスマートフォンに通知することが可能である。                          | 動画         | 人物画像          | 国内  | 実用化済   |

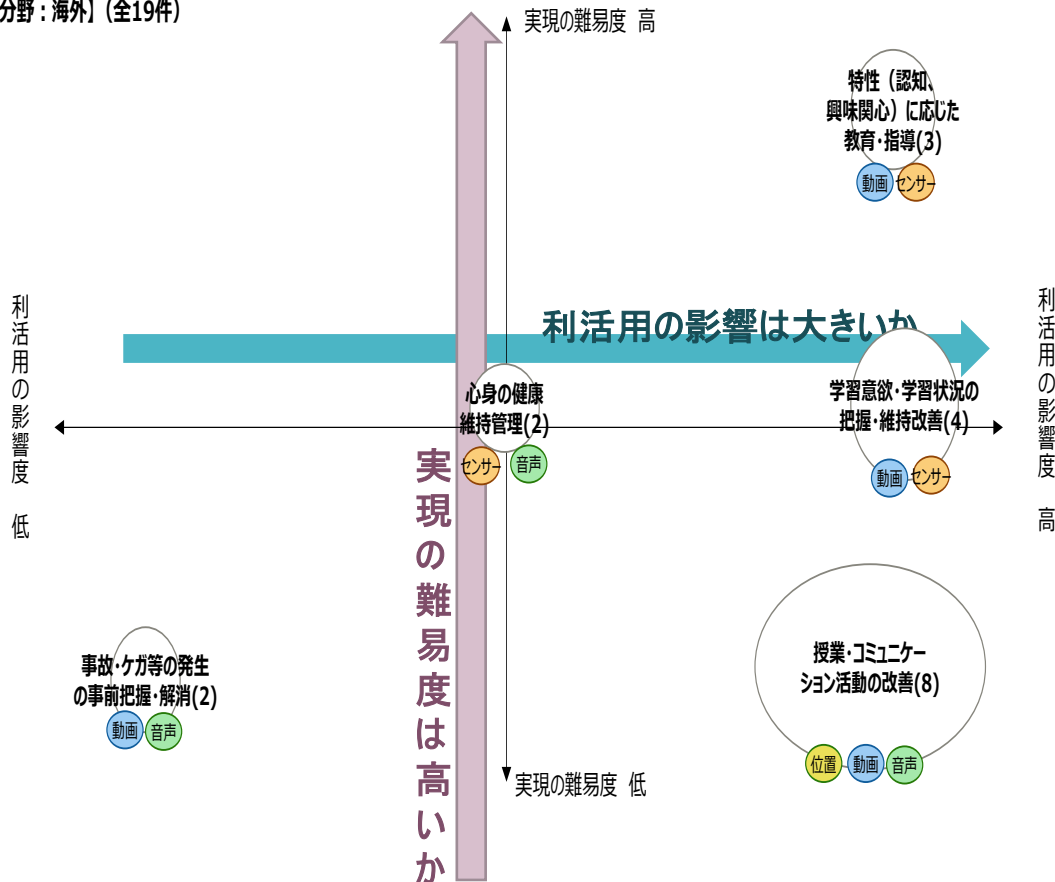
# 事例マッピングの試行 – 実現の難易度と利活用の影響度

受託者：  
株式会社富士通総研

- 本調査では、教育分野・教育分野以外・国内・国外のそれぞれにおいて①2つの分析方法（個別/全体）、②主な期待効果、③データ種別、④データ対象で整理した。
- これに加えて実現の難易度（コスト（金銭）、プライバシー、カスタマイズ、心理的安全性）の観点と、利活用の影響度（利活用のしやすさ、業務の範囲、利活用の価値）の観点で試行的に整理を行った。

## 【例：国外の教育分野の事例の場合】

【教育分野：海外】（全19件）



| 分析方法              | 期待効果                 | データ種別 |    |     |      |      | 主なデータ対象     | 実現の難易度 | 利活用の影響度 |
|-------------------|----------------------|-------|----|-----|------|------|-------------|--------|---------|
|                   |                      | 動画    | 音声 | 静止画 | センサー | 位置情報 |             |        |         |
| ①個別データに対する分析      | 特性（認知、興味関心）に応じた教育・指導 | ○     | ○  |     | ○    |      | 表情、声、心拍数、視線 | 高      | 高       |
|                   | 学習意欲・学習状況の把握・維持改善    | ◎     | ○  |     | ○    |      | 視線、表情、生体情報  | 中      | 高       |
|                   | 授業・コミュニケーション活動の改善    | ○     | ◎  |     |      | ○    | 声、動作        | 低      | 高       |
| ②全体データに対する特定要素の検出 | 心身の健康維持管理            |       | ○  |     | ○    |      | 声、生体情報      | 高      | 中       |
|                   | 事故・ケガ等の発生の事前把握・解消    | ○     | ○  |     |      |      | 顔、声、視線      | 低      | 低       |

### 【国外・教育まとめ】

授業中の教師・生徒の音声データを使ったコミュニケーション活動の分析や授業改善に関する事例が多い。

- 教育分野においては、国内外ともに授業中の教師・生徒の音声データを使ったコミュニケーション活動の分析や授業改善に関する事例・研究が多い。（発話バランス・発話量の可視化）
- また、表情・視線やセンサーで生徒の生体情報を分析することで興味関心や集中度を把握する研究も一部進んでおり、海外ではサービス化されているものもあった。
- さらに国外と比較すると、教育分野ではその活用の範囲や、結果の提供内容など、発展の伸びしろがあることが伺えた。

|         | 国内の教育分野   | 国外の教育分野 <u>（より進展している部分）</u>  |
|---------|---|--|
| 活用事例    | <ul style="list-style-type: none"> <li>■ 音声データを使ったコミュニケーション（発話バランス・発話量等）の分析事例が多く、生徒の学習意欲（集中度）や認知に関する分析は少ない。</li> <li>■ ただし、一部オンライン授業（家庭教師）で生徒の状態（感情の変化、集中度など）を把握したものが実用化されている。</li> </ul> | <ul style="list-style-type: none"> <li>■ 音声データを使ったコミュニケーションの分析事例に加え、動画やセンサーを用いて、<u>生徒の学習意欲（集中度）や認知に関する研究が多く、実用化されているもの複数ある。</u></li> </ul> |
| データ取得対象 | <ul style="list-style-type: none"> <li>■ 生徒が中心となっており、教師のみをターゲットとした事例は少ない。</li> </ul>  | <ul style="list-style-type: none"> <li>■ 生徒だけではなく、<u>教師の動作・動線や発話品質（話し方等）を分析している事例も多く見られる。</u></li> </ul>                                   |
| 分析結果の提供 | <ul style="list-style-type: none"> <li>■ 分析は結果については、AI等による音声データの可視化・数値化等に留まっている場合が多い。</li> </ul>  | <ul style="list-style-type: none"> <li>■ 音声データを使った分析においても、動線や話し方等を教師にAIが評価した上で<u>フィードバックするものがある。</u></li> </ul>                            |

- なお、非教育分野の事例からは、データ取得対象者、データ利用者から見る利点と、データ取得にかかるコストの2要素が実用化の発展程度に影響を及ぼしている可能性が伺える。

## 利活用に際し、今後課題となりうる論点

- 今後、非構造化データの利活用を進める場合には、各観点別の課題について検討を進める必要がある。

| 課題・制約事項<br>の観点 | 課題・制約事項                     |
|----------------|-----------------------------|
| 人材・体制面         | 人手・作業時間の調整・確保               |
|                | データ利活用における専門知識・スキルを持った人材の確保 |
| コスト面           | 導入に係るコストの確保                 |
|                | 維持・運用に係る費用の確保               |
| セキュリティ面        | 個人情報への対応                    |
|                | プライバシー面への対応                 |

| 課題・制約事項<br>の観点 | 課題・制約事項                         |
|----------------|---------------------------------|
| 環境面            | ノイズ等が入らないデータ取得環境の整備             |
|                | 精度を維持する条件を満たすことが必要              |
| データ管理面         | データ量に応じた適切な管理単位・環境の整備           |
| 業務・利用面         | データ利活用の業務（授業等）への取り込みを上手く行うことが必要 |
|                | フォーマットの統一の必要性                   |
|                | 利用用途の制約                         |

## 今後に向けて

- 現在は実証研究中のものも多いが、既に実用化も進みつつある事例もあるため、これらを参照しつつ**利活用シーンをより明確化**することで、普及が進む可能性がうかがえる。
- 但し、この利活用シーンの具体化やユースケースの蓄積に向けては、**利活用の価値をデータ取得対象者が納得できる**状況を創出することも重要と考えられる。
- いずれにしても、特に教育分野における非構造化データ活用はまだ研究途上の段階にあり、一層の研究が進むことが期待される。