

第2講データの種類と収集方法

尾関智恵

(岐阜大学 高等研究院 航空宇宙生産技術開発センター)

令和7年度

AI（人工知能）概論【II】
～データサイエンスから見える新たな学びの未来像～



もくじ

01 教育研究におけるデータとは

02 データの種類と特徴

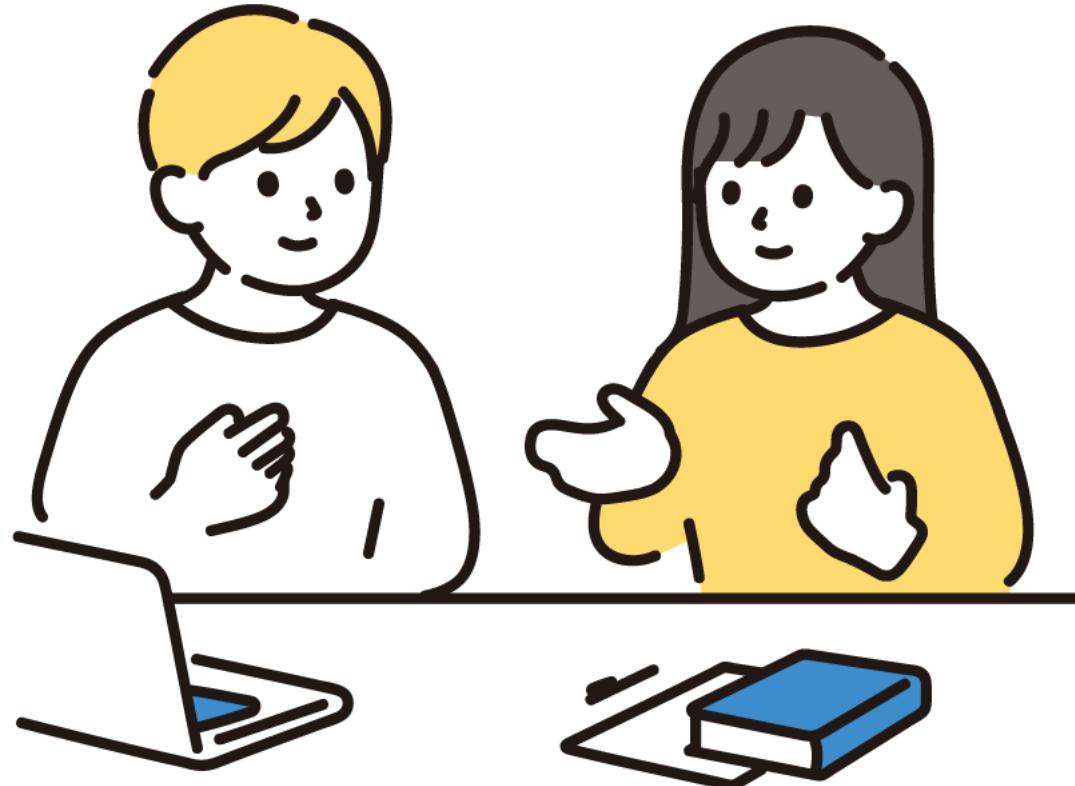
03 データの収集方法

04 データの信頼性と妥当性

05 まとめ



教育研究におけるデータとは



- 主観を客觀化する
- 経験を共有可能な知識に変換する
- 教育実践の改善サイクルを回す基盤

教育研究における統計の役割

1. 客観的な根拠を提供する

- ・「この学習活動は効果がありそう」という印象を、数値で検証できる
- ・個人の経験や直感だけでなく、データに基づいた判断ができる

2. 一般化可能な知見を得る

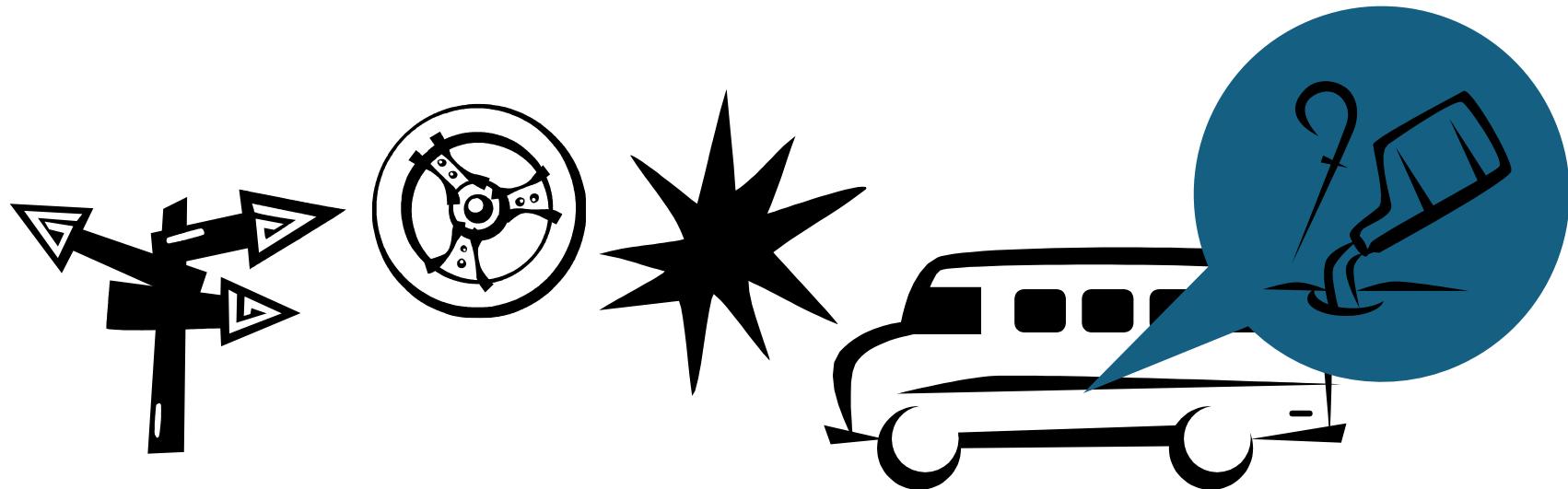
- ・目の前の30人のクラスでの結果が、他のクラスや学校でも当てはまるかを推測できる
- ・特定の文脈での観察を超えて、より広い適用可能性を議論できる

3. 複雑な現象を整理・理解する

- ・多くの要因が絡み合う教育現場で、何が重要な要因かを見極める
- ・データから意味のあるパターンを抽出したり、要因の相対的な影響力を比較

データの種類と特徴

世の中は・・・物質／エネルギー／情報



「情報」の定義

- 1 ある物事の内容や事情についての知らせ。
- 2 文字・数字などの記号やシンボルの媒体によって伝達され、受け手に状況に対する知識や適切な判断を生じさせるもの。
- 3 生体系が働くための指令や信号。神経系の神経情報、内分泌系のホルモン情報、遺伝情報など。

(大辞林より抜粋)



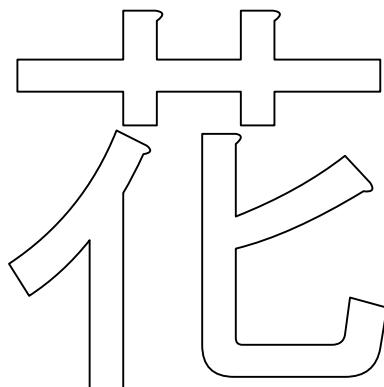
たとえば



パターン



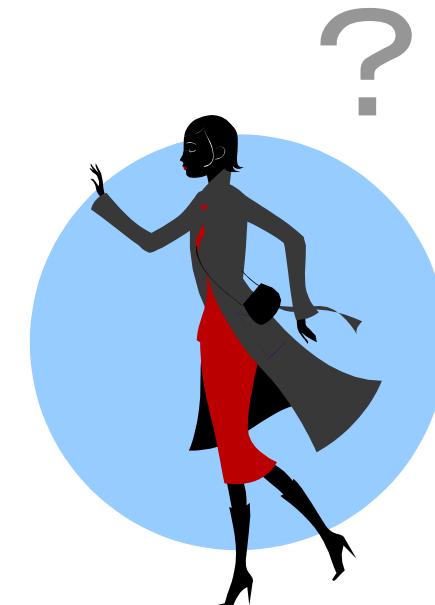
記号



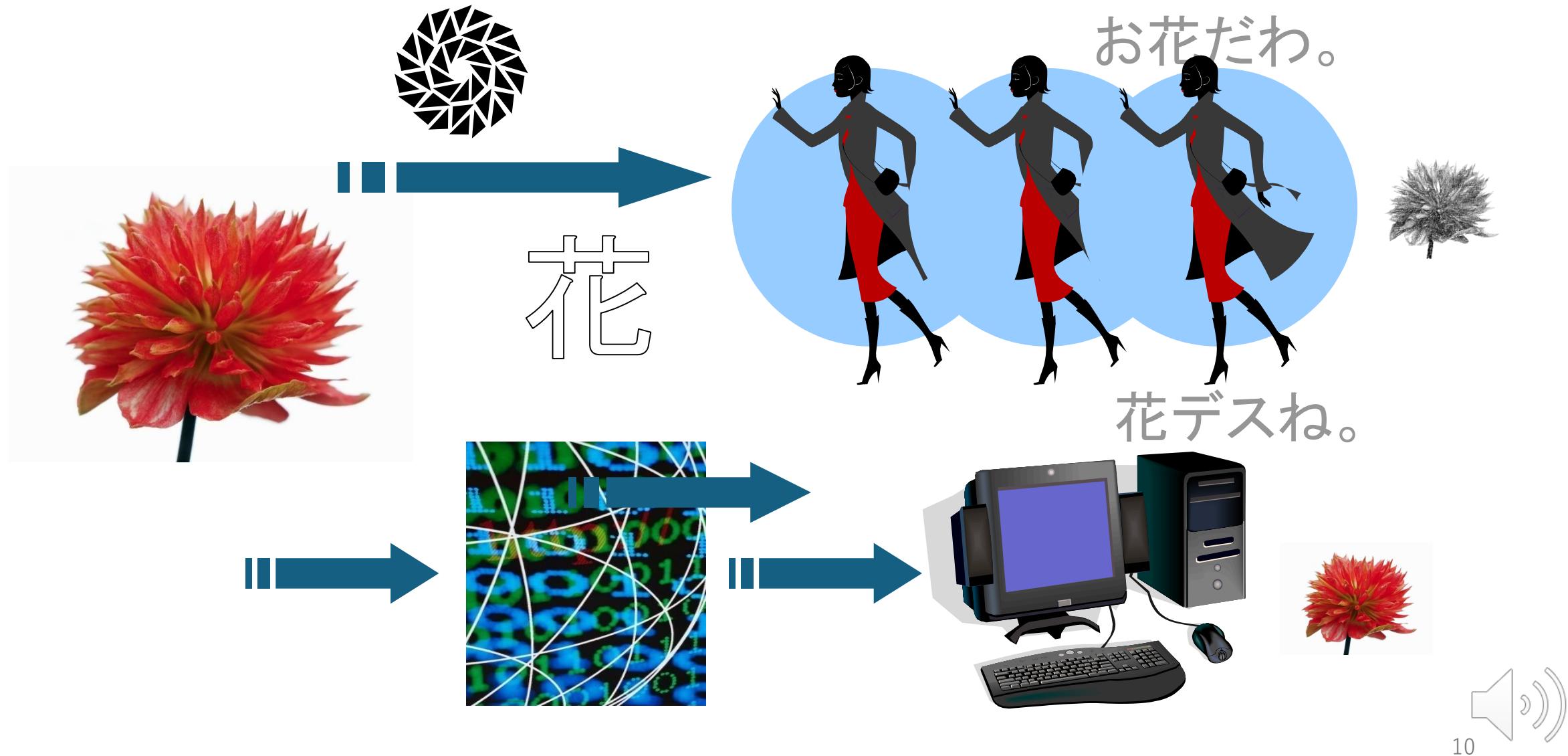
知識

種子植物の生殖器官。一定の時期に枝や茎の先端などに形成され、受精して実を結ぶ機能を有するもの
きれい。
春に咲く。
美しいものを形容することば。

情報を再度活用するためには



長い期間情報を伝えるには・・・



デジタル

- ・離散的に整数値（すなわちdigit）で表現すること
- ・デジタルコンピュータにおいては、2進数（0と1だけ）で表す

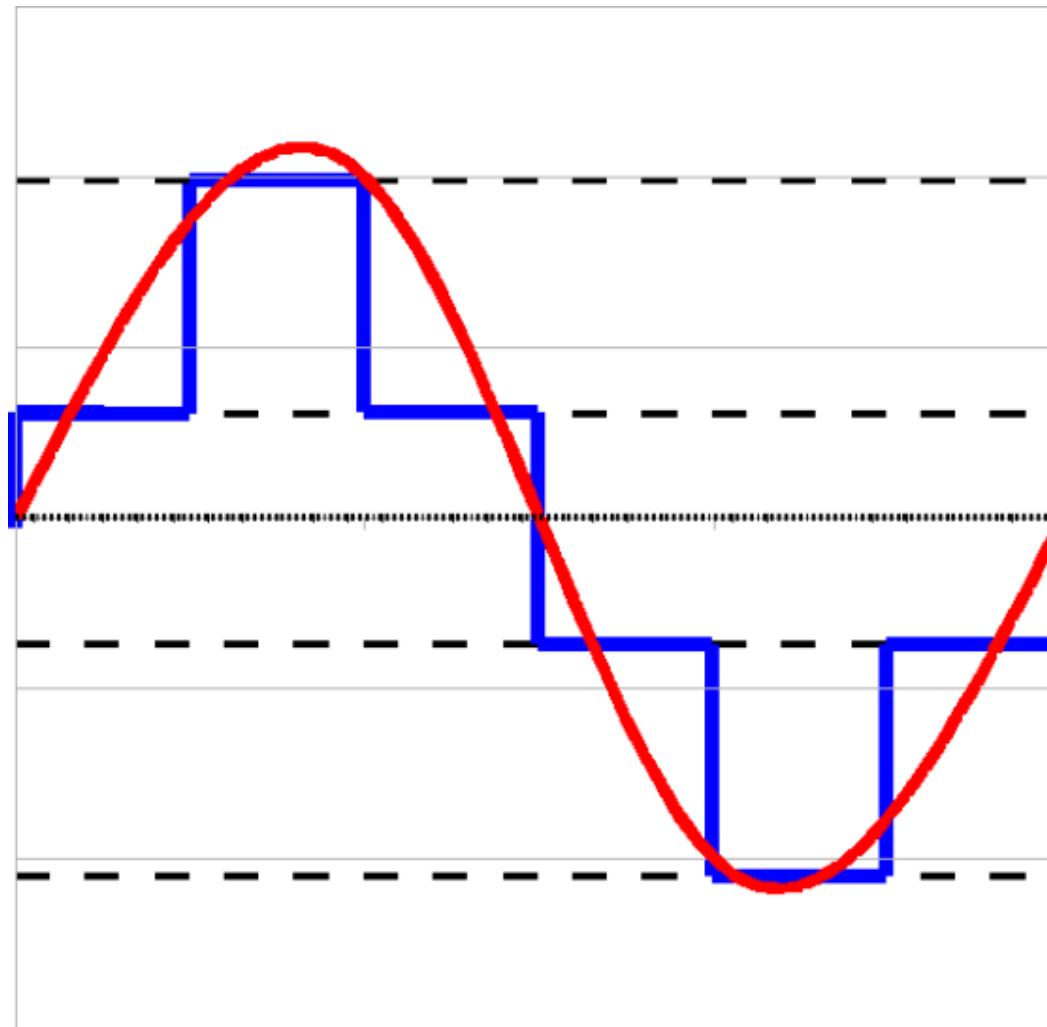
アナログ

- ・連続した量（例えば時間）を他の連続した量（例えば角度）で表示すること。
- ・時計や温度計などがその例で、あいまいな表現が可能



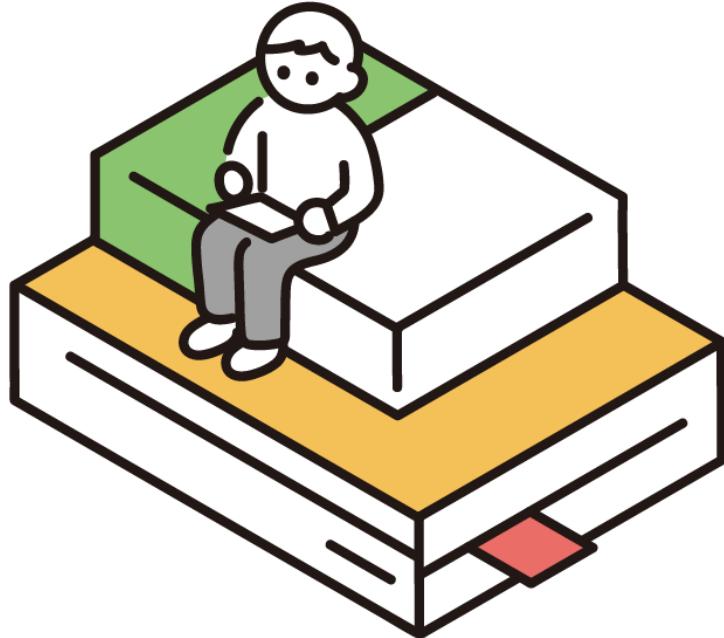
デジタルとアナログ

離散値
連続値



11
10
01
00

量的データと質的データ



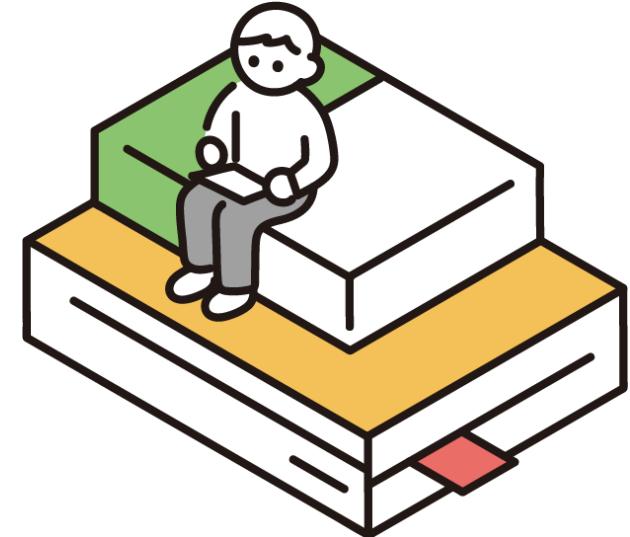
数値化できる



数値化が難しい

量的データ

- テストスコア、出席率、課題提出数
- アンケートの評定値（5段階評価など）
- 学習時間、発言回数



特徴：
統計分析が可能、客観的、大規模調査に向く

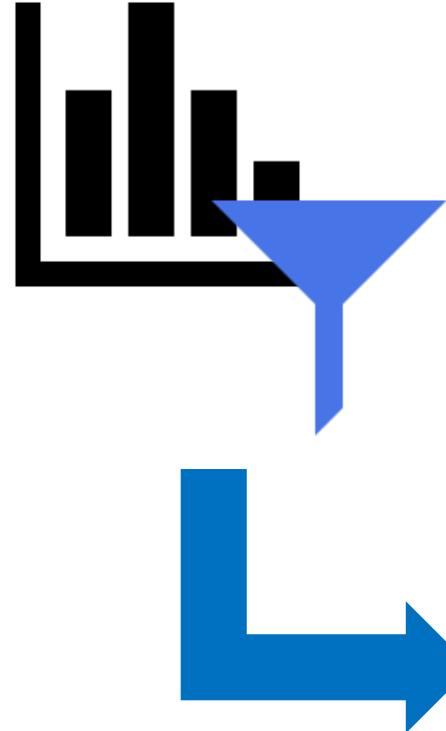
質的データ

- ・学生の自由記述回答
- ・インタビューの発話内容
- ・授業観察のフィールドノート



特徴：
文脈や意味の深い理解、少数でも深い洞察

入力された情報の特徴抽出（パターン認識）
制約を利用する論理的な処理



データ駆動型処理

客観科学的

主観科学的

解釈
理解
学習

相互に影響

既有知識・経験・期待（スキーマ・バイアス）
概念を基にした確認処理



概念駆動型処理

尺度について

名義尺度（カテゴリー）

- 単なる分類・ラベル
- 例：性別、専攻、出身地、学習スタイル
- 統計：度数、割合、カイニ乗検定

順序尺度（順位）

- 順序に意味があるが、間隔は等しくない
- 例：成績（A・B・C）、
理解度（理解できた～理解できない）
満足度（とても満足～全く満足しない）
- 統計：中央値、順位相関

間隔尺度（等間隔）

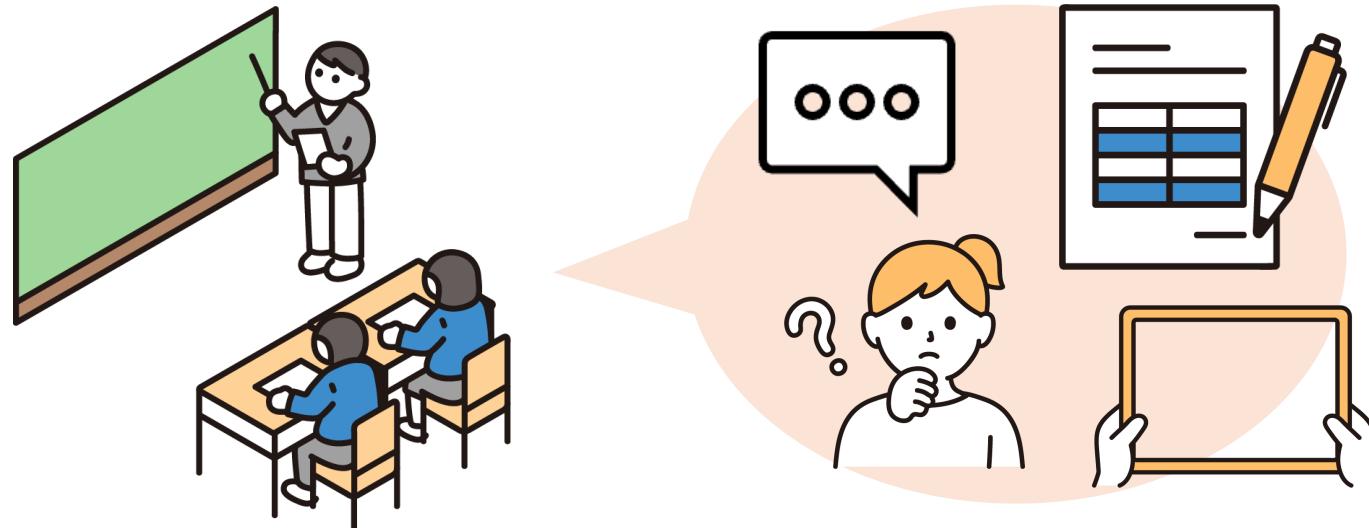
- 間隔が等しいが、絶対的な0点がない
- 例：偏差値、IQスコア、温度（摂氏）
- 統計：平均値、標準偏差、相関係数、t検定

比率尺度（絶対0点あり）

- 間隔が等しく、絶対的な0点がある
(0 = 何もない状態)
- 例：テスト得点、学習時間、年齢
- 統計：すべての統計手法が使用可能、
比率の計算も可能

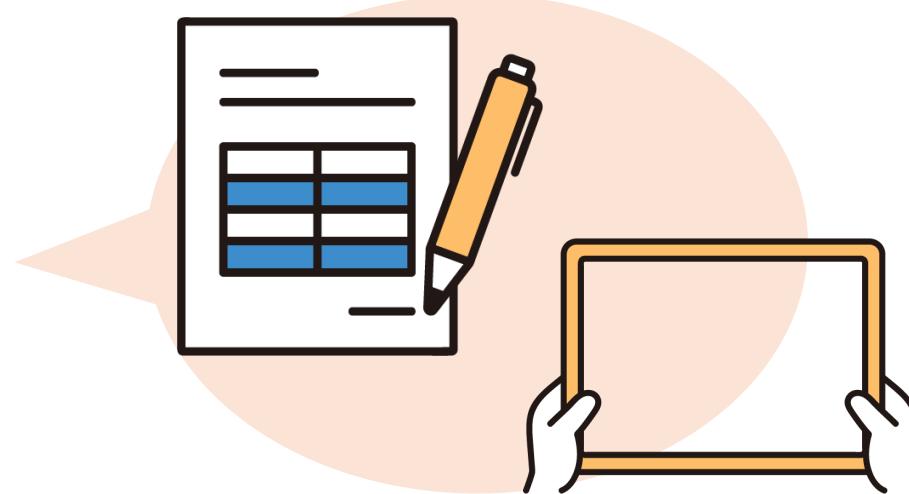
データの収集方法

- 学習過程（発散や収束などの段階が観察できるもの）や教材との提供タイミングや説明等、学習活動が行われている際に発生する情報は、より質の高い学習を構築・検討するために必要



学習成果物・質問紙

- 学習過程を捉えるために有効
- 学習成果を直接確認できる
- ループリックを使った評価など解釈方法が難しい



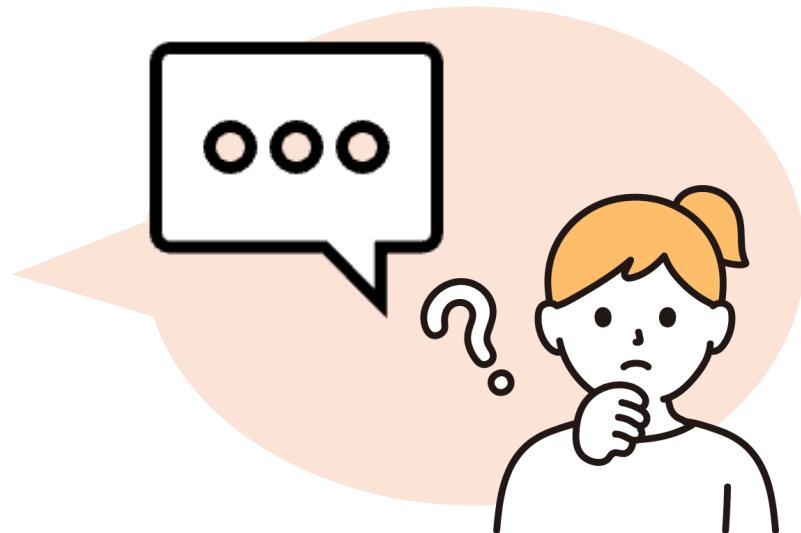
観察

- ・実際の行動を直接捉えられる
- ・無意識の行動や相互作用を把握
- ・自己報告では得られない情報



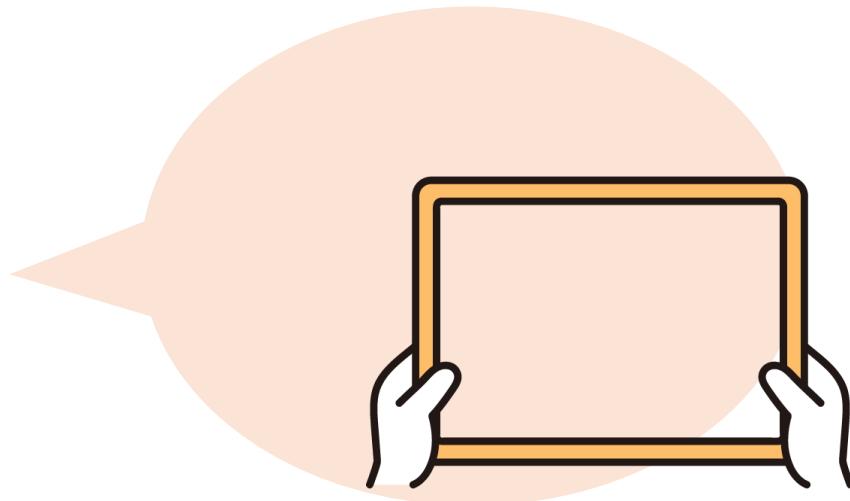
インタビュー

- ・深い理解が得られる
- ・予想外の発見がある
- ・「なぜ」を探れる



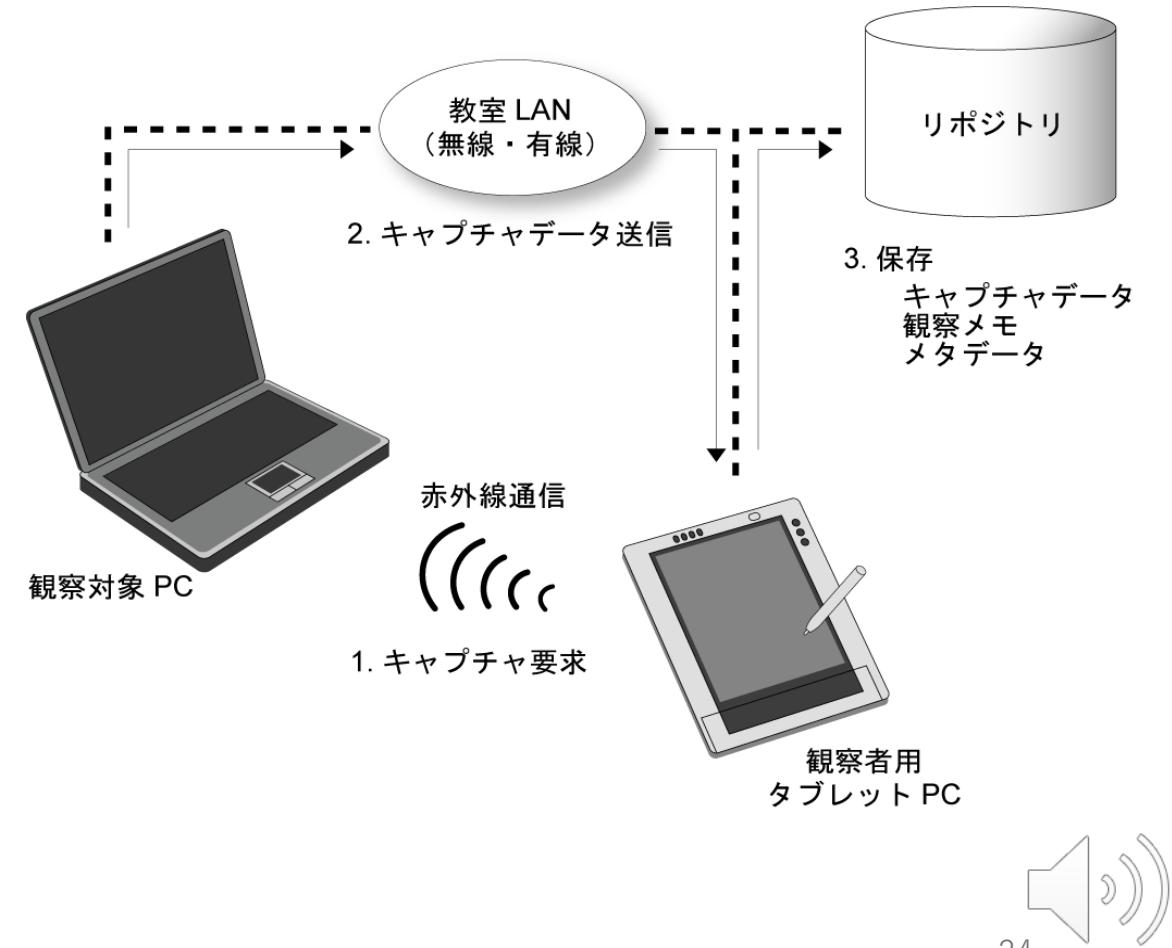
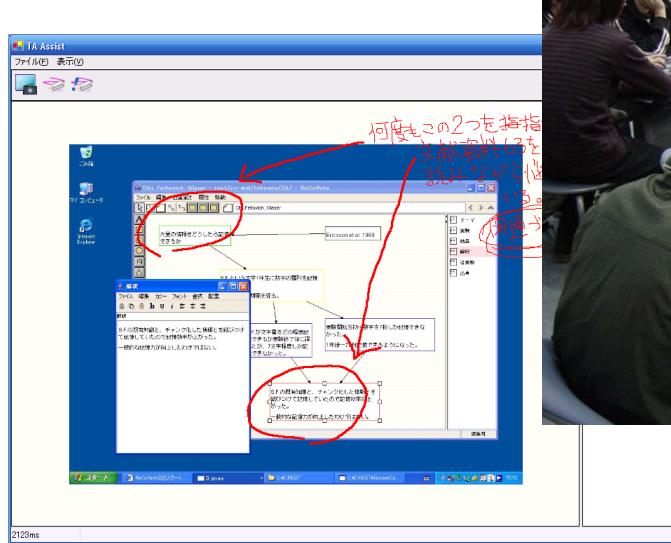
その他の既存データ

- 成績データベース、出席記録
- LMS（学習管理システム）のログデータ
- 研究目的で収集していないものは注意が必要（研究倫理）



収集データの例

協調学習場面における観察者の解釈を電子的に記録・共有する支援ツール (ROG)



平成11年度CREST「高度メディア社会の生活情報技術」研究領域 三宅なほみ 「高度メディア社会のための協調的学習支援システム」

- 高度メディア社会は、**斬新な教育方法**を必要としており、同時にそれを可能にする技術的な基盤を持っている。本研究はその技術的な基盤を有効に利用して、情報社会の実現に必要な**高度な知力を育成する新しい教育方法を、協調的な学習理論に基づいて提案**する。

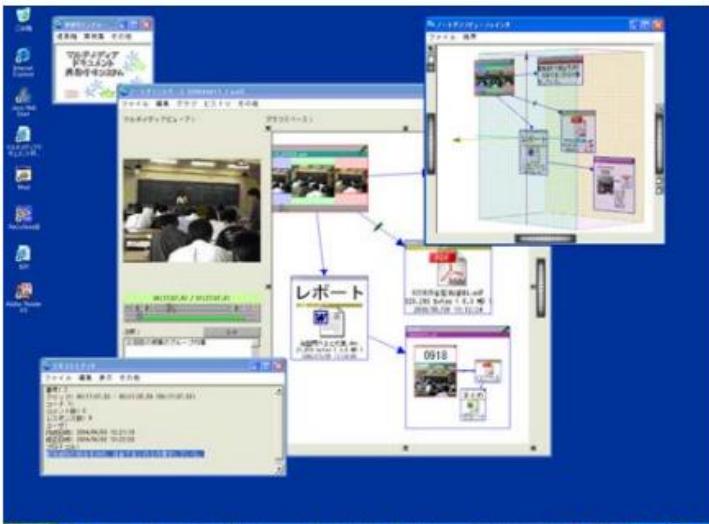


図 3-5 マルチメディアドキュメント共有吟味システム

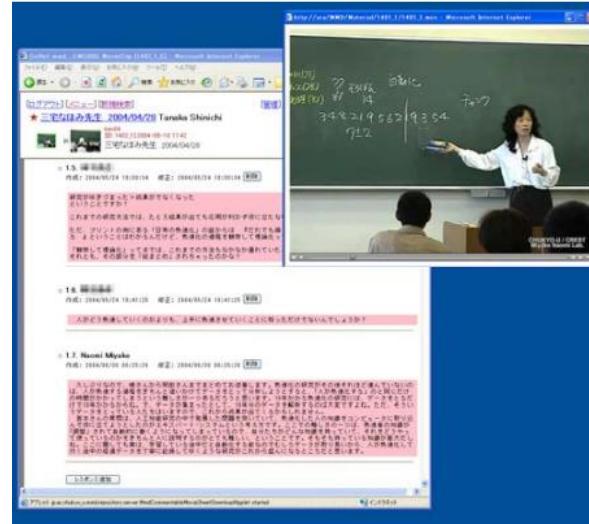


図 3-7 CMSonBBS

	2000年度入学生	2001年度入学生	2002年度生	2003年度生	2004年度生
2000年前期	認知科学入門				
2000年後期	認知科学研究法 I				
2001年前期	認知科学研究法 II	認知科学入門			
2001年後期	認知科学2・応用統計学	認知科学研究法 I			
2002年前期		認知科学研究法 II	認知科学入門		
2002年後期		認知科学研究法 I	認知科学研究法 I		
2003年春期			認知科学中級	認知科学入門 A・B	
2003年秋期			認知科学研究法 I		
2004年春期			認知科学上級	認知科学初級 A・B	
2004年秋期			認知科学研究法 II (応用統計学)		
				認知科学中級	認知科学入門 A・B
				認知科学研究法 I	
				認知科学上級	
				認知科学研究法 II (応用統計学)	認知科学初級 A・B

図 3-15 年度別入学生が受講した授業
(各授業名が週 1 コマ 90 分単位の1授業を表す)

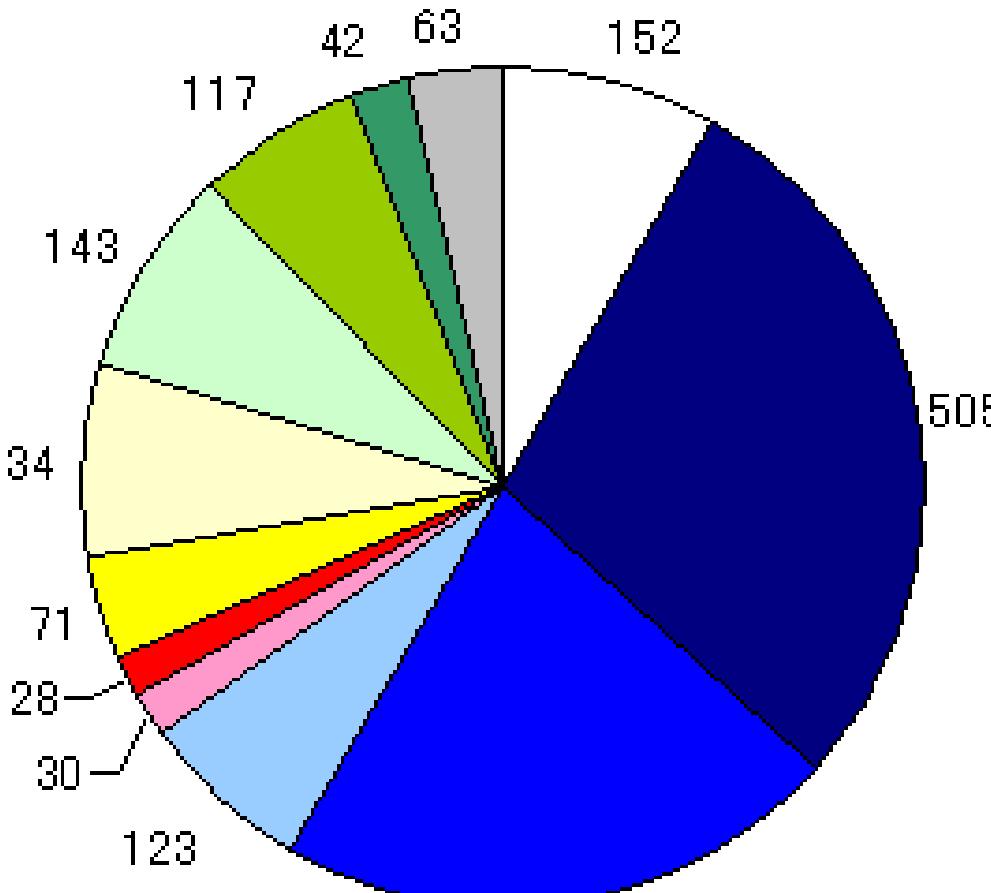


図 3-14a ワークノート例

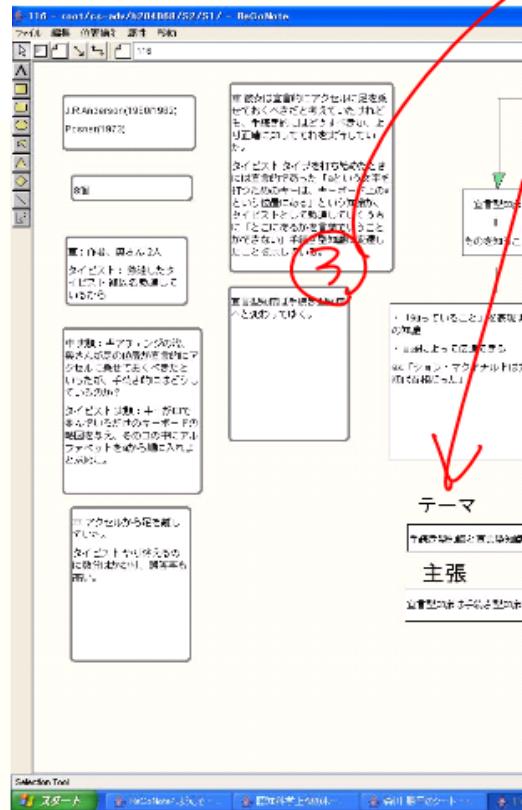
情報	内容
教案 (wordデータ)	教員が事前に用意している講義案、時間配分・目標・話す内容などで構成
配布物 (wordデータ)	学生に配布する資料、作業用紙、講義内容や資料、作業スペース等で構成
観察メモ (紙面)	講義中に教員・TAが作成する記録、スケジュール変更、学生の学習状態の観察、出席者数、行わせた作業の様子、次回教案作成時に考慮すべき情報が手書き
学生ノート (PDF)	講義中に学生が記入したノート、現在は上記の配布物にノート(メモ)部分を用意しておき、これに記入することを学生に要請している
ReCoNote利用履歴 (テキストデータ)	講義中に利用した学習システム履歴、学生の作業履歴を収集が可能
音声データ (mp3など)	講義中の教員・学生の発話データ、教員はICレコーダを常備し、解説などを記録している。学生同士の話し合いも録音
映像データ (QTなど)	講義風景の映像データ、教室の後ろから教壇を中心に講義風景をビデオカメラにて記録

図 3-14b ワークノート・フォルダ

メモの項目別カテゴリわけ



せりふ



- 授業について指示・流れ
- 観察された事実(対象1名)
- 観察された事実(対象2以上のグループ)
- 観察された事実(教室全体)
- 観察内容に対する解釈、予測
- 観察内容に対する評価
- 観察内容などから考えられた仮説
- 観察内容に対する疑問
- コメント・感想
- TA作業
- ROGに対する感想・コメント

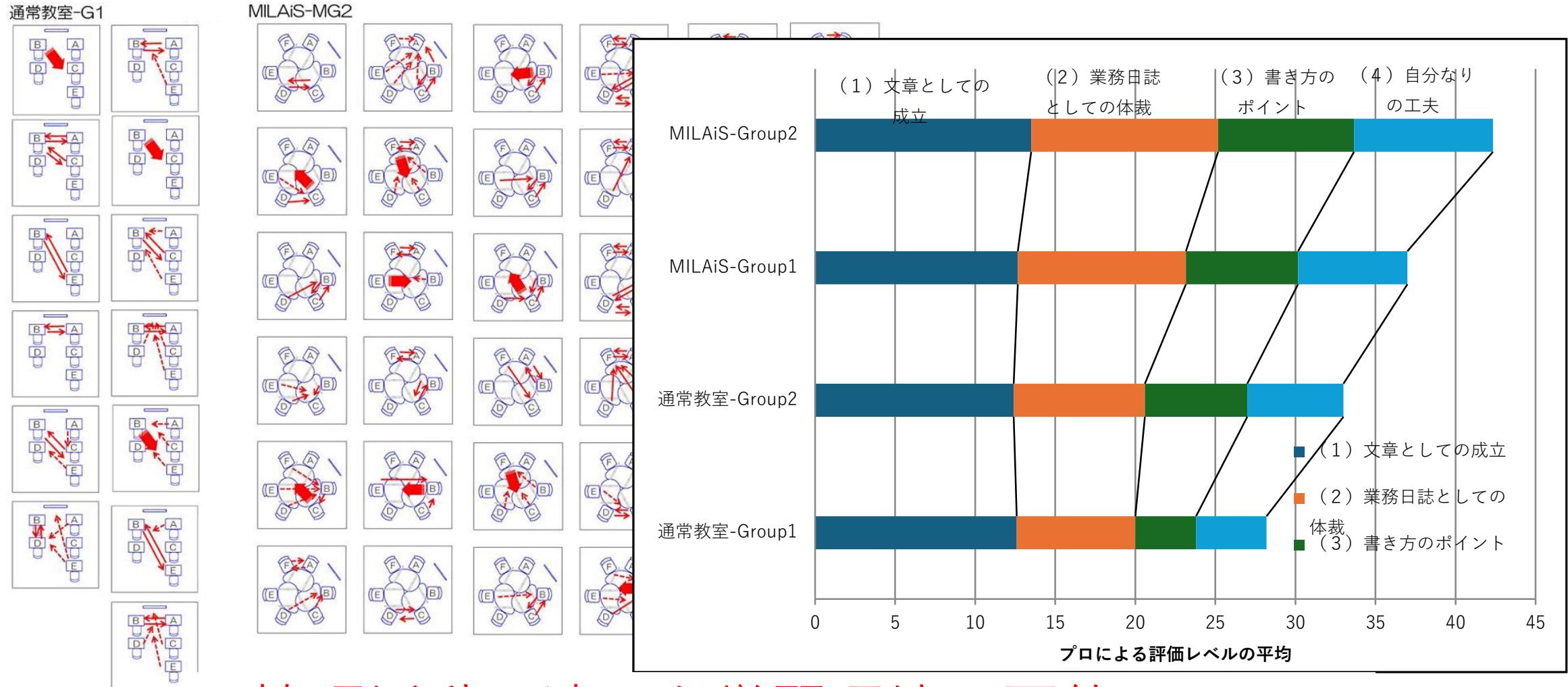
→ 学習過程を詳細に分析・解釈した経験から学生の状態を把握し、最適な学習を作る形成的評価の実現



学習活動のデータの例

ものづくり人材. 高度エンジニア教育にむけた
インタラクティブな学習環境および
協調学習を組み込んだ教育実践とその効果



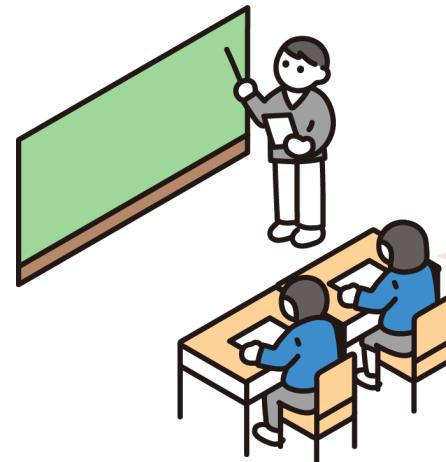
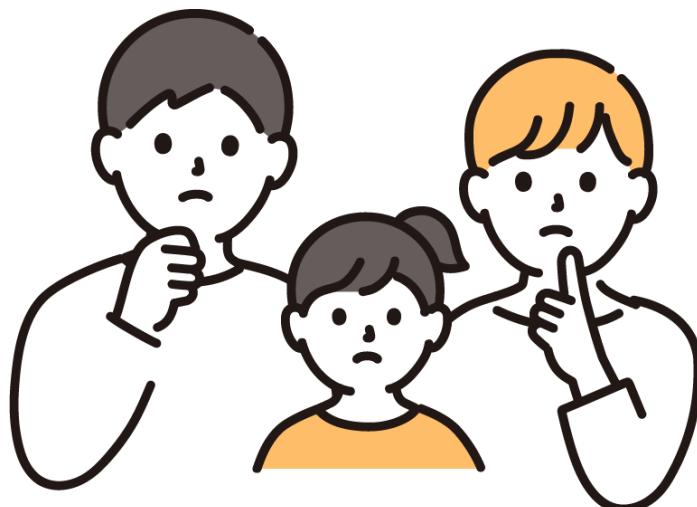


→協調活動に適した学習環境の要件
アイデアや知識の外化→共有→再吟味の
活動設計プロセスの詳細分析

データの信頼性と妥当性

- 再テスト信頼性、一貫性、評定者間信頼性
- 内容的妥当性、基準関連妥当性、構成概念妥当性

学習者の状況や過程を捉えるために適切なデータであるか



まとめ

- データの種類（デジタル・アナログ、量的・質的）
- 尺度は便利だが適切に使うこと
- データの収集方法は目的と用途に応じて選択する
- 信頼性と妥当性、倫理的配慮
- 良いデータ収集のポイント
 - ✓ 研究目的に合った方法を選ぶ
 - ✓ 複数の方法を組み合わせる
 - ✓ 限界を認識し、解釈に反映させる
 - ✓ 実現可能性（時間・コスト・人員）を考慮

引用文献

- 総務省統計局, レベル別テキスト,
<https://www.stat.go.jp/dss/getting/textbook.html>
- 総務省統計局,なるほど統計学園,
<https://www.stat.go.jp/naruhodo/index.html>
- ソコスト, <https://soco-st.com/>