「セキュアな環境における生成AIの校務利用の実証研究事業」
（Ⅳ）生成AIと校務支援システムのダッシュボード機能等との連携に関する開発の実施

事業報告書（概要版）

株式会社内田洋行

1 事業概要

事業概要と目的

子供の学びの充実や教職員の負担軽減に向けて、生成 AI の活用に伴うリスクへの対応を考慮した先進的な取組を多面的に進め、好事例や留意事項の整理・発信等を通じて教育現場での効果的・適切な生成 AI の利活用の実現を支援する必要がある。

教育現場における生成AIの利活用の可能性をさらに拡大するため、**①個人情報を含む重要情報を安全に取り扱うためのセキュアな環境の実現**とともに、**②繁忙期を含む1年間を通しての校務業務における生成AIの活用を促進する機能の開発・実装**を行う。

実施方法

各工程の対応内容は以下の通り。

実施工程	対応内容
① 要件定義	・試作版開発にあたって、学校現場や教職員の考えに即した回答を生成 AI が生成できるようにするため、現状の教職員の業務の流れや考え方、観点、順序等についてのヒアリングを実施 ・ヒアリング結果を受けて、設計・調整を実施
② 有識者ヒアリング	生成 AI を活用する環境及び試作版の機能の考え方について、セキュリティの観点で、課題や留意点についてヒアリングを実施。
③ 設計	要件定義および有識者ヒアリングの結果を受けて、試作版の開発を実施。
④ 開発・テスト	設計に基づいて、試作版の開発を実施。開発後、生成 AI の回答精度について、調整・テストを実施。
⑤ 効果確認	開発した試作版について、要件定義を実施した団体にデモ・ヒアリングを実施

2 実証環境

UCHIDA

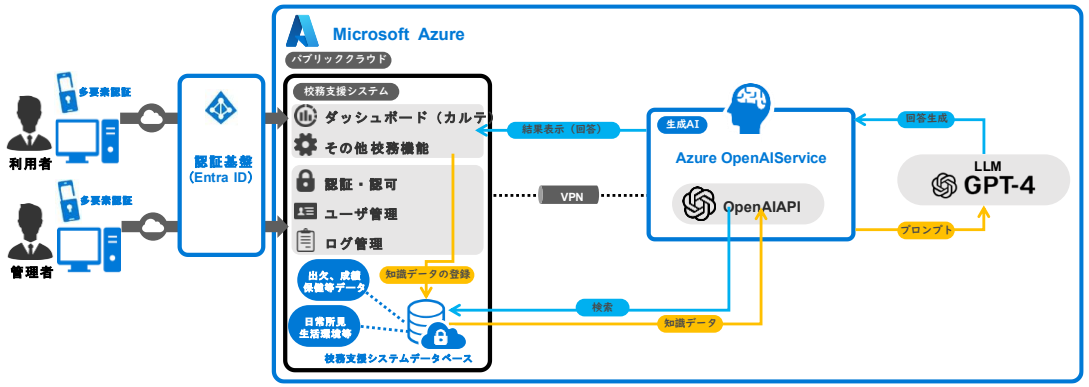
セキュリティ要件

セキュリティ要件について、有識者と意見交換の上、下記のとおり整理した。

	ネットワーク	アクセス権限	仮名処理	トレーサビリティ	その他
必須	・校務支援システムと生成 AI 間は閉域接続とする ・校務系データがインターネット上を経由しない	・校務支援ス テムの既存権限 設定に準拠した 制限 ・業務上必要な 範囲へのデータアクセスの制限 ・生成AI の回答 をそのまま操作者が閲覧できない設計とする	・操作者が個人を特定可能 な情報を入力しない・児童生徒識別コードを利用し仮名化処理する・識別コードは操作者が意識・閲覧できない設計とする	・生成 AI 機能を利用した事実をログで特定可能であること ・利用画 面単位での操 作履歴を管理すること	・生成AIがデータを学習利用しない
任意	・閉域接続の技術の方式選択 ・同一テナント内での構成	・操作者の権限に応じて、生成AI の機能のふるまいを変えるなどの細かな権限制御	・テキスト入力のガイドライン整備 ・学校向け利用ポリシーの文書化 ・自動マスキング・入力チェック機能の実装	・生成 AI の出力内容の完全保存 ・生成 AI の複数回答のうち、どの回答 を選択したか の記録	

システム構成

本実証事業におけるシステム構成は下記の通りとした。



▲ 実証事業における校務支援システム・生成AI環境

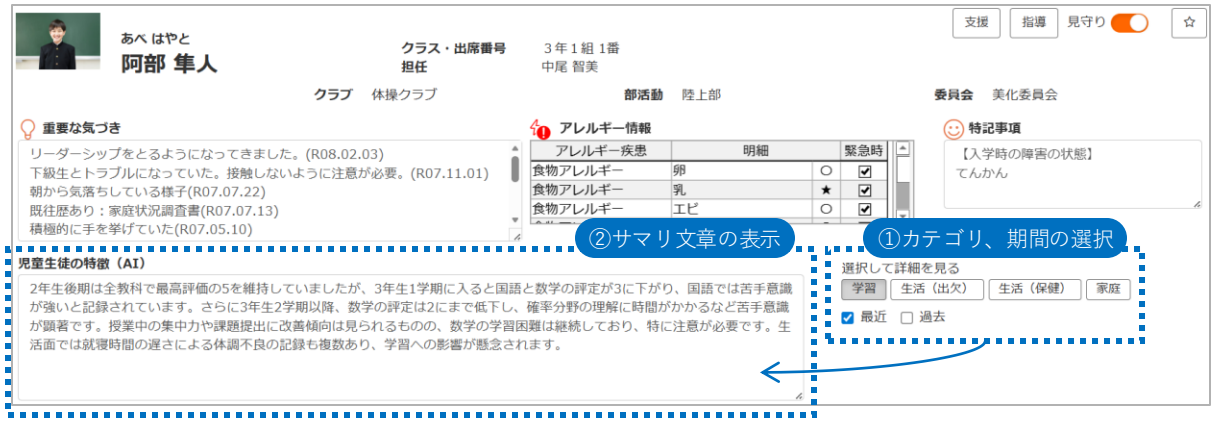
3 開発した試作版

機能一覧

機能	業務	業務スケジュール
ダッシュボード（カルテ）のサマリ文章表示	生徒指導、引き継ぎ	随時
通知表の文章案の作成、チェック	通知表所見	各学期末
学級編成の提案、チェック	クラス編成	年度末、年度初め

ダッシュボード（カルテ）のサマリ文章表示

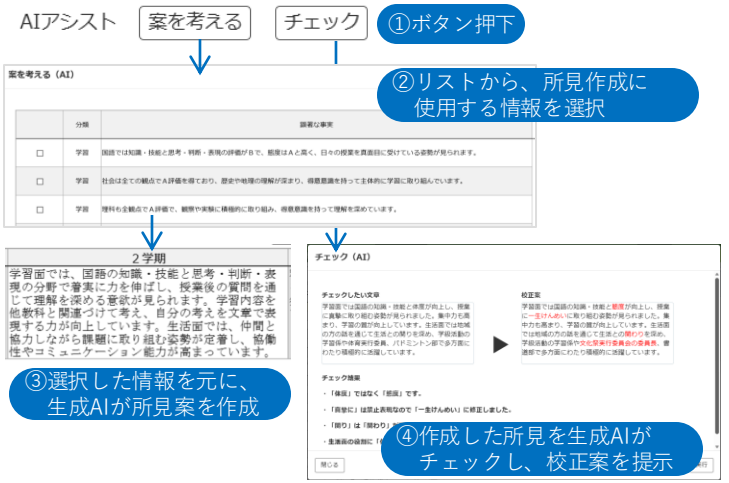
校務支援システムの知識データをもとに、生成AIが児童生徒の特徴をサマリ化する。支援が必要な児童生徒への対応や職員会議での活用を想定し、現場の声を踏まえて、留意すべき点を優先的に整理する。利用者や利用シーンに応じて、カテゴリや期間を選択して利用できる。（複数選択可）



▲ ダッシュボード（カルテ）のサマリ文章表示　画面イメージ

通知表の文章案の作成、チェック

校務支援システムの知識データをもとに、生成AIが児童生徒の成長情報をリスト化する。教職員はその中から所見に用いる内容を選択し、選択情報をもとに生成AIが通知表の所見案を作成する。また、作成した所見については、誤字脱字や衍字、事実誤認の有無、学校で推奨・禁止されている表現の含有を生成AIが確認し、校正案を提示する※。

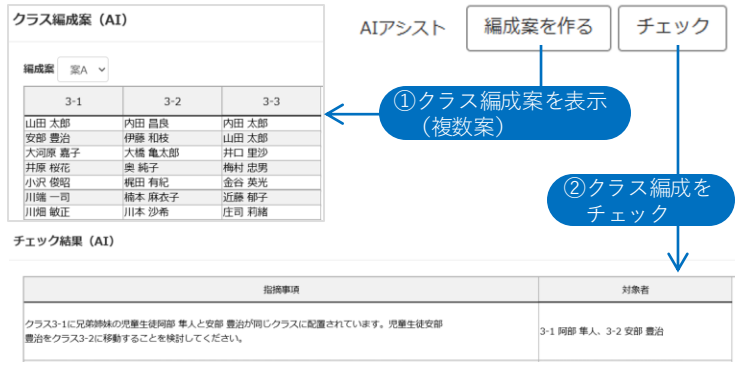


▲ 通知表の文章案の作成、チェック　画面イメージ

※本機能はあくまでも案の提示に留まるため、実運用に際しては、教職員による最終的な判断が必要不可欠である。

学級編成の提案、チェック

校務支援システムの知識データをもとに、生成AIがクラス編成案を複数作成する。実際のクラス分け業務の進め方を踏まえ、クラス編制情報からクラス間のバランスや、同一クラスにしてはいけない組み合わせがないかなどチェック・指摘する。



▲ クラス編成提案、チェック　画面イメージ

4 効果確認①

実施方法

校務支援システム利用ユーザ5団体に対して、要件定義時および開発後の効果確認時に対面ヒアリングを実施した。あらかじめ用意した質問について確認を行い、一部インタビュー形式で、より深堀した意見を徴収した。

問3【デジタイゼーション促進について】
この機能が利用できる場合、積極的にデータ(成績、気づき、特別活動)を入力しようと思いますか
□とても思う
□思う
□あまり思わない
□思わない

問7【感情面・体験面について】
本機能の説明・デモを受けて、生成 AI を業務利用することに抵抗は感じますか
□とても抵抗を感じる
□抵抗を感じる
□あまり抵抗は感じない
□抵抗は感じない

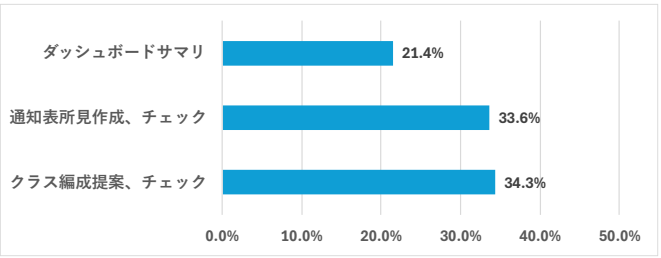
問27【意識・文化の変化について】
生成 AI の活用を通して、自分の仕事のやり方は変わりそうですか？
□とても変わる
□変わる
□あまり変わらない
□変わらない

▲ 効果確認ヒアリング質問の例

時間効率性

生成AIの機能を利用することにより、通知表所見案の作成、クラス編成提案の機能は3割強の業務時間の削減効果が見込まれるという意見だった。具体的な業務がイメージされる機能ほど、効果を実感しやすい結果となった。

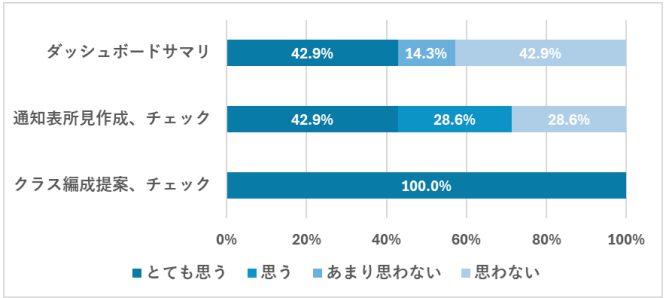
生成AI機能の利用により、
どの程度業務時間が削減されますか？



デジタイゼーションの促進

クラス編成提案と通知表所見案の作成の機能は、8割以上が紙で管理していた情報を積極的にシステムにデータ入力するという結果だった。ダッシュボードサマリについては4割強にとどまり、既存の報告業務との二重管理運用や、サマリ化の過程で重要な情報が見落とされることを心配する意見が挙がった。

この機能を利用できる場合、積極的に
データを入力しようと思いますか？

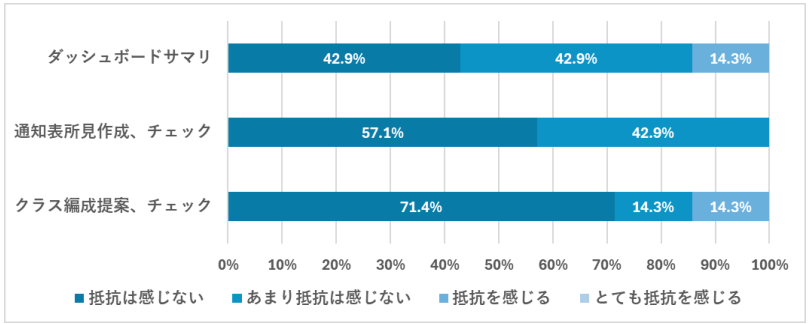


4 効果確認②

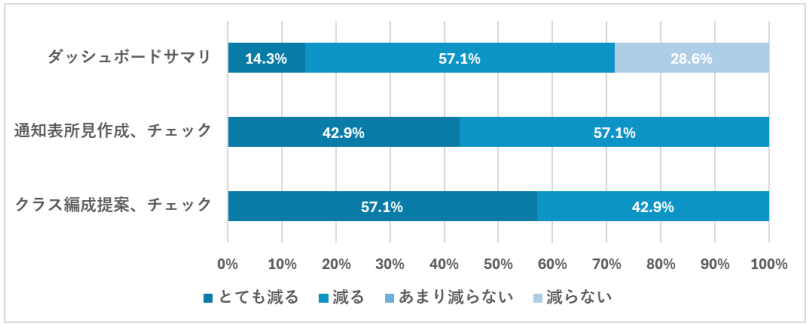
感情面・体験面

生成AI機能に対する感情面・体験面の評価として、①業務利用への抵抗感、②業務ストレス・不安の低減、③業務での利用意向の3点を確認した。
その結果、生成AIの業務利用に対しては8割以上が抵抗を感じないと回答し、高い受容性が確認された。また、通知表所見案作成・チェックやクラス編成提案・チェックなどの機能については、業務ストレスや不安が「減る」との回答が100%となり、心理的負担の軽減効果が評価された。
さらに、業務で利用したいとの回答は全機能で7割以上となり、特に通知表所見案作成・チェック機能の評価が高かった。

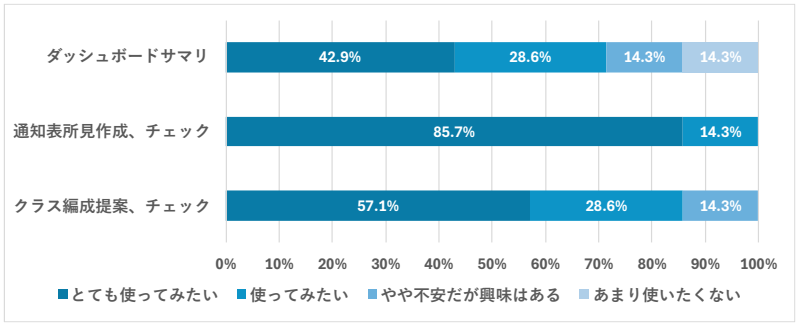
生成AIを業務利用することに抵抗は感じますか？



作業に対するストレスや不安は減りそうですか？



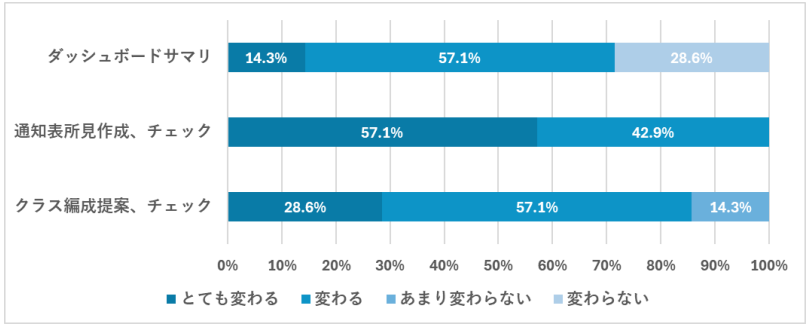
この機能を業務でも使ってみたいですか？



意識・文化の変化

全体的に7割以上が生成AI機能の利用に前向きで、特に通知表所見作成・チェックが高い結果となった。
日常所見の記録を、システムに登録するようになる点や、管理職視点でチェックの効率が良くなる点について前向きな意見が挙げられた。

自身の仕事のやり方は変わりそうですか？



5 回答精度評価①

ダッシュボード（カルテ）のサマリ文章表示

① サマリ文章の記載内容の正確性の確認

概要

特徴の異なる4種類のペルソナを設定し、各ペルソナに対してカテゴリ別のサマリ文章生成を複数回実施した。出力されたサマリを文節単位で事実・解釈・提言の3つに区分した上で、「事実」および「解釈」について、知識データとの整合性を確認した。

事実	1年生の6月以降、頭痛や喉の痛み、咳などの病欠が断続的に見られ、
解釈	加えて、2年生では係活動の引き継ぎが不十分で責任を抱え込みやすい傾向があり、
提言	精神的な負担軽減と健康管理の支援が求められます。

結果

全体の総文節（352件）に対して、区分・整合性は下記の通りとなった。

区分		件数	
事実・解釈	正しい回答	273件	（77.6%）
	誤った回答	9件	（2.6%）
提言		70件	（19.8%）
計		**352件**	

確認された誤りは、病欠回数の誤認、病欠発生時期の誤認の2点だった。

② データ量の違いによる回答への影響確認

概要

成績のカテゴリを対象に、登録データ量を段階的に変化させた8ケースを設定し、各ケースごとにサマリ文章生成を複数回実施した。生成AIが作成したサマリ文章について、回答への影響（データ有無、回答内容）を確認した。

ケース	登録データ量
1	評価評定登録なし、気づき登録なし
2	主要五教科のみ評価評定あり、気づき登録なし
3	全教科評価評定あり、気づき登録なし
4	全教科評価評定あり、気づき 4件
5	全教科評価評定あり、気づき 8件
6	全教科評価評定あり、気づき 16件
7	全教科評価評定あり、気づき 32件
8	全教科評価評定あり、気づき 64件

結果

ケース1及び2においては、サマリ文章にデータが不足（未記入）である旨の回答を返し、ケース3以降は知識データを元にした回答が生成された。回答内容については、ケース3までは評価評定の推移に関する定型的な回答となり、ケース4以降は気づきの情報を反映した記述となった。気づきの件数が少ないほど回答は単調となり、気づきの件数が増えるほど回答は具体且つバリエーションが確認された。

5 回答精度評価②

通知表の文章（所見）案の作成

① 顕著な事実※の正確性の精度確認

概要

特定のペルソナに対して、学年および学期を変更しながら、複数回の顕著な事実の抽出操作を行い、回答の正確性を確認した。抽出結果単位で確認し、知識データとの整合性を確認した。

※「顕著な事実」とは、児童生徒の当該学期および前学期の知識データをもとに、生成AIが成長に関する情報を抽出したもの又はリスト化したものを指す。

結果

全体の顕著な事実（306件）に対して、結果は下記の通りとなった。

	件数	
正しい回答	268件	（87.6%）
誤った回答	38件	（12.4%）
計	**306件**	

誤った回答を分析した結果、主に以下の3類型に整理された。

ⅰ 前学期情報の混入

前学期の内容を今学期の事実として抽出する事象。前学期から評価評定が変化していない場合、継続している事実として誤って引用される傾向が見られた。

ⅱ 異なる評価の記述

C評価をB評価として扱うなど、評価を取り違える事象。複数教科の情報をまとめて抽出する際に、一部該当しない教科を含めてしまう傾向が確認された。

ⅲ 事実と提言の混同

教員が気づきに記載した提言を「事実」として抽出する事象。改善提案を含む記述を、実際の評価内容として扱ってしまう傾向が見られた。

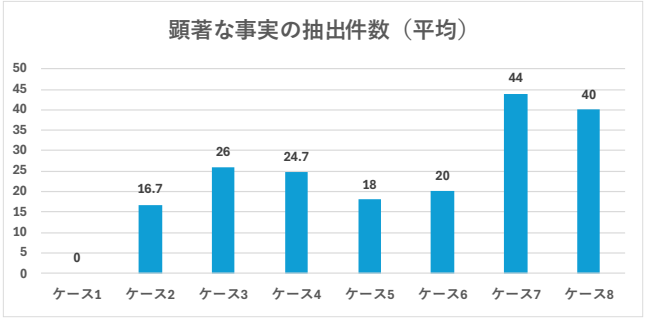
② データ量の違いによる回答への影響確認

概要

特定のペルソナに対して、登録データ量を段階的に変化させた8ケースを設定し、各ケースごとに顕著な事実の抽出操作を複数回実施した。生成AIが抽出した顕著事実について、回答への影響（回答件数）を確認した。

結果

気づきの件数が中程度の範囲では増減は限定的であるが、件数が多く登録されている場合は、抽出される顕著な事実の件数も増える傾向にあった。



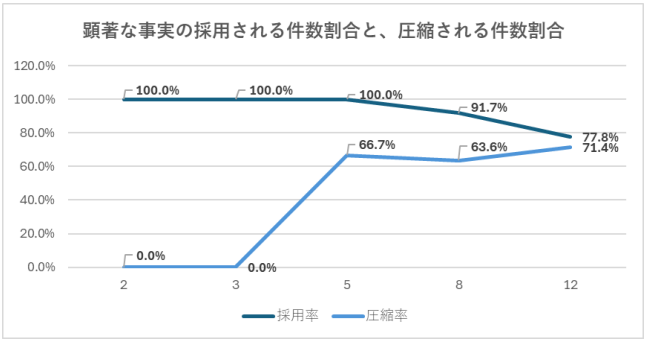
③ 選択する件所な事実に対する、所見案の精度の確認

概要

特定のペルソナに対して、選択する顕著な事実の件数を段階的に増加させ、所見案への反映精度および記載傾向がどのように変化するかを確認した。

結果

顕著な事実の選択が5件以下の場合は、すべて所見案に記載された。一方、8件以上では記載されない内容が生じ、選択件数が増えるほど記載の割合は減少した。また、選択が3件以下では内容がほぼそのまま反映されたが、5件以上になると複数の事実を統合・要約する傾向が見られた。



5 回答精度評価③

通知表の文章（所見）案のチェック

① チェック対象文章に対する指摘・訂正の精度の確認

概要

あらかじめ 1つの所見に4〜5 件の指摘事項（訂正ポイント）を含む所見文章を作成し、生成AI によるチェックを複数回実施した。チェック観点を「誤字・脱字・衍字」「事実との乖離」「学校固有表現の有無」の3項目とした。

結果

各チェック観点の結果は下記の通りとなった。一部、接続詞の誤用を指摘出来なかったケースが確認されたが、概ね指摘・訂正が行われた。

チェック観点	設定件数	指摘件数	正しく訂正した件数
誤字・脱字・衍字	50件	49件 （98%）	49件 （100%）
事実との乖離	40件	40件 （100%）	40件 （100%）
学校固有表現の有無	52件	52件 （100%）	52件 （100%）

学級編成の提案、チェック

① クラス編成の精度確認

概要

あらかじめ105名の児童生徒が所属する学年のデータを用意し、生成AIによるクラス編成の提案を複数回実施した。設定した編成ルールおよび均衡条件を満たしているかを確認した。

結果

検証の結果、すべての回で編成ルールを完全には満たさない結果となった。主な事象として、児童生徒の並び替え・入替処理における削除・追加漏れ、同一人物の重複配置、一部児童生徒の未配置、成績等の均衡条件が満たされないケースが確認された。生成AIは状態（集合）を厳密に保持しながら逐次的な入替や再配置を行う処理を得意としない傾向が見られ、集合の整合性が維持されない事象が確認された。

② クラス編成の精度確認

概要

あらかじめ105名の児童生徒が所属する学年のデータを用意し、編成を行っている状態から、生成AIによるクラス編成チェックを複数回実施した。設定した編成ルールおよび均衡条件をもとに指摘を行うか確認した。

結果

検証の結果、同一の編成状態に対して複数回チェックを行った場合でも、指摘内容にばらつきが見られ、回答の再現性は十分に安定しているとは言い難い結果となった。「同一クラスに配置してはならない児童生徒の組み合わせ」については、すべてのチェックにおいて適切な指摘が行われていた。