

教育データの利活用に関する有識者会議（第9回）

令和3年度

**オンライン学習システムの全国展開，先端技術・教育データの利活用推進事業
（教育データの分析・利活用に係る実証事業）**

～オンライン学習システムを活用した教育データの分析事業～

概要説明

みずほリサーチ&テクノロジーズ株式会社
デジタルコンサルティング部



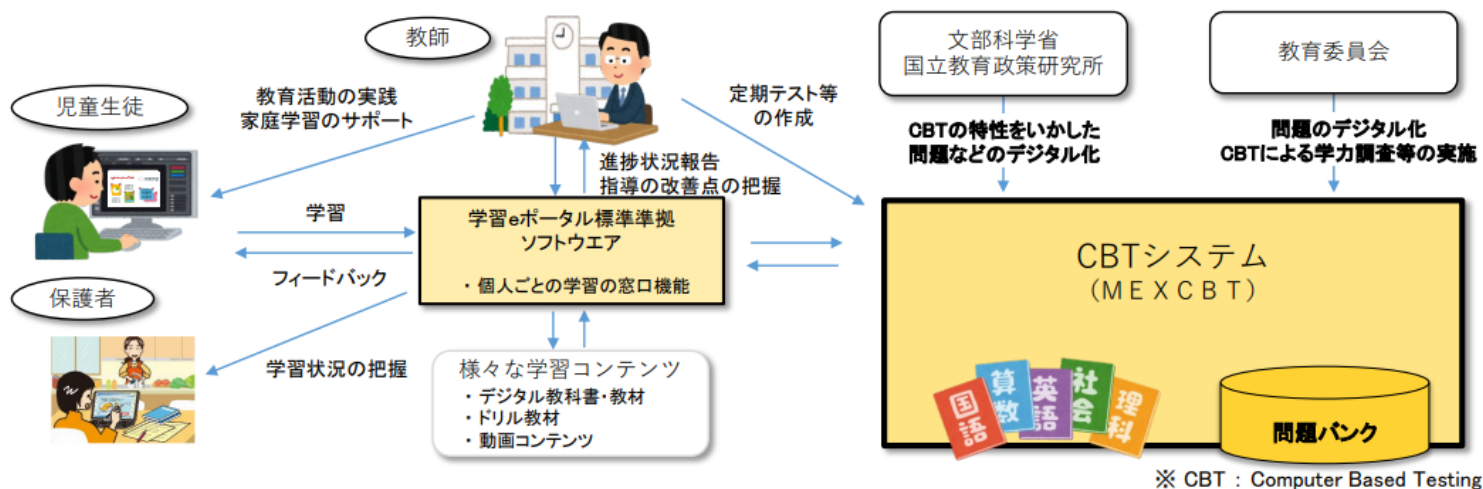
昨年度MEXCBT事業概要のご紹介

MEXCBTとは文部科学省開発のCBTシステムであり、民間事業者提供の学習eポータルを通じアクセス可

文部科学省CBTシステム（MEXCBT：メクビット）について

- 小・中・高等学校等の子供の学びの保障の観点から、**児童生徒が学校や家庭において、学習やアセスメントができるCBTシステム**
- 文部科学省が開発（事業者連合体のコンソーシアムに委託）
- 国や地方自治体等の公的機関等が作成した問題を活用可能
- 「GIGAスクール構想」により実現する「**1人1台端末**」を活用した「**デジタルならではの学び**」を実現

MEXT + CBT
文部科学省 Computer Based Testing



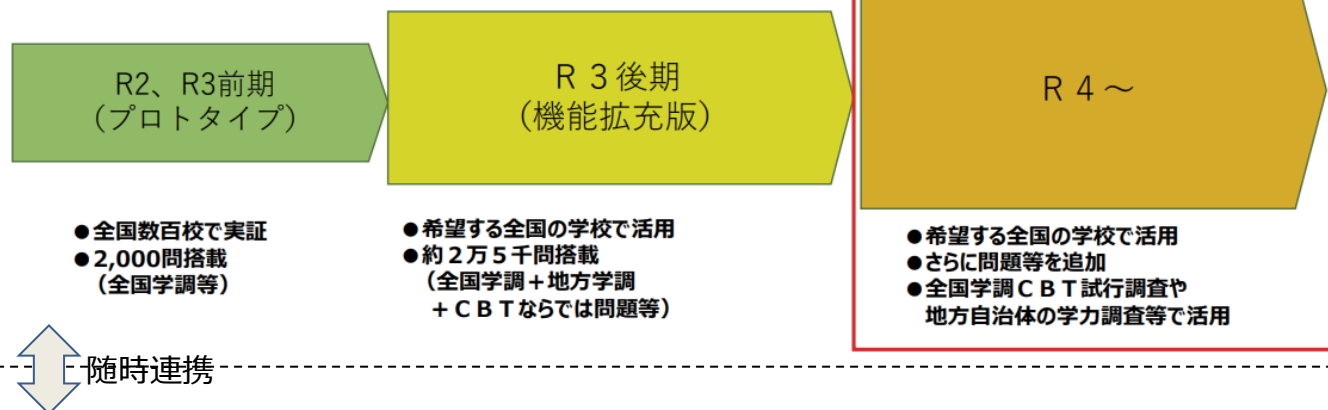
昨年度MEXCBT事業概要のご紹介

事業2年目の昨年度（R3年度）、開発・運用事業と並行しMEXCBTデータの有効性について検討実施

MEXCBTの全体スケジュール

- 令和2年度からプロトタイプ（試行版）を開発
- 公的なC B Tプラットフォームとして、徐々にコンテンツや機能を拡充
- MEXCBT(機能拡充版)については、12月頭から、全国の希望する小中高校等での活用を順次開始
- 令和3年度は、約8500校、約300万人が登録。
- 令和4年度は、利用申込を通年受付。

MEXCBT運用・開発事業



MEXCBTデータ分析事業
※分析対象はR2、R3前期のプロトタイプ
において収集されたデータ

R3後期

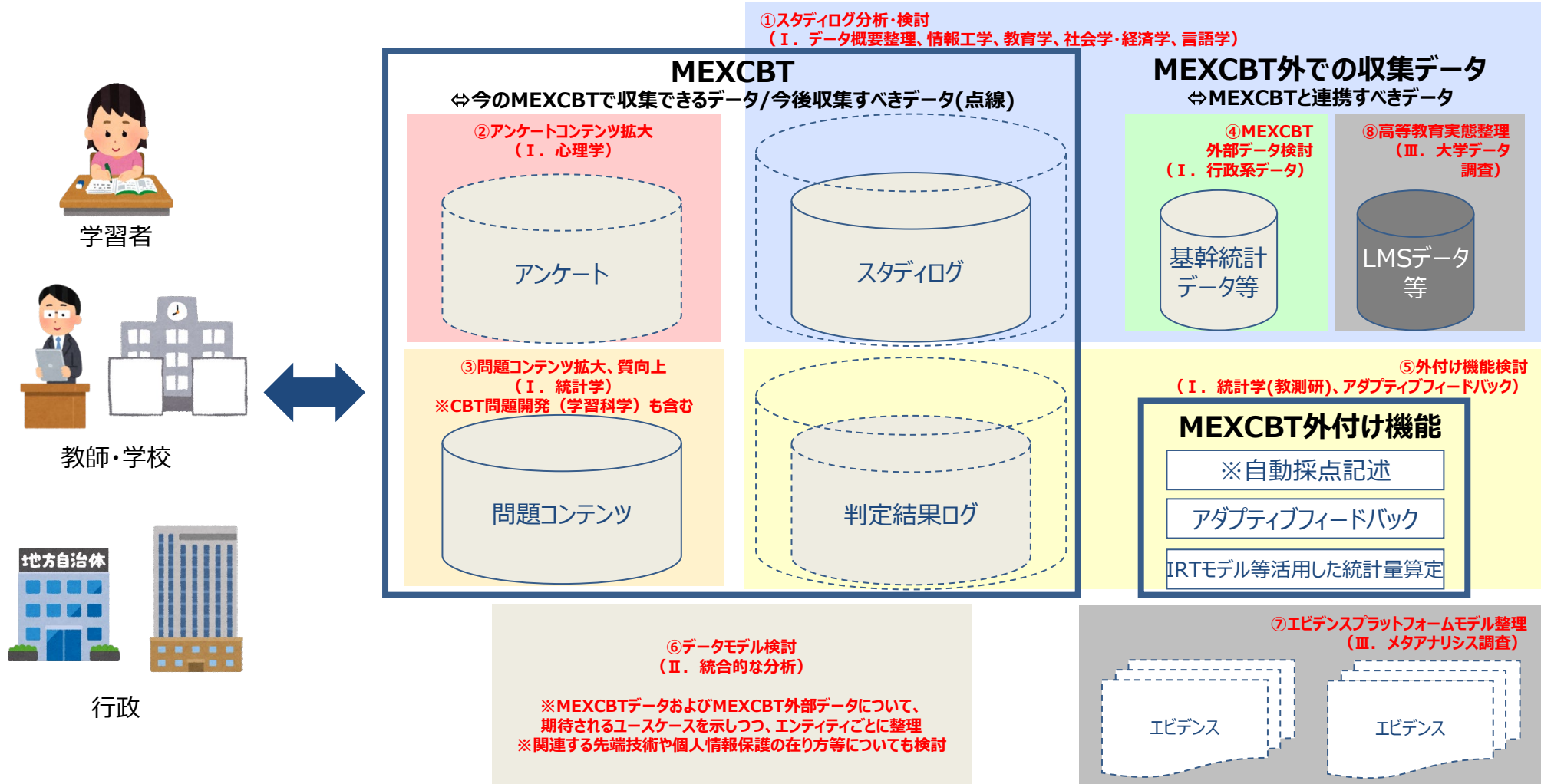
昨年度MEXCBT事業概要のご紹介

**MEXCBT(プロトタイプシステム)は以下のような項目のUUID単位の回答データが取り扱われる仕様
(利用者個人を特定するような属性情報は取り扱ってはいない)**

項目 (UUID単位)
CBTシステムに登録されているテストに対する回答履歴から取得可能な情報 ※R2,R3前期のプロトタイプシステムにおける情報項目
テストタイトル
正解/不正解
回答日時
回答形式 (単一選択、記述式など)
回答内容
ページ遷移順 (利用者がテスト内の各ページを遷移した順番)
ページの表示時間
見直しの有無 (利用者がテスト内の同一ページを複数回閲覧 = 見直しを実施したと判定したもの)
見直し時間 (利用者によるテスト内の各ページの閲覧が2回目以降のレコードの表示時間)
解説参照時間
完答したかどうか (利用者が回答したテストを提出完了したか判定したもの)

昨年度MEXCBT事業概要のご紹介

MEXCBTデータ利活用によってもたらされる便益、および次年度以降のアクションプランについて、期待される事項を多様な学術分野の知見をベースに、主要ステークホルダーの観点、①～⑧の観点から整理



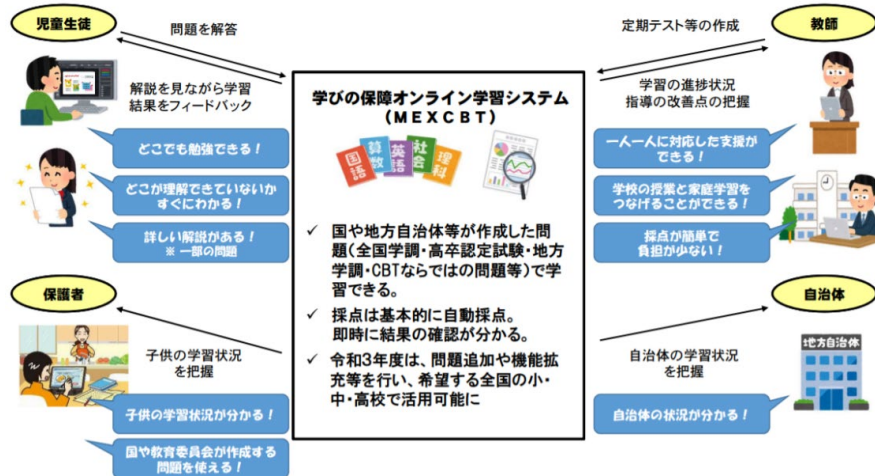
※印の項目については別事業にて取り組まれている

昨年度MEXCBT事業概要のご紹介

公表資料等に記載のある期待される利用シーンも意識し、分析・検討事項の方向性を整理

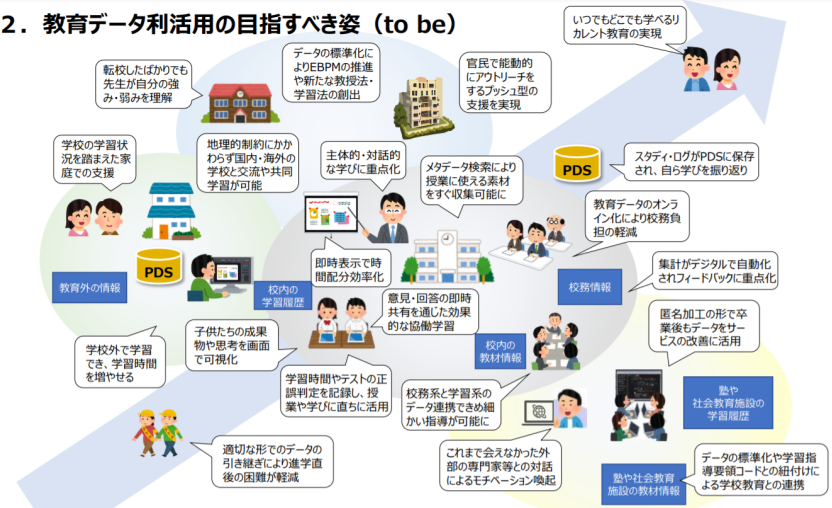
- 対外的に公開されているMEXCBT説明資料及び教育データの利活用シーンをインプットとして、本プロジェクトにおける、多様な観点における検討の具体化を図り、論点の整理と想定される今後MEXCBTに期待されるような実装機能やデータ項目について整理

令和3年7月公表「学びの保障オンライン学習システム (MEXCBT : メクビット)の活用に関する説明会」資料



令和3年1月7日公表「教育データ利活用ロードマップ」

2. 教育データ利活用の目指すべき姿 (to be)



スタディログ等の分析・検討の主な成果

現行のMEXCBTデータによってわかること（分析成果）

	教育工学（ラーニングアナリティクス）	教育学
①MEXCBTデータ分析目的	<ul style="list-style-type: none"> 学習者本人と教員に対する支援を重視し、児童生徒の普段の生活や学習の状況を表すデータを活用することで、児童生徒個々に合わせたフィードバックを行うこと PBTでは得られなかった情報をCBTデータ上で捉えるためのデータ収集と分析を行なっていくこと 	<ul style="list-style-type: none"> 「教育方法学」「教科教育学」を中心に、本事業では「算数・数学」について、教育学・教育現場における課題・通説を整理し、その仮説検証に将来的に資すること <ul style="list-style-type: none"> ✓ 児童生徒の理解度把握 ✓ より正確な問題解決能力の把握 ✓ 中一ギャップ問題 ✓ 単元ごとの習熟難易度 ✓ 単元・教育カリキュラムの見直し
②MEXCBTデータ分析成果（概要）	<ul style="list-style-type: none"> MEXCBT上のスタディログ情報からCBTならではのデータ（解答時間等）分析することによって、フィードバックに有用な児童生徒の解答プロセスを捉えられる可能性を確認 	<ul style="list-style-type: none"> 児童生徒の回答時の動態等を分析することで、理解度や問題解決能力向上の確かな計測の実現に資する可能性を確認 児童生徒の回答結果と各問題の単元情報を紐づけた分析を実施した。「割合」や「単体量当たりの大きさ」といった単元による正答率の傾向を確認できた一方で、これらの要素を求められる問題間においても正答率の差や全体正解率変化に対する影響度に関する差異が確認
③必要とされるデータ粒度やデータ連携の在り方等（抜粋）	<ul style="list-style-type: none"> 教科書会社や教材会社等が検討を進めている学習要素リストレベルの細かい粒度のメタデータを連携させることで、児童生徒がどのような知識をどこまで理解し、習得したか、どこでつまづいているのかをデータから分析することができる 学習eポータル側へMEXCBTスタディログを出力することで、学習eポータル上の様々なデータと紐づけることができ学習者本人と教員に対するフィードバックをより充実させる 	<ul style="list-style-type: none"> 問題ごとの解説参照時間 授業に関する教師・指導方法等に関する情報 単元情報（学習指導要領コードに該当） 単元を超えた類似性の要因となりうるメタデータ 例：資質・能力の三つの柱（「個別の知識・技能」、「思考力・判断力・表現力等」、「学びに向かう力、人間性等」）に関する体系的情報等

スタディログ等の分析・検討の主な成果

MEXCBTへの問題搭載・改善等を通じて今後分かれると期待されること（机上検討結果）

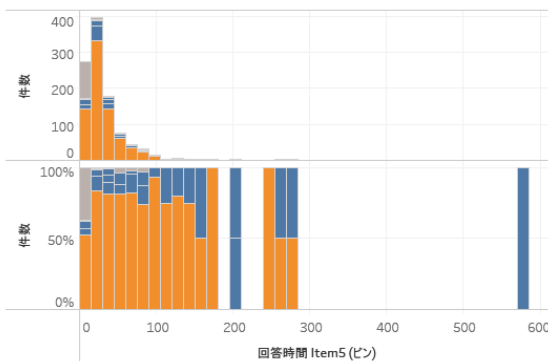
	社会学・経済学	統計学	言語学
①データ分析目的	<ul style="list-style-type: none"> MEXCBTデータのデータ分析を通じ、学習者・教師といった学校現場、あるいは基礎自治体・教育委員会や国といった行政機関にそれぞれのようなメリットあるユースケースが考え得るか、必要なデータ項目は何か、そしてそのデータ整備の在り方等について、経済学・社会学の観点から整理する 	<ul style="list-style-type: none"> 学習者にとって、品質が一定担保されたコンテンツを提供できるような、IRTモデル等統計手法を活用した技術的仕組み構築を目指す 	<ul style="list-style-type: none"> 児童生徒の言語能力の理解に資する分析とその利活用 問題の質向上等のための分析 自然言語処理技術の適用と技術開発のための取り組み
②期待されるデータ分析成果（概要）	<ul style="list-style-type: none"> 回答ログ情報と自治体間の教育施策の特徴などを組み合わせることで、特定の施策の因果関係を精緻に推定し、直接的に成績向上に役立つ知見を蓄積していくことが期待 年に複数回簡単な（共通）調査といった高頻度データの利用を実施することで、年度内での伸びやつまづきの発見につながると期待 wellbeing指標なども普段からタブレットで1週間に1回聞くなどで普段から定点観測を実施 	<ul style="list-style-type: none"> 心理測定学的な指標である「信頼性」「妥当性」の2観点を測定・分析することを想定 確認の手段としては、古典的テスト理論や項目反応理論（IRT）を採用することが有効と史料 	<ul style="list-style-type: none"> 言語能力の育成の状況について、「知識・技能」「思考力・判断力・表現力等」「学びに向かう力・人間力等」といった区分ごとに、MEXCBTデータから分析していく方向性が考えられる ただし、言語能力がデータ分析上どのように定義できるか検討する必要がある
③必要とされるデータ粒度やデータ連携の在り方等（抜粋）	<ul style="list-style-type: none"> 体系的に整理された自治体間の教育施策の特徴情報 年に複数回、できれば毎日実施できるような簡単な（共通）調査より収集される匿名化された学習者情報（学力情報、wellbeing指標等） 義務教育課程～社会に出るまでの個人に関する情報のパネル化 	<ul style="list-style-type: none"> 仕組みの目的意義をどうするかによることもあるが、今回の検討範囲では追加すべき情報項目は特になく見込み <ul style="list-style-type: none"> 基本的には「正解/不正解」に着目する想定、ただしその他項目も分析しながら着目すべき項目がないかを検討できると望ましい またコンテンツの教科は把握する必要があるため、問題データの属性情報に含まれると望ましいが、テストタイトルから推察する方法も一案 	<ul style="list-style-type: none"> MEXCBTで取り扱われる児童生徒の解答はスタディログとしてテキスト形式で保存されるため、作文などの言語的な分析を実績するためのコーパスを作成することは比較的ハードルが低いと思われるものの、句読点の有無など分析しやすいものであるのか言語学の専門家とともに確認しながらデータ整備を行うことが期待される

スタディログ等の分析・検討の主な成果

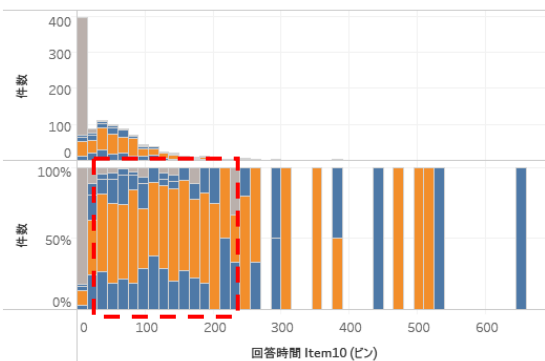
例：教育工学（ラーニングアナリティクス）

- 回答ログを分析することによって従来のPBTでは捉えることのできなかった児童生徒の回答傾向をつかむことができる
- 問題を解くことをすぐに諦めてしまう動きや、時間をかけて正答にたどり着く児童生徒の動きや、問題の傾向を捉えることができる可能性が示唆
 - どのような要因によって回答行動がもたらされているかについては、深掘りが必要

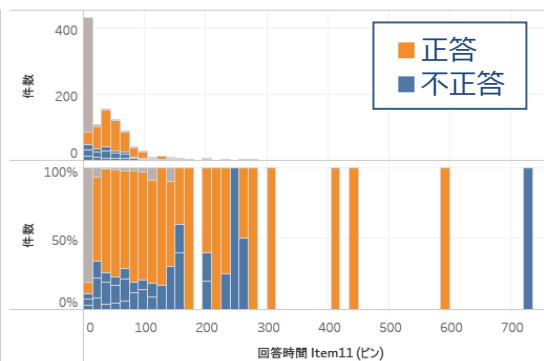
Item05.回答時間_01回答内容



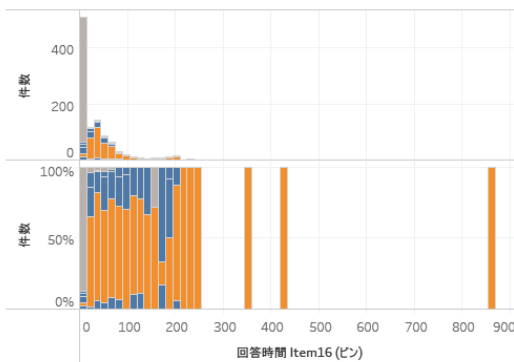
Item10.回答時間_01回答内容



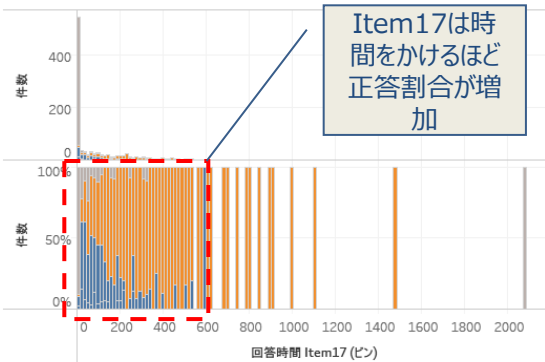
Item11.回答時間_01回答内容



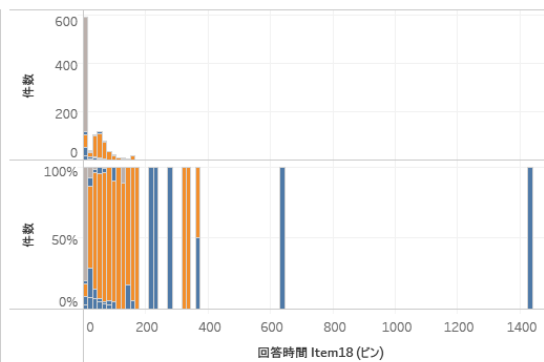
Item16.回答時間_01回答内容



Item17.回答時間_01回答内容



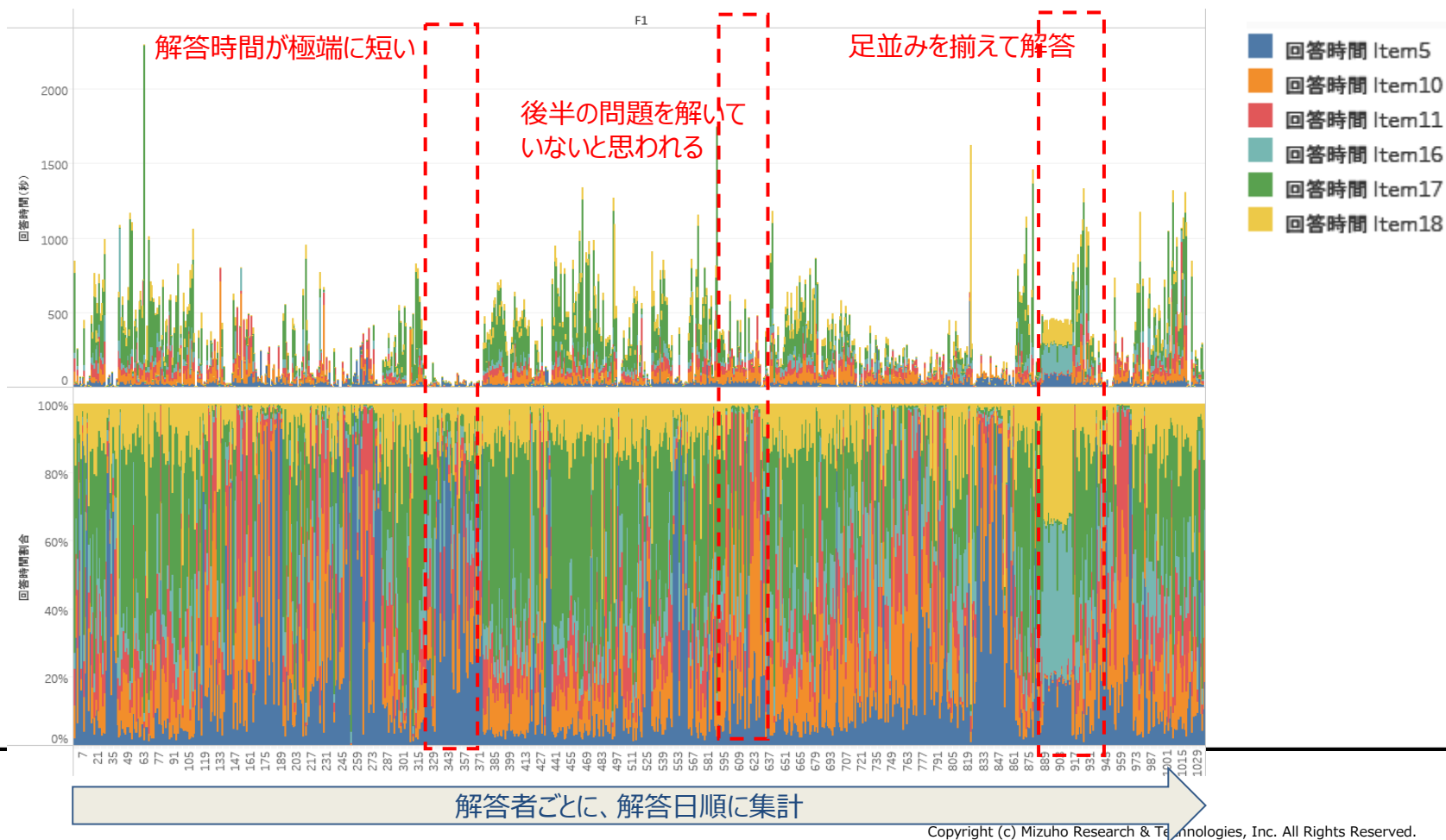
Item18.回答時間_01回答内容



スタディログ等の分析・検討の主な成果

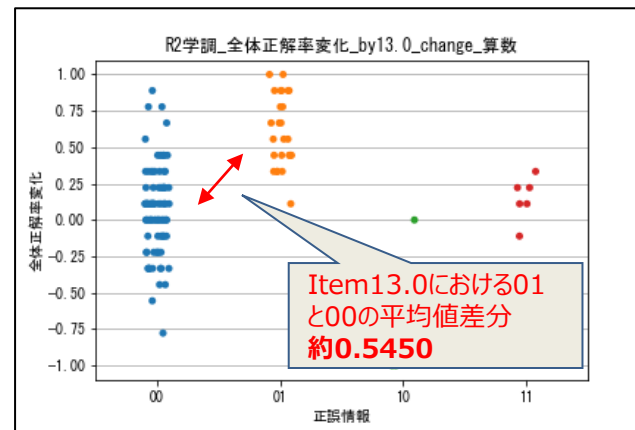
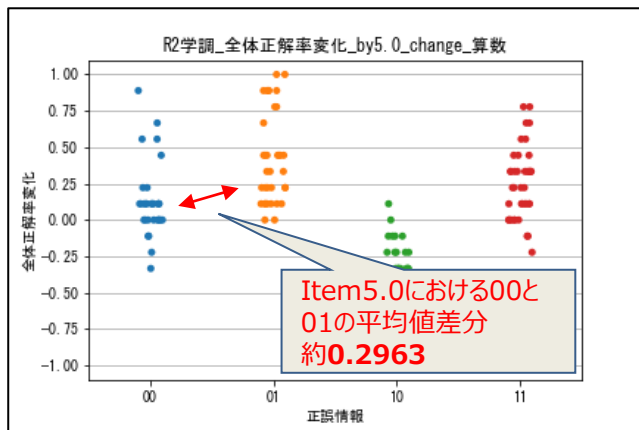
例：教育工学（ラーニングアナリティクス）

- 回答ログのうち解答時間に着目し集計すると、解答時間が短かつ最初の問題のみに取り組んだものや、解答時間がUUID間でほぼ同じであることから授業等で児童生徒が一緒に解いていると思われるデータ等、様々な傾向が把握できることが示唆された
 - 見られた傾向の深堀においては、回答ログごとのMEXCBTの利用シーンの違い等を把握する必要があると史料



例：教育学

- 算数におけるある2単元に関する問題を2回解いた際の正誤変化、および2回解いた際の問題全体の正解率変化の平均値を比較した結果、「 $\times \rightarrow \times$ 」に変化する際と「 $\times \rightarrow \bigcirc$ 」に変化する際の平均値の差分の値が異なった
 - 平均値差分の大きい問題の単元については、解答する上で、その単元上の知識に加え、推論力や論理的思考力の複雑さ等の算数全体として必要なスキルも求められる可能性が示唆できた
 - 解答する上で必要なスキル等の属性情報をシステム側に持たせることができれば、データ分析の成果の更なる積み重ねること、分析成果のさらなる精緻化が期待できる

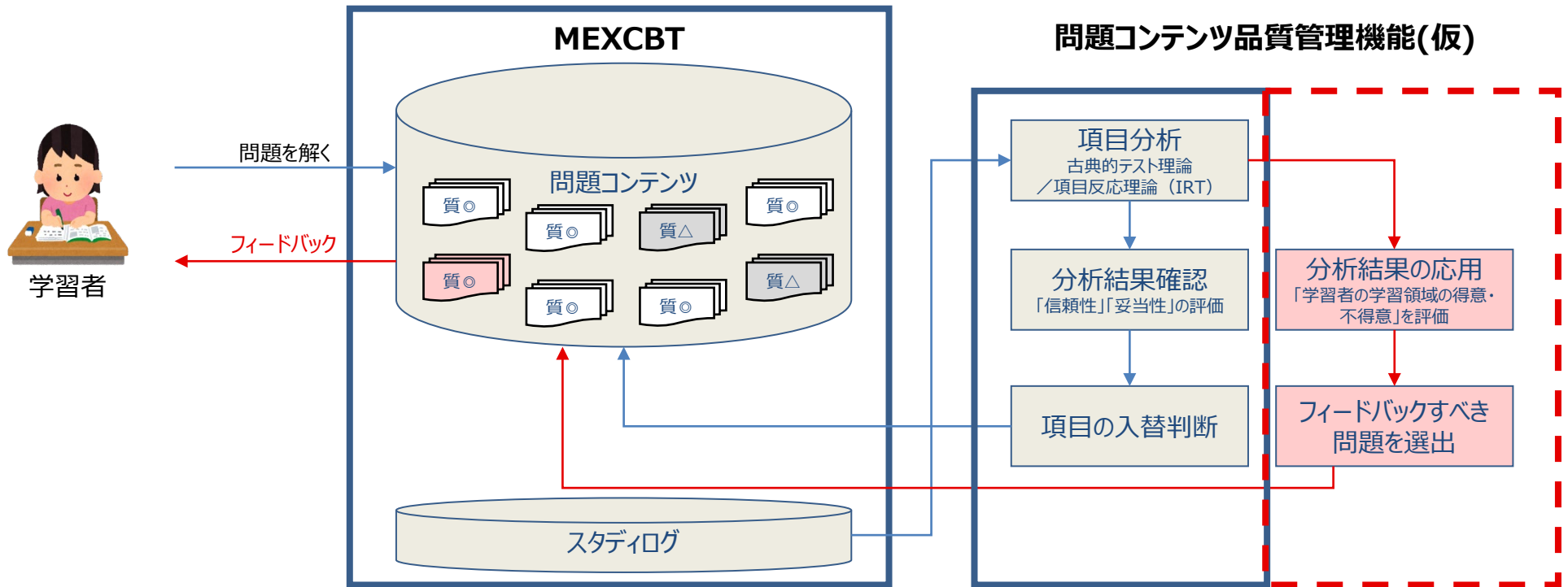


「0」は誤答、「1」は正答を意味する。
従って、「01」は、1回目の解答時には誤答したが、2回目の解答時に正答している生徒を意味する。

スタディログ等の分析・検討の主な成果

例：統計学

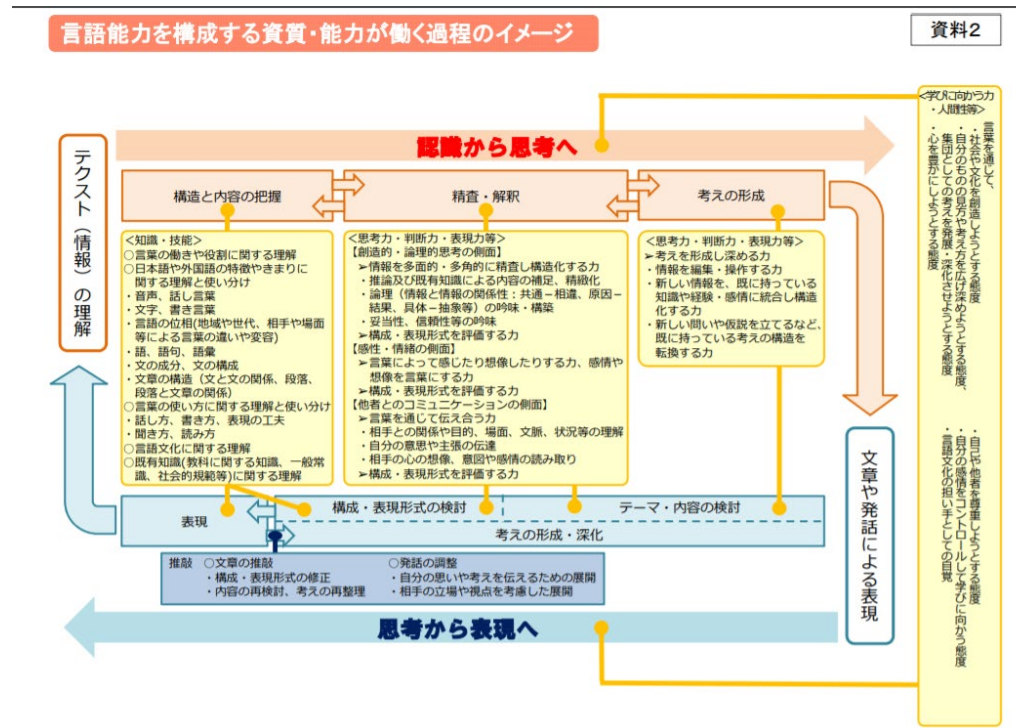
- 統計学上の問題コンテンツの質を「信頼性」「妥当性」と定義し、スタディログの分析を通じ、各コンテンツの質を推定
 - 分析結果確認については、仕組み構築当初は特に、項目分析結果を確認する複数の専門家等の検討の場があることが望ましい
 - 項目の入替判断は、MEXCBTの運用上の目的に照らし、品質の変化の見られる項目の入替や修正の判断を行うことを想定
 - 中長期的には、分析結果を応用し、赤字・赤字にあるような学習者への効果的なフィードバックの品質担保等の他の機能拡充の検討の余地があると思料



スタディログ等の分析・検討の主な成果

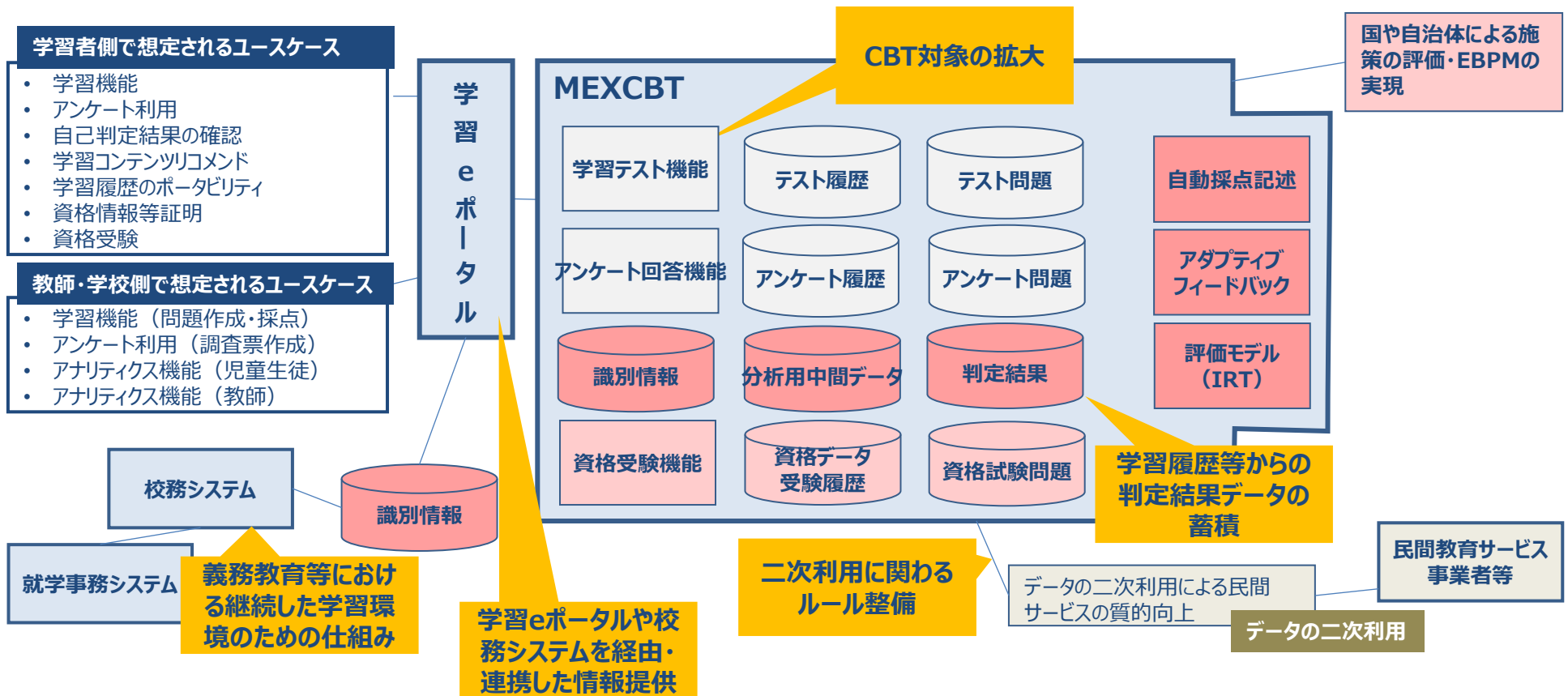
例：言語学

- 公表資料によれば、「言語能力」向上のためには、「テキスト(情報)の理解」と「文章や発話による表現」の2系統に分けられるうち各過程において必要な能力を向上させる必要があると理解
- 言い換えれば、「言語能力」の育成の実態把握においては、以下の示すどこの過程における資質・能力に着目しているかを事前に整理することが重要



データモデル検討の成果

昨年度事業の検討結果をふまえ、ユースケース等からの必要機能やデータ項目追加等を網羅的に整理



成果をふまえ期待される今後の方針

以下2つの方針があり得ると思料

- 昨年度事業成果もインプットに、MEXCBTの機能において改善すべき点について、各種ステークホルダーの意見もふまえた整理が重要
 - 本事業では既存の公開情報等における教育データ利活用への期待もふまえ、MEXCBTデータの有効性について検討を行った
 - 利用者個人を特定するような属性情報は取り扱っていないこともあり、MEXCBTデータのみ利活用によってわかることは限定的ではあるものの、MEXCBT利用者の学習状況や搭載された問題コンテンツの設問別回答傾向の可視化ができることを確認した
 - また、多様な学術分野の知見をふまえた分析を継続的に行うことで、学習状況や設問別回答傾向のモデリングの可能性も示唆された
 - モデリングを実行する上では、「問題コンテンツ」の属性情報や、MEXCBTを利用している際の状況に関する情報（例えば「授業中に一斉的に利用している」のか「家庭学習中に利用している」のか等）等の新規に必要な情報項目があることも整理した
 - また分析成果を用い、MEXCBTに搭載ないコンテンツ（例えば「取扱いのない問題」や「デジタル教材」等）のうち、利用者に最適なものをフィードバックするような学習の保障を実現するには、学習eポータル等のMEXCBT外の仕組みとの連携の在り方を検討することが重要
 - 全国の学習者の学習を保障するためのMEXCBTとして、こういった機能改善が望ましいか、引き続き、各種ステークホルダーの意見もふまえつつ整理していく必要がある
- （上記検討に併せて）机上検討に終わってしまった範囲を中心に、昨年度成果のさらなる深ぼりした検討が重要
 - 昨年度事業において仮説として出たMEXCBTデータが有効な分析方向性のうち、実際に分析作業を行った範囲は限定的
 - よって、上記担うべき役割の検討成果に併せて、特に技術的に担うことが可能な役割は何かを検討するため、机上検討に終わった範囲等の検討のさらなる深ぼりが重要と思料
 - 新規に必要な情報項目のほか、連携が期待されるデータについても言及が必要

成果をふまえ期待される今後の方針

普及展開イメージでの対象範囲からの拡張も想定しつつ、現時点での教育範囲の質的向上に資するものについて優先順位が高いものとして整理（案）

- これまで公表されてきたMEXCBTの普及展開イメージにおける機能拡張のステップ感を認識しつつ、より詳細な検討・実装工程に整理するもの、教育データ利活用ロードマップに示されるような将来的なデータ利活用イメージもふまえつつ、MEXCBTがどのような役割を担うか、定義とともに検討が必要なものを整理し、より前段については機能拡張の優先順位の高いものとして整理することが一案。
- また、長期の利用や教育の質的向上をはかるためにMEXCBTとその他教育系システム等との連携（公的なもの、あるいは民間の教育サービスも含む）のあり方について、関連ステークホルダーの意見も徴収しながら検討すべき事項について優先的に整理を進めることが期待される。

	MEXCBTが現在対象としている初等教育の質的向上、あるいは参加者増加に資するもの（およそR4以降のスコープに入っているもの）		R4以降に記載されている大学等教育の場あるいはそれ以外の場での活用でも期待されるもの	
	CBT由来の機能	CBTから拡張される機能	CBT由来の機能	CBTから拡張される機能
1. MEXCBT現機能、データ活用の高度化	直近実装したほうがよいと考えられるが、自己学習の伸長、教師等の積極的な支援、分析機能等優先順の整理要	学習eポータルを通じた他システムの連携のあり方、運用ルールについて検討、整備	学調等のCBT化や初等、高等教育での学習対象が拡大されてからの実装みきわめ	優先順位高
2. 保障する教育対象の拡大	テスト形式の多様化等が想定され費用対便益を考慮しつつ対象範囲と実装工程を検討	CBTの役割の認知・理解が進んだところで拡張するものとしてその施策や実装工程を検討	初等・高等教育での対象範囲を踏まえ、専門性の強い学習コンテンツのどこまでをCBT化するか検討	学習教育のどこまでをMEXCBTがカバーするか定義しつつ施策や実装工程を検討
3. 長期の利用、利用者の連続性を担保	教育の質的向上、長期の利用に備え、CBTで蓄積した学習履歴等をどう保持・活用可能とするためのデータポータビリティのあり方を検討			