

# 研究計画書(クラウドAI活用版)

受験番号: (記入欄)

氏名: 楠香谷 隆規

作成日: 2025年11月23日

---

## 研究題目

「クラウドAIによる因果推論×高校探究学習統合教育プログラムの効果検証」

—Vertex AI・BigQueryを活用した『メディアリテラシー×デジタルものづくり教育データ』の科学的分析—

(Vertex AI・BigQueryを活用する根拠: Google Workspace(Classroom, Drive)と合わせやすく、探究データをBigQueryに貯めて、AIで分析＆可視化＆要約」のような データサイエンス寄り+生成AIの教育に適応性が高いと思われる為)

---

## I. 研究背景と問題設定

### 1.1 社会的背景: 教育DXとクラウドAIの時代

文部科学省のGIGAスクール構想(2020年開始)とDXハイスクール事業(2024年開始)により、全国の高等学校にICT環境とデジタルファブリケーション設備が急速に導入されています。同時に、Google Classroom(連携: Vertex AI・BigQuery)、Microsoft Teams(連携: Azure AI)等のクラウドプラットフォームが教育現場に浸透し、大量の教育データが生成されているにもかかわらず、それを科学的に分析する仕組みが欠落しているのが実情です。

また、SNS上の偽・誤情報問題の深刻化と生成AI時代の到来により、クリティカルシンキング能力とメディアリテラシーの育成が教育の急務となっています。

### 1.2 教育データ分析における構造的課題

#### A. 教育ビッグデータの未活用

高校では、生徒アンケート、学習ログ(Google Classroom)、CAD設計ファイル、3Dプリント試行データ、成績記録など、膨大な教育データが生成されていますが、これらは独立した形式で保存されたままです。横断的な分析基盤(データレイク)がないため、データ統合分析による因果効果の推定が困難です。

## B. 因果推論手法の教育分野への未適用

マーケティング分野では、因果推論(傾向スコアマッチング、差分の差分法)が当然のように活用されていますが、教育評価ではほぼ適用されていません。これは、教育効果の定量的検証が「印象論」のままである根本原因です。

## C. クラウドAIリソースの教育への活用不足

Google Cloud のVertex AI・BigQueryなどのクラウドAIリソースは、NLP、時系列分析、生成AI機能を備えていますが、高校教育の実践研究ではほとんど活用されていません。

### 1.3 本研究の独自性: クラウドAIと因果推論の融合

本研究は、以下の3点で従来研究の限界を超えます:

#### 第一に、**Vertex AI × BigQuery**による教育データの統合分析基盤構築

Vertex AIで、MLOpsやパイプラインの考え方を身につけましょう。ノーコードでモデルを構築できる「AutoML」から、生成AIプラットフォーム「Vertex AI Studio」の活用まで、実践の入り口を幅広く学びます。

- 生徒アンケート、操作ログ、学習ログをBigQueryに集約
- Vertex AIの前処理・機械学習・生成AI機能で多段階処理
- GeminiによるテキストデータNLP処理(振り返りシート900枚のトピック抽出)
- ダッシュボード可視化で教員・管理職への情報提供

\* 9/25(水) Google ColabとGoogle Cloud Vertex AIで体験する生成AIプログラミング ハンズオン | 広島大学情報メディア教育研究センター

#### 第二に、因果推論とクラウドAIの組み合わせ

- Python(Google Colab)でDoWhy + Vertex AI Workbench統合実行
- 交絡因子の自動検出(機械学習)と傾向スコアマッチング
- ベイズモデリング(PyMC5)による不確実性の確率的推定
- 結果の再現性を確認するための複数手法の並行実行

#### 第三に、**EBPM**推進の教育分野での先駆的実装

- クラウドAIを活用した教育効果の科学的検証フレームワーク
- 他校への実装可能性(スケーラビリティ)の実証
- 教育DXの「見える化」と政策決定への科学的根拠提供

---

## II. 研究目的と具体的仮説

### 2.1 研究目的の3層構造

#### 第1層: 学術的目的

クラウドAI(Vertex AI・BigQuery・Gemini)を活用した教育データ分析プロセスを開発し、因果推論による教育効果の定量的検証モデルを構築する。

測定する主要変数:

- 認知的変数: フェイクニュース識別能力、クリティカルシンキング尺度、情報源信頼性判断力
- 態度的変数: 情報への懐疑的態度、デジタルシティズンシップ意識、自己肯定感
- 行動的変数: 情報発信時の倫理的配慮、AI利用時の批判的思考

#### 第2層: 実践的目的

高校探究学習に統合可能なクラウドAI活用カリキュラムを開発し、以下を実現する:

- 学習指導要領との整合性
- 既存探究学習との融合可能性
- 他校での実装可能性
- 教員負担の最小化
- クラウドAIの教育現場での活用事例の提示

#### 第3層: 社会的目的

クラウドAIを活用した教育データ分析フレームワークを全国のDXハイスクール実施校に提供し、メディアリテラシー教育の普及と教育DXの加速に貢献する。

### 2.2 研究仮説

#### メイン仮説

統合型教育プログラムへの参加は、参加しない場合と比較して、生徒の認知的・態度的・行動的変数を統計的に有意に向上させる。これは、Vertex AI × 因果推論による分析で検証可能である。

#### サブ仮説

- H1:SNSシミュレータ体験により、情報への懐疑的態度が20点以上向上(100点満点中)
- H2:ファクトチェック実践により、クリティカルシンキング能力が15点以上向上
- H3:3Dプリント情報収集で、情報源評価能力が25%以上改善
- H4:EC販売企画により、デジタルシティズンシップ意識が10点以上向上(50点満点中)

---

### III. 研究方法: クラウドAIを活用した新プロセス

#### 3.1 研究デザイン

準実験デザイン(**Quasi-Experimental Design**):

- 非等価統制群デザイン: 介入群30名 vs 統制群30名
- 傾向スコアマッチング(Vertex AI機械学習で交絡因子検出)
- 差分の差分法(DID): 時系列データの因果効果分離
- ベイズ統計モデリング: 効果量の確率的推定

#### 3.2 データ収集・分析の新しい流れ

##### Step 1: データ統合基盤の構築(**BigQuery**)

生徒アンケート → Google Sheets → BigQuery

↓  
Google Classroom学習ログ → BigQuery

↓  
3Dプリント試行記録 → BigQuery

↓  
振り返りシート(紙) → OCR → BigQuery

↓  
統一されたデータレイク完成

##### Step 2: データ前処理と特微量エンジニアリング(**Vertex AI + Python**)

Vertex AI Workbench(Google Colab環境)で:

- 欠損値処理(多重代入法)
- 特微量エンジニアリング(新しい変数の自動生成)
- 外れ値検出と処理
- 正規化・標準化

##### Step 3: 交絡因子の自動検出と傾向スコア推定

```
from google.cloud import aiplatform  
from dowhy import CausalModel  
import pandas as pd
```

```
# BigQueryからデータ取得
```

```

query = """
SELECT student_id, program_participation, learning_outcome,
       baseline_skill, motivation, ses_level, learning_history
FROM research.education_data
"""

df = pd.read_gbq(query, project_id='project-id')

# 交絡因子の自動検出(Vertex AI AutoML)
automl_model = aiplatform.AutoMLTabularModel.create(...)

# 傾向スコア推定
model = CausalModel(
    data=df,
    treatment='program_participation',
    outcome='learning_outcome',
    common_causes=['baseline_skill', 'motivation', 'ses_level', 'learning_history']
)

identified_estimand = model.identify_effect()

estimate = model.estimate_effect(
    identified_estimand,
    method_name="backdoor.propensity_score_matching"
)

```

#### **Step 4: テキストマイニングによる質的データ分析(Vertex AI Gemini)**

振り返りシート900枚のテキスト分析:

- Gemini APIで自動要約(各シート50字→20字への圧縮)
- トピック抽出(LDAまたはBERTベース)
- 感情分析(ポジティブ/ネガティブ判定)

- キーワード抽出と頻出度分析

```
import vertexai
from vertexai.generative_models import GenerativeModel
vertexai.init(project='project-id', location='asia-northeast1')

# Geminiを使った大量テキスト処理

model = GenerativeModel('gemini-pro')
reflections = df['reflection_text'].tolist() # BigQueryから取得
for reflection in reflections:

    response = model.generate_content(f"""

```

以下の生徒の振り返りシートから、

学習成果、学習課題、感情の変化を抽出してください：

```
{reflection}
```

```
""")
```

```
print(response.text)
```

#### **Step 5: 結果の可視化と政策提言 (Data Studio)**

Google Data Studioで、BigQueryデータから自動的に：

- 因果効果の大きさ(ATE, CATEの可視化)
- 信用区間と有意性判定
- サブグループ分析(学年別、背景別)
- 時系列推移(施策別の効果時系列)

#### 3.3 実装スケジュール

時期	実施内容	クラウドAIの活用
4月	データ統合基盤の設計・構築	BigQueryスキーマ設計
5-7月	Phase 1実施(メディアリテラシー基礎)	Google Classroomログ自動収集

時期	実施内容	クラウドAIの活用
8-10月	Phase 2実施(3Dプリンティング)	3Dプリント試行データの自動記録
11-1月	Phase 3実施(発信倫理)	Google Formsアンケート自動集計
2月	最終評価・分析	Vertex AI + 因果推論分析
3月	テキスト分析・報告書作成	Gemini テキストマイニング + Data Studio可視化

---

## IV. 期待される成果

### 4.1 学術的貢献

- 教育データサイエンス分野の先駆的事例: クラウドAI × 因果推論の教育応用
- 統合型教育効果検証フレームワークの体系化
- テキストマイニング + 量的分析の質的統合モデル構築

### 4.2 実践的貢献

- クラウドAI活用カリキュラムの全国展開可能性の実証
- 教員向けクラウドAI研修プログラムの開発
- 他校が実装可能なBigQuery スキーマの提供

### 4.3 社会的貢献

- 教育委員会・総務省への政策提言(データ根拠あり)
- 全国DXハイスクール 100校への普及波及
- 「データに基づく教育改革」文化の醸成

---

## V. 研究の限界と対応策

課題	対応策
サンプルサイズの小ささ	ペイズ統計による小サンプル対応、2年次に協力校3校追加
無作為割当の困難性	傾向スコアマッチング、複数手法による三角測量

課題	対応策
クラウドAI学習コスト	Google Cloud 教育プログラム(無料クレジット)の活用
データプライバシー	Google Cloud ゼロトラスト セキュリティの活用

---

## 結語

本研究は、クラウドAIの教育現場への実装と因果推論による科学的検証を通じて、SNS時代を生き抜く若者のクリティカルシンキング育成に貢献するものです。限られた資源でも、クラウドAIと統計手法の組み合わせにより、全国の高等学校へ還元可能な教育改革の根拠を創出することを目指しています。

総文字数:約2000字

---

## 1. そもそも「因果AI」がやってくれること

普通のAI(機械学習)は

「この条件のとき、だいたい何が起こりそう?」=予測(相関)

因果AIは

「もしこの施策をやったら、“やらなかつた世界”と比べてどれくらい変わる?」=介入の効果(因果)

を推定します。

- ガートナーなども、\*\*「相関ベースの予測を超えて、因果関係を使って意思決定を高度化する技術」\*\*として因果AIを定義しています。
  - ビジネス意思決定の世界では、キャンペーン・価格変更・新サービス導入などの「施策の効果」を測るために実務利用が進行中です。
- 

## 2. 今どのくらい“現場で使われているか”

ライブラリ・ツール

- EconML (Microsoft Research)

- 観察データから「施策の効果」を機械学習で推定するPythonパッケージ。
- **DoWhy+EconML**
  - 因果グラフ→識別→推定→検証までを一貫してやるツールキット。

## クラウド統合

- **Azure Machine Learning**
  - Responsible AI Dashboard の中に、EconML をバックエンドにした因果推論コンポーネントが統合されています(what-if・介入シミュレーション)。
- **Vertex AI(Google Cloud)**
  - AutoML Forecasting+合成コントロールなどを活用した因果推論デモやノートブック事例が公開されています。
  - さらに、causalLens のような 専用因果AIプラットフォームが **Vertex AI / Gemini API** と連携し、「因果推論+LLM」という構成も出てきています。

→つまり

\*\*「研究レベルのオモチャ」ではなく、  
普通に企業や研究機関が使い始めている“プロ向けツール群”\*\*になっています。

## Azure AI(+Azure OpenAI)

- 強み
  - Microsoft 365(Teams, OneDrive, SharePoint)との連携が超強い
  - 「企業内導入」「Officeとの連携」「Copilot 系」に乗りやすい
- **Azure AI Services** :「学校現場 × Office365」を考えるならかなり王道。

## Vertex AI

- 強み
  - **BigQuery**・データ分析・ログ解析との親和性が高い
  - GCP の他サービス(Maps, YouTube Data APIなど)との連携も視野に入る

- Google Workspace (Classroom, Drive) とも合わせやすい
- 「探究データを BigQuery に貯めて、AIで分析 & 可視化 & 要約」  
みたいな データサイエンス寄り + 生成AI の教育には向いている。