

デジタル庁教育関連データ連携 実施報告書

2024 年 3 月 31 日

株式会社サイバーリンクス

1. はじめに

本報告書は、デジタル庁が主導する教育関連データ連携実証事業において、本年度実施した名寄せ作業の概要と成果についてまとめたものです。名寄せ作業とは、学習 e ポータルと校務支援システムで管理されている児童生徒情報において、同一児童生徒を識別するための手順を指します。名寄せ作業にて、学習 e ポータルが保持する UUID の校務支援システムへの書き戻しにより、システム間識別用 ID 統合による年度更新の簡便化、教育関連データの活用や分析の効率化や精度向上が期待されます。

2. 実施するシステム環境

本実証事業では、以下のような 2 自治体の環境で実施した。

- A 自治体
学習 e ポータルと校務支援システムの名簿情報は連携せずそれぞれ登録している環境
※小規模自治体にて、小学校 1 校・中学校 1 校のデータを利用
- B 自治体
学習 e ポータルと校務支援システムの名簿情報は連携しており、独自の外部連携用のキーによる紐づけが行われている環境
※大規模自治体にて、約小学校 180 校・中学校 70 校のデータを利用

3. 名寄せ手順

本実証事業では、以下のような名寄せ手順にて実施した。

- A 自治体の手順
各ファイルレイアウトは以下となります。

○校務支援システム出力レイアウト

| 出力項目 |
|--------|
| 学校cd |
| 児童生徒id |
| 姓ふりがな |
| 名ふりがな |
| 姓 |
| 名 |
| 生年月日 |
| 性別 |
| 学年 |
| 学級 |
| 整理番号 |
| 交流学年 |
| 交流学級 |
| 出席番号 |
| 入学日 |
| 編入学日 |
| 転入学日 |
| 転退学日 |
| 退学日 |

○学習eポータル出力レイアウト（項目と説明）

| 項目名 | 項目説明 |
|---------------------|---------------------|
| pfid | UUID |
| family_name | 姓 |
| given_name | 名 |
| full_name | 氏名 |
| family_name_kana | 姓（かな） |
| given_name_kana | 名（かな） |
| full_name_kana | 氏名（かな） |
| display_name | 表示名 |
| gender | 性別（m/f 又は 男/女） |
| role | 役割（student/teacher） |
| login_name | ユーザーID |
| password | パスワード |
| federation_id | 外部認証ID |
| grade_name | 学年名 |
| homeroom_name | クラス名 |
| number_in_homeroom | 出席番号 |
| graduate | 卒業済み（卒業の場合x） |
| email | メールアドレス |
| note | メモ |
| parent_display_name | 保護者アカウントの表示名 |
| parent_login_name | 保護者アカウントのID |
| parent_password | 保護者アカウントのパスワード |

○校務支援システム取込用レイアウト

| 項目名 | 項目説明 |
|--------|--------------------|
| 児童生徒id | 校務支援システム用キー |
| pfid | 学習eポータルから出力されたUUID |

- I. 校務支援システムにて全校分の児童生徒名簿 CSV ファイル（1 ファイル）をバッチ処理にて出力
- II. 学習 e ポータルにて学校単位の Excel ファイル（2 ファイル）を出力、共有フォルダにてデータ受理
※学習 e ポータル管理業者にて出力

- III. 名寄せマクロ（ver.11.3）にて名寄せの条件等を設定し、名寄せ処理を実行
以下の名寄せ条件にて設定にて実行した。

○パターン 1

| | 項目名① | 項目名② | 項目名③ | 項目名④ | 項目名⑤ |
|---------|------------------|-----------------|------------|---------------|--------------------|
| 校務支援 | 姓ふりがな | 名ふりがな | 学年 | 学級 | 整理番号 |
| 学習eポータル | family_name_kana | given_name_kana | grade_name | homeroom_name | number_in_homeroom |

○パターン 2

| | 項目名① | 項目名② | 項目名③ | 項目名④ | 項目名⑤ |
|---------|-------------|------------|------------|---------------|--------------------|
| 校務支援 | 姓 | 名 | 学年 | 学級 | 整理番号 |
| 学習eポータル | family_name | given_name | grade_name | homeroom_name | number_in_homeroom |

○パターン 3

| | 項目名① | 項目名② | 項目名③ | 項目名④ | 項目名⑤ |
|---------|-------------|------------|------------|---------------|--------------------|
| 校務支援 | 姓ふりがな | 名ふりがな | 交流学年 | 交流学級 | 出席番号 |
| 学習eポータル | family_name | given_name | grade_name | homeroom_name | number_in_homeroom |

○パターン 4

| | 項目名① | 項目名② | 項目名③ | 項目名④ | 項目名⑤ |
|---------|-------------|------------|------------|---------------|--------------------|
| 校務支援 | 姓 | 名 | 交流学年 | 交流学級 | 出席番号 |
| 学習eポータル | family_name | given_name | grade_name | homeroom_name | number_in_homeroom |

○パターン 5

| | 項目名① | 項目名② | 項目名③ | 項目名④ | 項目名⑤ |
|---------|------------|---------------|--------------------|--------|------|
| 校務支援 | 交流学年 | 交流学級 | 出席番号 | 性別 | |
| 学習eポータル | grade_name | homeroom_name | number_in_homeroom | gender | |

○パターン 6

| | 項目名① | 項目名② | 項目名③ | 項目名④ | 項目名⑤ |
|---------|------------------|-----------------|------------|--------|------|
| 校務支援 | 姓ふりがな | 名ふりがな | 学年 | 性別 | |
| 学習eポータル | family_name_kana | given_name_kana | grade_name | gender | |

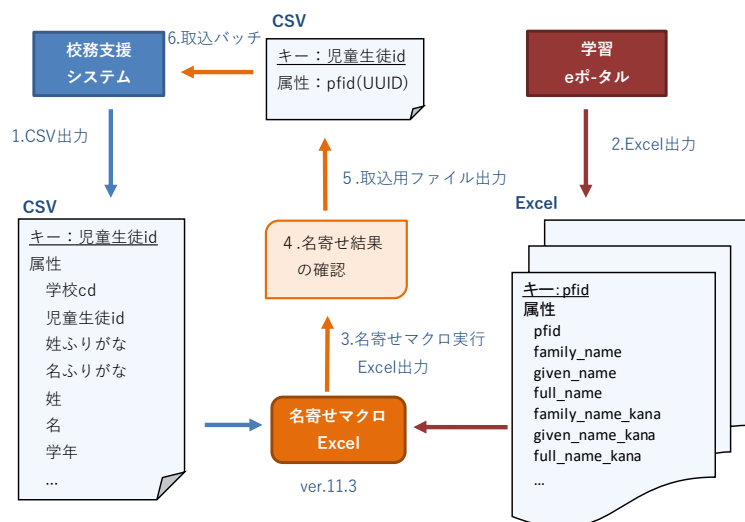
○その他全体条件

| 除去対象データ | 項目名 | 除去したい文字 |
|---------|---------------|---------|
| 学習eポータル | grade_name | 年 |
| 学習eポータル | homeroom_name | 組 |

- IV. 出力された名寄せの結果（名寄せ Output_yyyymmdd_hhmmss.xlsx）を確認
名寄せに成功したデータ（条件一致）や失敗したデータ（校務支援のみ・e ポータルのみ）を確認する。また、名寄せに失敗したデータについては、手動で修正や確認を行う。
- V. 名寄せの結果ファイルより校務支援システムへの取込用ファイルを出力
※今回は提供いただいた「インポート CSV 作成ツール_ver.1.00.xlsm」は利用せず実施。

- VI. 校務支援システムへの取込用ファイルにて校務支援システムに UUID を取込
校務支援システムのキーである児童生徒 id より学習 e ポータルの UUID の取込を実施した。

○名寄せの手順



● B 自治体の手順

各ファイルレイアウトは以下となります。

○校務支援システム出力レイアウト

| 出力項目 |
|--------|
| 外部連携ID |
| 学校cd |
| 児童生徒id |
| 姓ふりがな |
| 名ふりがな |
| 姓 |
| 名 |
| 学年 |
| 学級 |
| 整理番号 |
| 交流学年 |
| 交流学級 |
| 出席番号 |
| 入学日 |
| 編入学日 |
| 転入学日 |
| 転退学日 |
| 退学日 |

○学習eポータル出力レイアウト（項目と説明）

| 項目名 | 項目説明 |
|---------------------|---------------------|
| pfid | UUID |
| family_name | 姓 |
| given_name | 名 |
| full_name | 氏名 |
| family_name_kana | 姓（かな） |
| given_name_kana | 名（かな） |
| full_name_kana | 氏名（かな） |
| display_name | 表示名 |
| gender | 性別（m/f 又は 男/女） |
| role | 役割（student/teacher） |
| login_name | ユーザーID |
| password | パスワード |
| federation_id | 外部認証ID |
| grade_name | 学年名 |
| homeroom_name | クラス名 |
| number_in_homeroom | 出席番号 |
| graduate | 卒業済み（卒業の場合x） |
| email | メールアドレス |
| note | メモ |
| parent_display_name | 保護者アカウントの表示名 |
| parent_login_name | 保護者アカウントのID |
| parent_password | 保護者アカウントのパスワード |

○校務支援システム取込用レイアウト

| 項目名 | 項目説明 |
|--------|--------------------|
| 外部連携ID | 校務支援システムの外部連携キー |
| pfid | 学習eポータルから出力されたUUID |

- I. 校務支援システムにて全校分の児童生徒名簿 CSV ファイル（1 ファイル）をバッチ処理にて出力
- II. 学習 e ポータルにて全校の Excel ファイル（1 ファイル）を出力、共有フォルダにてデータ受理
※学習 e ポータル管理業者にて出力

III. 名寄せマクロ（ver.11.3）にて名寄せの条件等を設定し、名寄せ処理を実行 以下の名寄せ条件にて設定にて実行した。

○パターン1

| | 項目名① | 項目名② | 項目名③ | 項目名④ | 項目名⑤ |
|---------|------------|------|------|------|------|
| 校務支援 | 外部連携ID | | | | |
| 学習eポータル | login_name | | | | |

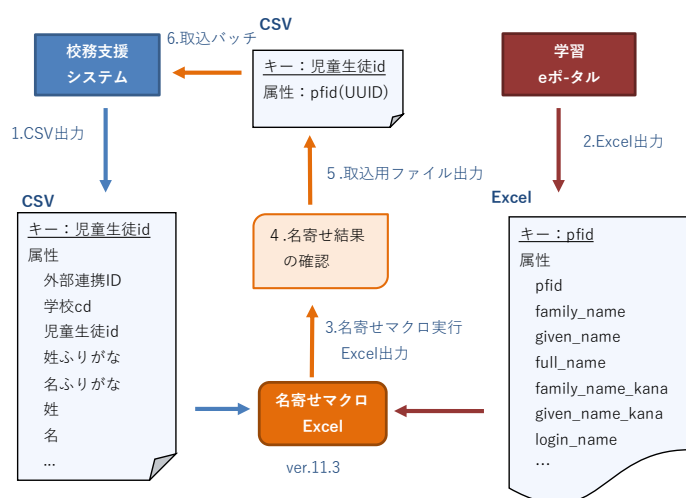
※児童生徒数が10万件以上のため名寄せマクロが負荷のため処理失敗

（以降の想定していた名寄せの手順については、A自治体と同様）

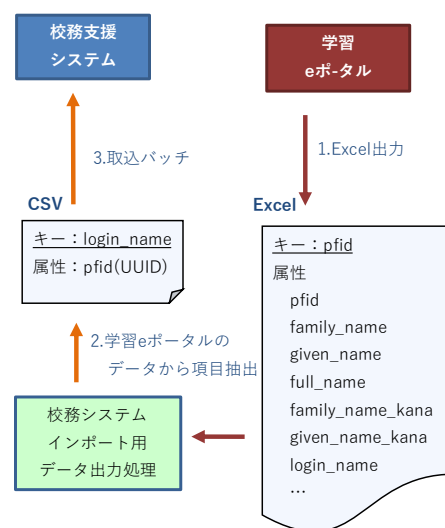
※今回は名寄せマクロが利用できなかったため学習eポータルから出力されたExcelデータより校務支援システム取込用のデータ作成処理を代替え利用

名寄せ利用不可のためのExcelから校務システムインポート用出力.vbs

○想定していた名寄せの手順



○代替え手順（名寄せ未実施）



4. 名寄せ作業の工数

本実証事業では、A自治体の名寄せ作業の工数を計測した。

A自治体：小学校1校（児童数：約200人）、中学校1校（約100人）

| 手順 | 作業内容 | 工数 | 備考 |
|----|---|-----|--|
| 1 | 名寄せ条件の検討 | 30分 | 各データよりどの条件が最適か確認する必要があるため、自治体により大きく工数の変動の可能性がある。 |
| 2 | 名寄せの実行 ※1回1校当たり処理時間は約2分程度。 ※6パターン実施 | 24分 | 名寄せ件数：約300人中292人 ※名寄せ条件を複数実行した結果より最も名寄せ率が高かった場合 ※姓名のふりがなが入っていない学級がある等の条件で大きな差異が発生した。 |

| | | | |
|---|----------------------------------|----|---|
| 3 | 名寄せに失敗したデータの確認 ※1人1分程度 | 8分 | 名寄せの確認作業は教職員が行うほうが工数削減となるがツール利用等を想定した場合業者が実施するため工数が増える可能性がある。 ※教育委員会や学校への確認作業が必要と場合も考えられる。 |
| 4 | 名寄せに失敗したデータの手動での書き換え ※1人30秒程度 | 4分 | 校務支援システム・学習eポータルデータの修正が実際には必要であるが、業者に再出力を依頼する必要がでるなど難しい場合がある。手動で変更した場合はシステムとの差異に注意が必要である。 |

以上の表から、名寄せマクロを用いることで、名寄せ作業の工数は削減できることがわかります。最適な条件を設計することで、約97%のデータが正しく名寄せされました。ただし、名寄せの条件の設計や手動での書き換えについては、運用によって条件や件数に大きな差異が出るため注意が必要である。

5. 名寄せ作業における課題や留意事項

本実証事業では、以下のような名寄せ作業における課題や留意事項を把握した。

● 運用（データ）について

- 名簿情報が校務支援システムと学習eポータルで連携（データインポート）されておらず、個別に登録されている場合、名寄せのための項目に差異が発生しているため条件一致しない場合がある。

今回の名寄せ作業で発生した条件一致しないケースは以下となりました。

| | |
|---|---|
| 1 | 個別に登録されているため登録情報に差異が発生している場合 ※姓名ひらがなの濁点や学級情報の誤り等 |
| 2 | 校務支援システムの姓名に外字が登録されているが学習eポータルでは内字で登録されている場合 ※姓名のふりがなが登録されておらず、漢字で名寄せが必要となるために発生。 ※校務支援システム側には内字の姓名登録箇所はあるが利用されていないために発生。 |
| 3 | 姓などの情報が変更となった場合に両システムで統一されていない場合 ※更新頻度やタイミングによって、名寄せの結果が変わる可能性がある。 |
| 4 | 特別支援学級の児童生徒の学級や出席番号の運用が統一されていない場合 ※交流学級の設定が行われているかどうか。 |
| 5 | 学習eポータルの出力データに学校情報が含まれておらず市内間転出入や進学時の紐づけで転出側や進学前の学校と紐づくリスクがある。 ※実証事業で検証した学習eポータルの場合のため他製品の場合は不明。 |

● 名寄せマクロ（ver11.3）について

- 名寄せマクロの利用や名寄せの条件には、専門的な知識や経験が必要である。名寄せの対象となるデータ項目や条件を適切に選択することで、名寄せの精度や効果を向上させることができる。しかし、名寄せの条件やルールの設定には、抽出された児童生徒名簿データの内容や特性を熟知している必要がある。また、名寄せの設定やルールの設定には、データ提供者や利用者との調整が必要である。
- 名寄せマクロの処理可能件数に課題が残る。小規模自治体であれば学校単位での処理も可能であるが、大規模自治体の場合には運用は難しいため改善の必要がある。

端末、条件にて処理した結果を以下に記載する。

端末性能：Windows11 64bit、Core i7 3.4GHz、メモリ 32GB

◆条件（1セットは入力必須）

| | ID名 |
|---------|------------|
| 校務支援 | 外部連携id |
| 学習eポータル | login_name |

| データ件数 | 処理状況 |
|---------|----------------------------------|
| 120,000 | [No:6]オーバーフローしました。 のメッセージ表示で処理停止 |
| 100,000 | [No:6]オーバーフローしました。 のメッセージ表示で処理停止 |
| 50,000 | [No:6]オーバーフローしました。 のメッセージ表示で処理停止 |
| 20,000 | Excel がフリーズしたため処理未完了。 |
| 5,000 | 処理時間 10 分程度、正常に処理完了。 |

- 名寄せの結果の確認には、人的な判断や作業が必要である。名寄せマクロは、名寄せの結果を自動的に生成するが、その結果が正しいかどうかは、人間が確認する必要がある。特に、名寄せに失敗したデータについては、手動で修正や確認を行う必要がある。名寄せに失敗した「校務支援のみ」「eポータルのみ」データを加工した場合、「条件一致」に移行できる仕組みがあれば名寄せ作業の工数を大幅に削減できると考えられる。
- 名寄せマクロは作業端末で行う必要があるため、個人情報や機密情報の取り扱いに注意が必要である。

● インポート CSV 作成ツール（ver.1.00）について

- 実施した名寄せ作業ではツールを利用することはできませんでした。
ツールは学習 e ポータルの項目のみを出力可能な設定となっており、校務支援システムのキー項目（児童生徒 id や外部連携 id）を対象とすることができなかった。
校務支援システムのキー項目をもとに児童生徒を紐づけし、学習 e ポータルの UUID を取り込む必要がある。今回のツールでは新規登録を対象としているが、校務支援システムへの新規登録には学習 e ポータルでは持っていない他必要項目があるため取込対象とすることは難しい。

6. まとめ

本報告書では、デジタル庁教育関連データ連携実証事業において、本年度実施した名寄せ作業の概要と成果についてまとめました。名寄せ作業では、提供いただいた名寄せ・インポートツールのみでの取込を

行うことはできませんでしたが、名寄せマクロにより効率的に名寄せすることができました。

本実証では以下の2つの課題が大きな問題となりました。

- 名寄せ条件が複雑である。「姓名」「ふりがな」「学籍番号」「性別」などの複数の項目の組み合わせによる判定条件を調査する必要がある。また、学校や学級単位で登録されている項目に差異が発生した。
- 名寄せの件数が多いと処理が完了できないため、複数に分けて処理する必要がある。

これらの課題を解決するための提案を以下に記載する。

- 名寄せ条件を簡素化する。名寄せに必要な項目を絞り込み、標準ルールまたは運用ルールを決定する。例えば「ふりがな」「学年」「生年月日」などを必須項目として校務支援システム・学習eポータルの運用を定めておく。また、簡素化した名寄せで対象外となったデータのみを追加条件で処理する仕組みを検討する。
- 名寄せ作業を定期化する。作業を一定間隔で行うことにより、名簿の整合性を保つことができる。また一定間隔で同じ条件での作業を行うことで自動化の検討が可能である。
- 名寄せに利用しない不要なデータを削除する。名寄せ作業前に項目を削除できる仕組みを用意することで負荷を軽減することができ、一度に処理できる件数を増やすことが可能である。

これらの案にて名寄せ作業の負担軽減となるように具体的な実施方法を引き続き検討する必要があると考えます。