

銀行データを活用するための産学連携研究プラットフォームの構築

小林 学[†] 中原悠太[†] 佐藤絵里子[‡] 高谷親信[‡] 塚本好宏[‡]早稲田大学データ科学センター[†] みずほ銀行デジタルマーケティング部[‡]

1. はじめに

株式会社みずほ銀行（以下みずほ銀行）と早稲田大学は 2018 年 7 月にデータサイエンス活用の裾野拡大に向けた研究・教育に関わる学術交流協定を締結した[1]。本協定の目的はデータサイエンスにおける共同研究やデータサイエンス教育プログラムの共同開発等，研究・教育が一体となり，学術的な新しい価値の創出や新たなビジネスを生み出すことも視野に入れた産学連携の仕組みを構築することである。この理念の下みずほ銀行と早稲田大学データ科学センター（以下 CDS）の協働プロジェクトとして，早稲田大学内で共通的にみずほ銀行データを分析・活用するための産学連携研究プラットフォーム（以下銀行データ活用プラットフォーム）の構築を行なった。なお CDS は「高度なデータ分析スキル」を持った研究者と様々な領域における「深い専門知識」を持った研究者の融合による，データ駆動型の最先端の研究を推進するとともに，専門知識とデータ分析スキルの両方を兼ね備えた実践的な人材の育成を目的として 2017 年 12 月に早稲田大学に設置された組織である[2,3]。なお実際にこのプラットフォームを活用したいくつかの研究プロジェクトが実施されており，例えばその成果の一つとして文献[4]では新型コロナウイルス感染症緊急経済対策のための特別定額給付金に関してみずほ銀行データを分析した結果が報告されている。

本稿ではこの銀行データ活用プラットフォームに焦点をあて，その理念と必要性，具体的なプラットフォームの仕組みと運用方法及び評価を示す。

2. データサイエンスと産学連携研究

情報通信技術の発展及び統計学や機械学習などデータの分析技術の発展により，様々なデータが収集・分析・活用されるようになってきている。ここで大学を考えてみると，様々な学術領域でデータを利用した研究・教育が盛んに行われるようになっており，教員及び研究者の間で研究・

教育に利用するデータの必要性が高まっている。一方で実社会においてデータを収集・保有できるのは企業であることが多く，そのような大規模かつ貴重なデータを大学の教員あるいは研究者が個人的に入手して分析に利用することは難しい場合が多い。一方企業側において，データを分析・活用できる人材を十分に確保できるのは一部の企業に限られており，多くの企業では収集したデータをどのように事業活動の発展に活用していくかが課題となっている。このような課題を解決するために，共同研究等によって企業が大学の研究者にデータを提供し，それを利用して得られた研究成果を企業や社会にフィードバックする動きがある[5]。このような状況において，CDS はデータの提供を希望・許諾する企業と，データを活用した研究・教育を行いたい大学の教員・研究者との間に立ち，データ活用に関する契約や相談，データの保持や提供等々を行うサービス WIRP を実施している[6]。

3. 銀行データ活用プラットフォーム

3.1 構築したシステムの概要

前節のようなデータサイエンスにおける産学連携研究の重要性が高まっている中，2018 年からみずほ銀行データを早稲田大学の中で共通的に活用するためのプラットフォーム（図 1 参照）の構築を行ってきた。本節では主に構築したシステムについて説明する。

システムの構成要素は大別して 1) みずほ銀行と CDS の間に位置する「Cloud 分析環境」，2) CDS 内に設置している「データアクセス室」2 部屋，3) データアクセス室内で利用する「データ分析 PC」である。「Cloud 分析環境」には分析用の銀行データがアップロードされる。早稲田大学側の分析を実施する利用者（以下利用者）は，「データアクセス室」に来訪し，「データアクセス室」において分析専用の「データ分析 PC」を用いて「Cloud 分析環境」にアクセスし，データ分析を実施する。ここで「データアクセス室」と「Cloud 分析環境」は専用線で繋いでおり，「データ分析 PC」は「Cloud 分析環境」のみとしか接続できない設定としている。また「データ分析 PC」は管理者以外がファイルを取り出すことができない設定にしている。さらに不正を防止することを目的として，各データアクセス室には

Construction of academia-industry collaborative research platform to utilize bank data

[†]Manabu Kobayashi, Yuta Nakahara, Waseda University

[‡]Eriko Sato, Chikanobu Takaya, Yoshihiro Tsukamoto, Mizuho Bank

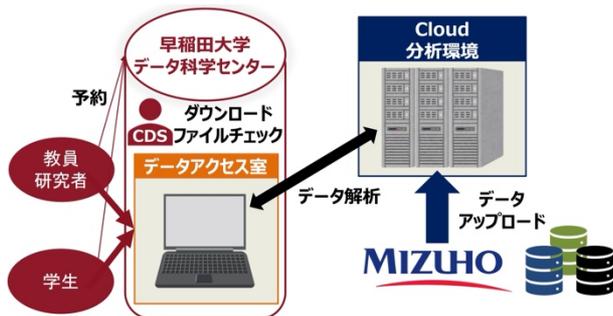


図1：銀行データ活用プラットフォーム



図2：データアクセス室

監視カメラが設置されている。「Cloud 分析環境」における分析は、RやPythonの他、必要に応じてより高度な分析を実施する場合のために IBM Watson Studio も利用可能な設計になっている。また「データ分析 PC」においてはExcelの他、必要に応じて STATA 等の分析ソフトウェアを利用可能にしている。

3.2 プラットフォームの運用

まず利用者は CDS 窓口に銀行データ利用の相談の申請を行い、その後みずほ銀行担当者、CDS 担当者を含めた三者で、研究目的やデータの種類、具体的な分析方法を含めて綿密な打合せを実施する。その後研究プロジェクトの申請が行われ、問題がなければ正式に承認と契約が行われる。次のステップにおいて利用者並びに研究プロジェクトごとに Cloud 上にアカウントや個別の分析環境が用意される。また承認された銀行データが「Cloud 分析環境」にアップロードされる。ここで全データは個人情報など個人を特定できる情報が全てマスキングされている点に注意されたい。ここまでが大まかな準備である。

次に利用者が「データアクセス室」並びに「データ分析 PC」を利用する場合、予め CDS 担当者に利用日時と利用するプロジェクトメンバーを伝え、承認を得る。利用日時に受付を行った後「データアクセス室」において分析を実施する。図2に「データアクセス室」を示す。利用が終わった後に再び受付において利用終了の報告を行う。

なお、利用者が分析のためのプログラムや関連データのアップロードを行いたい場合や、分析後のファイルのダウンロードを行いたい場合には CDS に申請を行い、規約に反していないかどうかファイル内容を管理者が確認した上で、管理者がファイルのアップロードやダウンロードを実施する。これは万一にもインシデントが起こらないようにするための予防措置である。

3.3 プラットフォームの評価

本プラットフォームは貴重な銀行データを十全

に分析・活用できる仕組みであることと、厳しいセキュリティ要件を満たすことに主眼を置いている。その上で他に類を見ない銀行データの分析事例が成果として出はじめている[4]。一方で既に複数の研究プロジェクトが同時進行するに当たり、データアクセス室の利用人数の制限が問題になりつつある。本課題については、邦銀初のデータビジネスである統計データ販売サービス Mizuho Insight Portal (Mi-Pot/ミーポット)[7]の利用も検討を行う予定である。

4. おわりに

本稿で構築したプラットフォームでは厳格な運用を実施しているが、これは非常に貴重かつ重要な銀行データに万一のことがあってはならないことを大前提とし、その上で銀行データを用いることにより種々の学術的価値の非常に高い分析が実施できること、ビジネス価値の創造並びに社会課題の解決へ向けた社会貢献など様々な目的を達成するためであり、本プラットフォームにより今後も種々の産学連携の一助になることを期待する。

謝辞

本研究は JPSP 科研費 JP19K04914 の助成を受けたものです。

参考文献

- [1] https://www.mizuhodobank.co.jp/release/pdf/20180713release_jp.pdf
- [2] 須子統太, 小林学, 堀井俊佑, 安田豪毅, 松嶋敏泰, "早稲田大学におけるデータ科学教育の取り組み", 日本経営工学会秋季大会予稿集, 2018.
- [3] <https://www.waseda.jp/inst/cds/>
- [4] So Kubota, Koichiro Onishi, Yuta Toyama, "Consumption responses to COVID-19 payments: Evidence from a natural experiment and bank account data", *Journal of Economic Behavior and Organization*, vol.188, pp.1-17, Aug.2021.
- [5] 大山敬三, 大須賀智子, 国立情報学研究所における研究用データセットの共同利用, *情報管理*, Vol. 59, No. 2, pp. 105-112, 2016.
- [6] 安田豪毅, 中原悠太, 野村亮, 小林学, 須子統太, "早稲田大学の研究・教育用データ提供プラットフォーム WIRP について", 日本経営工学会春季大会予稿集, 2021.
- [7] https://www.mizuhodobank.co.jp/release/pdf/20201110_2release_jp.pdf